Ergebnisse der Mathematik und ihrer Grenzgebiete.
3. Folge / A Series of Modern Surveys in Mathematics 63

Tuomas Hytönen
Jan van Neerven
Mark Veraar
Lutz Weis

Analysis in Banach Spaces

Volume I: Martingales and Littlewood-Paley Theory

Springer

Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics

Volume 63

Ergebnisse der Mathematik und ihrer Grenzgebiete, now in its third sequence, aims to provide summary reports, on a high level, on important topics of mathematical research. Each book is designed as a reliable reference covering a significant area of advanced mathematics, spelling out related open questions, and incorporating a comprehensive, up-to-date bibliography.

More information about this series at http://www.springer.com/series/728

The four authors during a writing session in Oberwolfach in November 2013.
Left to right: Mark Veraar, Lutz Weis, Tuomas Hytönen, Jan van Neerven

Tuomas Hytönen · Jan van Neerven
Mark Veraar · Lutz Weis

Analysis in Banach Spaces

Volume I: Martingales and Littlewood-Paley Theory

Springer

Tuomas Hytönen
Department of Mathematics and Statistics
University of Helsinki
Helsinki
Finland

Mark Veraar
Delft Institute of Applied Mathematics
Delft University of Technology
Delft
The Netherlands

Jan van Neerven
Delft Institute of Applied Mathematics
Delft University of Technology
Delft
The Netherlands

Lutz Weis
Department of Mathematics
Karlsruhe Institute of Technology
Karlsruhe
Germany

ISSN 0071-1136 ISSN 2197-5655 (electronic)
Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics
ISBN 978-3-319-48519-5 ISBN 978-3-319-48520-1 (eBook)
DOI 10.1007/978-3-319-48520-1

Library of Congress Control Number: 2016955329

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Over the past fifteen years, motivated by regularity problems in evolution equations, there has been tremendous progress in the analysis of Banach space-valued functions and processes. For so-called UMD spaces in particular, central areas of harmonic analysis, such as the theory of Fourier multipliers and singular integrals, were extended to operator-valued kernels acting on Bochner spaces, and basic estimates of stochastic analysis, including the Itô isometry and the Burkholder–Davis–Gundy inequalities, were generalised to Banach space-valued processes.

As it was long known that extensions of such sophisticated scalar-valued estimates are not possible for all Banach spaces, these results depended on essential progress in the geometry of Banach spaces during the 70s and 80s. The theory of Burkholder and Bourgain on UMD spaces became the foundation on which the recent theory we wish to report on was built; just as important are results of Kwapień, Maurey, and Pisier on type and cotype, since they link the structure of the Banach space to estimates for random sums which replace to some extent the fundamental orthogonality relations in Hilbert spaces.

For most classical Banach spaces, the UMD, type and cotype properties are readily available and therefore the results of vector-valued Analysis can be applied to many situations of interest in the theory of partial differential equations; they have already proved their value by providing sharp regularity estimates for parabolic problems. Our aim is to give a detailed and careful presentation of these topics that is useful not only as a reference book but can be used also selectively as a basis for advanced courses and seminars.

This project ranges over a broad spectrum of Analysis and includes Banach space theory, operator theory, harmonic analysis and stochastic analysis. For this reason we have divided it into three parts. The present volume develops the theory of Bochner integration, Banach space-valued martingales and UMD spaces, and culminates in a treatment of the Hilbert transform, Littlewood–Paley theory and the vector-valued Mihlin multiplier theorem.

Volume II will present a thorough study of the basic randomisation techniques and the operator-theoretic aspects of the theory, such as R-

boundedness, vector-valued square functions and radonifying operators, as well as a detailed treatment of the relevant probabilistic Banach space notions such as type, cotype, K-convexity and properties related to contraction principles. These techniques will allow us to present the theory of H^∞-functional calculus for sectorial operators and work out the main examples. This sets the stage for our final aim, a presentation of the theory of singular integral operators with operator-valued kernels and its applications to maximal regularity for deterministic and stochastic parabolic evolution equations, which will be the subject matter of Volume III.

The central theme in all volumes is the identification of the Banach spaces to which the key estimates of classical harmonic and stochastic analysis can be extended as those with the fundamental UMD property. The very definition behind this abbreviation is the unconditionality of martingale differences, a primarily probabilistic notion, and a number of different characterisations are formulated in purely probabilistic terms. However, this same property is also equivalent to the boundedness of the vector-valued Hilbert transform, the Littlewood–Paley inequality for vector-valued Fourier integrals, and several other estimates in the realm of classical harmonic analysis.

Each of these aspects of UMD spaces makes a substantial body of theory in its own right, and one could certainly produce respectable treatments of large parts of this material with a "clean" probabilistic or analytic flavour. However, rather than striving for such "purity", our aim is to emphasise the rich connections between the two worlds and the unity of the subject. For example, while martingales are traditionally regarded as a topic in Probability, we define and discuss them on σ-finite measure spaces from the beginning, so that they are immediately applicable to Analysis on the Euclidean space $\mathbb{R}^d$ without the need of auxiliary truncations or decompositions into probability spaces. Moreover, it is important to observe that even if we (or the reader) wanted to concentrate on the analytic side of UMD spaces only, we could hardly present a complete picture without an occasional reference to the probabilistic notions, at least at the present state of knowledge. For instance, although we know that both the Hilbert transform boundedness and the Littlewood–Paley inequality are equivalent to the UMD property, and therefore to each other, the only known way of proving the equivalence of these two analytic notions passes through the probabilistic UMD. There are numerous other such examples, and new frontiers of the theory have shown over and over again that it is the probabilistic definition of UMD spaces that lies at the centre and connects everything together.

So much said about the unity of Analysis and Probability (in Banach spaces), we should acknowledge the existence of a third side of the triangle, which is barely touched by the present treatise, namely: Geometry (of Banach spaces). Our choice of topics is not meant in any way to downplay the importance of this huge topic, both in its own right and in relation to analytic and probabilistic questions, but rather to admit our limits and to leave the proper account of the geometric connections for other treatments.

*

This book can be studied in a variety of ways and for different motivations. The principal, but not the only, audience that we have in mind consists of researchers who need and use Analysis in Banach spaces as a tool for studying other problems, in particular the regularity of evolution equations mentioned above. Until now, the contents of this extensive and powerful toolbox have been mostly scattered around in research papers, or in some cases monographs addressed to readerships with a rather different background from our focus, and we feel that collecting this diverse body of material into a unified and accessible presentation fills a gap in the existing literature. Indeed, we regard ourselves as part of this audience, and we have written the kind of book that we would have liked to have for ourselves when working through this theory for the first time.

Aside from this, parts of the book may also offer an interesting angle to the classical analysis of scalar-valued functions, which is certainly covered as a special case, and seldom required as a prerequisite or used as a building block for the Banach space-valued theory. For a classical harmonic analyst, the approach that we take, say, to the L^p-boundedness of the Hilbert transform, is possibly exotic, but not necessarily substantially more difficult than more traditional treatments in the scalar-valued case.

*

There are a couple of technical features of the book worth mentioning. Most of the time, we are quite explicit with the constants appearing in our estimates, and we especially try to keep track of the dependence on the main parameters involved. Thus, rather than saying that a particular bound "only depends on the UMD constant $\beta_{p,X}$", we prefer to write out, say, $(\beta_{p,X})^2$, or whichever function of $\beta_{p,X}$ appears from the calculation. We often go to the extent of writing, say, "2000" instead of "c, where c is a numerical constant", although we also might write "2000" instead of "1764", when there is no reason to believe that the latter constant, although given by a particular computation, would be anywhere close to optimal. Indeed, except for a few select places, we make no claim that our explicit constants cannot be improved; however, in many places, we have made an effort to present the best (order of) bounds currently available by the existing methods. We hope that making this explicit documentation might spur some interest towards research on such quantitative issues.

We also pay more attention than many texts to the impact of the underlying scalar field (real or complex) on the results under consideration. While this is largely irrelevant for many questions, it does play a role in some others, and we try to be quite explicit in pointing out the differences when they do occur, hopefully without insisting too much on this point when they do not.

*

This project was initiated in Delft and Karlsruhe already in 2008. Critical to its eventual progress was the possibility of intensive joint working periods in the serenity provided by the Banach Center in Będlewo (2012), Mathematisches Forschungsinstitut Oberwolfach (2013), Stiftsgut Keysermühle in Klingenmünster (2014 and 2015) and Hotel ’t Paviljoen in Rhenen (2015). All four of us also met twice in Helsinki (2014 and 2016), and a number of additional working sessions were held by subgroups of the author team. One of us (J.v.N.) wishes to thank Marta Sanz-Solé for her hospitality during a sabbatical leave at the University of Barcelona in 2013.

Preliminary versions of parts of the material were presented in advanced courses and lecture series at various international venues and in seminars at our departments, and we would like to thank the students and colleagues who attended these events for feedback that shaped and improved the final form of the text. Special thanks go to Jamil Abreu, Alex Amenta, Markus Antoni, Sonja Cox, Chiara Gallarati, Fabian Hornung, Luca Hornung, Marcel de Jeu, Marcel Kreuter, Nick Lindemulder, Emiel Lorist, Bas Nieraeth, Jan Rozendaal, Jonas Teuwen, and Ivan Yaroslavtsev who did detailed reading of portions of this book. Needless to say, we take full responsibility for any remaining errors. A list with errata will be maintained on our personal websites. We wish to thank Klaas Pieter Hart for LaTeX support.

During the writing of this book, we have benefited from external funding by the European Research Council (ERC Starting Grant “AnProb” to T.H.), the Academy of Finland (grants 130166 and 133264 to T.H., and the Centre of Excellence in Analysis and Dynamics, of which T.H. is a member), the Netherlands Organisation for Scientific research (NWO) (VIDI grant 639.032.201 and VICI grant 639.033.604 to J.v.N. and VENI grant 639.031.930 and VIDI grant 639.032.427 to M.V.), the Volkswagenstiftung (grant I/78593 to L.W.) and the Deutsche Forschungsgemeinschaft (grant We 2847/1-2 to L.W.). We also wish to thank the Banach Center in Będlewo and Mathematisches Forschungsinstitut Oberwolfach for allowing us to spend two highly productive weeks in both wonderful locations.

Delft, Helsinki and Karlsruhe,
November 14, 2016.

Contents

Symbols and notations

Sets

$\mathbb{N} = \{0, 1, 2, \ldots\}$ - non-negative integers
$\mathbb{Z}$ - integers
$\mathbb{Q}$ - rational numbers
$\mathbb{R}$ - real numbers
$\mathbb{C}$ - complex numbers
$\mathbb{K}$ - scalar field ($\mathbb{R}$ or $\mathbb{C}$)
$\mathbb{S} = \{z \in \mathbb{C} : 0 < \Im z < 1\}$ - unit strip
$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ - unit circle
$\overline{\mathbb{Z}} = \mathbb{Z} \cup \{-\infty, \infty\}$ - extended integers
$\mathbb{R}_+ = (0, \infty)$ - positive real line
B_X - open unit ball
S_X - unit sphere
$B(x, r)$ - open ball centred at x with radius r

Vector spaces

BMO - space of functions of bounded mean oscillation
c_0 - space of null sequences
C - space of continuous functions
C_0 - space of continuous functions vanishing at infinity
C_{b} - space of bounded continuous functions
C_{c} - space of continuous functions with compact support
C_{c}^{∞} - space of test functions with compact support
$\mathscr{C}^p$ - Schatten class
H - Hilbert space
$H^{s,p}$ - Bessel potential space
H^p - Hardy space
$\mathscr{H}(X_0, X_1)$ - space of homomorphic functions on the strip
ℓ^p - space of p-summable sequences

ℓ^p_N - space of p-summable finite sequences
L^p - Lebesgue space
$L^{p,\infty}$ - weak-L^p
$\mathscr{L}(X,Y)$ - space of bounded linear operators
$\mathfrak{M}L^p(\mathbb{R}^d;X,Y)$ - space of Fourier multipliers
$\mathfrak{M}(\mathbb{R}^d;X,Y)$ - Mihlin class
$\mathscr{S}$ - space of Schwartz functions
$\mathscr{S}'$ - space of tempered distributions
$W^{k,p}$ - Sobolev space
X, Y, $\dots$- Banach spaces
$X_{\mathbb{C}}$ - complexification
$X^{\gamma,p}_{\mathbb{C}}$ - Gaussian complexification
X^*, Y^*, $\dots$- dual Banach spaces
$X \otimes Y$ - tensor product
$[X_0,X_1]_\theta$ - complex interpolation space
$(X_0,X_1)_{\theta,p}$, $(X_0,X_1)_{\theta,p_0,p_1}$ - real interpolation spaces

Measure theory and probability

$\mathscr{A}$ - σ-algebra
$df_n = f_n - f_{n-1}$ - nth martingale difference
ϵ_n - signs in $\mathbb{K}$, i.e., scalars in $\mathbb{K}$ of modulus one
ε_n - Rademacher variables with values in $\mathbb{K}$
$\mathbb{E}$ - expectation
$\mathscr{F}$, $\mathscr{G}$, $\dots$- σ-algebras
$\mathscr{F}_f$ - collection of sets in $\mathscr{F}$ on which f is integrable
$\mathbb{E}(\cdot|\cdot)$ - conditional expectation
${}^\tau f_n = f_n - f_{\tau\wedge n}$ - started martingale
$f^\tau_n = f_{\tau\wedge n}$ - stopped martingale
$f^\star_n = \sup_{k\leqslant n}\|f_n\|$ - maximal function
γ - Gaussian variable
h_I - Haar function
μ - measure
$\|\mu\|$ - variation of a measure
$(\Omega,\mathscr{A},\mathbb{P})$ - probability space
$\mathbb{P}$ - probability measure
$\mathbb{P}(\cdot|\cdot)$ - conditional probability
r_n - real Rademacher variables
$(S,\mathscr{A},\mu)$ - measure space
$\sigma(f,g,\dots)$ - σ-algebra generated by the functions $f,g,\dots$
$\sigma(\mathscr{C})$ - σ-algebra generated by the collection $\mathscr{C}$
τ - stopping time
w_α - Walsh function

Norms and pairings

$|\cdot|$ - modulus, Euclidean norm
$\|\cdot\| = \|\cdot\|_X$ - norm in a Banach space X
$\|\cdot\|_p = \|\cdot\|_{L^p}$ - L^p-norm
$\langle\cdot,\cdot\rangle$ - duality
$(\cdot|\cdot)$ - inner product in a Hilbert space
$a \cdot b$ - inner product of $a, b \in \mathbb{R}^d$

Operators

D_j - pre-decomposition
$\varDelta$ - Laplace operator
$\mathscr{D}$ - dyadic system
$\partial_j = \partial/\partial x_j$ - partial derivative with respect to x_j
∂^α - partial derivative with multi-index α
$\mathbb{E}(\cdot|\cdot)$ - conditional expectation
$\mathscr{F} f$ - Fourier transform
$\mathscr{F}^{-1} f$ - inverse Fourier transform
H - Hilbert transform
$\widetilde{H}$ - periodic Hilbert transform
J_s - Bessel potential operator
$\mathscr{L}(X,Y)$ - space of bounded operators from X to Y
$\mathscr{L}_{\rm so}(X,Y)$ - idem, endowed with the strong operator topology
$\mathscr{R}_p$ - R-bound
R_j - jth Riesz transform
S, T, $\dots$- bounded linear operators
T^* - adjoint of the operator T
T_m - Fourier multiplier operator associated with multiplier m
T_v - martingale transform associated with predictable sequence v
$T \otimes I_X$ - tensor extension of T

Constants and inequalities

$\beta_{p,X}$ - UMD constant
$\beta^{\mathbb{R}}_{p,X}$ - UMD constant with signs ± 1
$\beta^{\pm}_{p,X}$ - upper and lower randomised UMD constant
$c_{q,X}$ - cotype q constant
$c^{\rm mart}_{q,X}$ - martingale cotype q constant
$\hbar_{p,X}$ - norm of the Hilbert transform on $L^p(\mathbb{R};X)$
$K_{p,X}$ - K-convexity constant
$\kappa_{p,q}$ - Kahane–Khintchine constant
$\tau_{p,X}$ - type p constant
$\tau^{\rm mart}_{p,X}$ - martingale type p constant
$\varphi_{p,X}(\mathbb{R}^d)$ - norm of the Fourier transform $\mathscr{F} : L^p(\mathbb{R}^d;X) \to L^{p'}(\mathbb{R}^d;X)$.

Miscellaneous

$\hookrightarrow$ - continuous embedding
$\mathbf{1}_A$ - indicator function
$a \lesssim b$ - $\exists C$ such that $a \leqslant Cb$
$a \lesssim_{p,P} b$ - $\exists C$, depending on p and P, such that $a \leqslant Cb$
C - generic constant
$\complement$ - complement
$d(x, y)$ - distance
$f^\star$ - maximal function
$\widetilde{f}$ - reflected function
$\widehat{f}$ - Fourier transform
$\check{f}$ - inverse Fourier transform
$f * g$ - convolution
$\Im$ - imaginary part
$K(t, x) = K(t, x; X_0, X_1)$ - K-functional
Mf - Hardy–Littlewood maximal function
$M_{\mathrm{Rad}}f$ - Rademacher maximal function
$p' = p/(p-1)$ - conjugate exponent
$p^* = \max\{p, p'\}$
$\Re$ - real part
$s \wedge t = \min\{s, t\}$
$s \vee t = \max\{s, t\}$
$v \star f$ - martingale transform of f by v
x - generic element of X
x^* - generic element of X^*
$x \otimes y$ - elementary tensor

Standing assumptions

Throughout this book, two of conventions will be in force.

1. Unless stated otherwise, the scalar field $\mathbb{K}$ can be real or complex. Results which do not explicitly specify the scalar field to be real or complex are true over both the real and complex scalars.

2. In the context of randomisation, a *Rademacher variable* is a uniformly distributed random variable taking values in the set $\{z \in \mathbb{K} : |z| = 1\}$. Such variables are denoted by the letter ε. Thus, whenever we work over $\mathbb{R}$ it is understood that ε is a real Rademacher variable, i.e.,

$$\mathbb{P}(\varepsilon = 1) = \mathbb{P}(\varepsilon = -1) = \frac{1}{2},$$

and whenever we work over $\mathbb{C}$ it is understood that ε is a complex Rademacher variable (also called a *Steinhaus variable*), i.e.,

$$\mathbb{P}(a < \arg(\varepsilon) < b) = \frac{1}{2\pi}(b - a).$$

Occasionally we need to use real Rademacher variables when working over the complex scalars. In those instances we will always denote these with the letter r. Similar conventions are in force with respect to Gaussian random variables: a *Gaussian random variable* is a standard normal real-valued variable when working over $\mathbb{R}$ and a standard normal complex-valued variable when working over $\mathbb{C}$.

1

Bochner spaces

In this first chapter we present the essentials of the integration theory of Banach space-valued functions. We begin by exploring the various possible definitions of measurability for such functions. It turns out that for separable Banach spaces X and measurable spaces $(S, \mathscr{A})$, a function $f : S \to X$ is measurable—in the sense that the pre-images

$$f^{-1}(B) := \{f \in B\} := \{s \in S : f(s) \in B\}$$

are measurable for every Borel set B in X—if and only if the scalar-valued function $\langle f, x^* \rangle$ is measurable for every functional x^* in the dual space X^*. This is essentially the content of the Pettis measurability theorem, which is proved in Section 1.1.

In Section 1.2 we proceed with the construction of the Bochner integral, which is the analogue of the Lebesgue integral for X-valued functions. It preserves all essential aspects of the Lebesgue integral, such as the availability of approximation results, convergence theorems and Fubini's theorem. The Banach spaces $L^p(S; X)$ of Bochner integrable functions provide the basic functional framework of our work. However, occasionally we shall also need the Pettis integral, which is defined in terms of the Lebesgue integrals of the functions $\langle f, x^* \rangle$.

For the most complete statement of the duality theory of the Bochner spaces $L^p(S; X)$ in Section 1.3 we have to restrict ourselves to the class of Banach spaces for which an analogue of the Radon–Nikodým theorem holds. This class is closely connected with the almost-everywhere differentiability of Lipschitz and absolutely continuous functions with values in X, a topic that will be discussed in the next chapter.

Throughout this book, X is a Banach space over the scalar field $\mathbb{K}$ which may be either $\mathbb{R}$ or $\mathbb{C}$. If we wish to emphasise a particular choice of scalar field we will speak of *real* and *complex* Banach spaces. It is often useful to note that a complex Banach space is also a real Banach space by restricting the complex scalar multiplication to the real numbers.

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 63, DOI 10.1007/978-3-319-48520-1_1

The norm of an element $x \in X$ is denoted by $\|x\|_X$, or, if no confusion can arise, by $\|x\|$. The Banach space dual of X is denoted by X^*. We shall use the notation $\langle x, x^* \rangle := x^*(x)$ to denote the duality pairing of the elements $x \in X$ and $x^* \in X^*$.

1.1 Measurability

In the context of Analysis in Banach spaces, several natural notions of measurability present themselves, such as measurability, strong measurability and weak measurability. In finite dimensions all three coincide, but in infinite dimensions they do not and our first task is to understand the way they are interrelated. The main result in this direction, and indeed one of the cornerstones of the theory, is the Pettis measurability theorem. It asserts that a function with values in a Banach space X is strongly measurable if and only if it is separably valued and weakly measurable. We shall present two versions of this result: a pointwise version for functions defined on a measurable space $(S, \mathscr{A})$ and a μ-almost everywhere version for functions defined on a measure space $(S, \mathscr{A}, \mu)$. In general we shall make some effort to present the results for arbitrary measure spaces, avoiding assumptions such as σ-finiteness whenever this is possible.

1.1.a Functions on a measurable space $(S, \mathscr{A})$

Measurability

The first definition of measurability for Banach space-valued functions that comes to mind is that of inverse images: a function with values in a Banach space X is said to be *measurable* if the pre-image $f^{-1}(B)$ of every Borel set B in X is measurable. As it turns out, in many respects this natural notion is not as useful as one might think, the reason being that the Borel σ-algebra $\mathscr{B}(X)$ is in general 'too large'. In fact, the σ-algebra generated by all continuous linear functionals on X may be strictly smaller than $\mathscr{B}(X)$. This presents an obstruction to applying the standard tools of functional analysis such as the Hahn–Banach theorem in an effective way.

Our first objective is to prove that if X is a *separable* Banach space, this problem does not occur. Given a subset Y of the dual space X^* we denote by $\sigma(Y)$ the σ-algebra generated by Y, i.e., the smallest σ-algebra in X for which every $x^* \in Y$ measurable. It is easy to see that $\sigma(Y)$ is generated by the collection $\mathscr{C}(Y)$ of all sets of the form

$$\big\{x \in X : \ \big(\langle x, x_1^* \rangle, \ldots, \langle x, x_n^* \rangle\big) \in B\big\}$$

with $n \geqslant 1$, $x_1^*, \ldots, x_n^* \in Y$ and $B \in \mathscr{B}(\mathbb{K}^n)$.

Recall that a linear subspace Y of X^* is dense with respect to the weak* topology of X (which, by definition, is the smallest topology in X^* for which the mapping $x^* \mapsto \langle x, x^*\rangle$ continuous for every $x \in X$) if and only if Y separates the points of X. We refer the reader to Appendix B for some background material on the weak* topology.

Proposition 1.1.1. *If X is separable and Y is a weak* dense linear subspace of X^*, then*

$$\sigma(Y) = \sigma(X^*) = \mathscr{B}(X).$$

Proof. Let G denote the set of all $x^* \in X^*$ having the property that the function $x \mapsto \langle x, x^*\rangle$ is $\sigma(Y)$-measurable. Then G is a linear subspace of X^* containing Y. Moreover, G is weak* sequentially closed. Therefore, $G = X^*$ by a corollary to the Krein–Šmulian theorem (Corollary B.1.14), which means that $x \mapsto \langle x, x^*\rangle$ is $\sigma(Y)$-measurable for all $x^* \in X^*$.

Now let $n \geqslant 1$, $x_1^*, \dots, x_n^* \in X^*$ and $B_1, \dots, B_n \in \mathscr{B}(\mathbb{K})$ be given. Put $B := B_1 \times \cdots \times B_n$. Then $B \in \mathscr{B}(\mathbb{K}^n)$ and the set

$$\big\{x \in X : \ \big(\langle x, x_1^*\rangle, \dots, \langle x, x_n^*\rangle\big) \in B\big\} = \bigcap_{k=1}^{n} \big\{x \in X : \ \langle x, x_k^*\rangle \in B_k\big\}$$

belongs to $\sigma(Y)$. Denote by Σ the collection of all $B \in \mathscr{B}(\mathbb{K}^n)$ having the property that

$$\big\{x \in X : \ \big(\langle x, x_1^*\rangle, \dots, \langle x, x_n^*\rangle\big) \in B\big\} \in \sigma(Y). \tag{1.1}$$

Then Σ is a σ-algebra in $\mathbb{K}^n$, and by the observation just made it contains all Borel rectangles $B_1 \times \dots, \times B_n$. Therefore $\mathscr{B}(\mathbb{K}^n) \subseteq \Sigma$.

We have shown that (1.1) holds for all finite sets $x_1^*, \dots, x_n^* \in X^*$ and all Borel sets $B \in \mathscr{B}(\mathbb{K}^n)$. It follows that $\sigma(X^*) \subseteq \sigma(Y)$. Since the reverse inclusion holds trivially, it follows that $\sigma(X^*) = \sigma(Y)$. It remains to be shown that $\sigma(X^*) = \mathscr{B}(X)$. Since every open set is the countable union of open balls and every open ball is a countable union of closed balls, it is enough to show that every closed ball $B(x_0, r) := \{x \in X : \ \|x - x_0\| \leqslant r\}$ belongs to $\sigma(X^*)$. Choose a norming sequence of unit vectors $(x_n^*)_{n\geqslant 1}$ in X^*. Then

$$B(x_0, r) = \big\{x \in X : \ \|x - x_0\| \leqslant r\big\} = \bigcap_{n\geqslant 1} \big\{x \in X : \ |\langle x - x_0, x_n^*\rangle| \leqslant r\big\},$$

and this set belongs to $\sigma(X^*)$. This completes the proof. □

When X is non-separable, strict inclusion $\sigma(X^*) \subsetneq \mathscr{B}(X)$ may indeed occur; this phenomenon will be discussed in the Notes at the end of the chapter.

Corollary 1.1.2. *If X is separable, then for a function $f : S \to X$ the following assertions are equivalent:*

(1) *f is measurable;*
(2) *$\langle f, x^*\rangle$ is measurable for all $x^* \in X^*$.*

Indeed, if (2) holds, then with the notations introduced in the above proof,

$$f^{-1}(B(x_0,r)) = \bigcap_{n\geqslant 1} \big\{s\in S:\ |\langle f(s)-x_0, x_n^*\rangle| \leqslant r\big\} \in \mathscr{A}.$$

Since the balls $B(x_0,r)$ generate $\mathscr{B}(X)$, this proves that f is measurable.

Strong measurability

The essential feature used in the construction of the Lebesgue integral for scalar-valued functions is that every measurable function can be approximated pointwise by a sequence of simple functions. Since, in the converse direction, pointwise limits of measurable functions are measurable, this suggests to tie up the notion of measurability with approximation by simple functions. This is precisely the idea taken up in the definition of *strong measurability.*

As before we let $(S,\mathscr{A})$ be a measurable space.

Definition 1.1.3. *A function $f: S\to X$ is called* simple *if it is of the form $f=\sum_{n=1}^N \mathbf{1}_{A_n}\otimes x_n$ with $A_n\in\mathscr{A}$ and $x_n\in X$ for all $1\leqslant n\leqslant N$.*

Here $\mathbf{1}_A$ denotes the indicator function of the set A and we use the notation

$$(f\otimes x)(s) := f(s)x$$

for functions $f: S\to\mathbb{K}$ and elements $x\in X$. We also define

$$F\otimes X := \Big\{\sum_{n=1}^N f_n\otimes x_n : f_n\in F,\ x_n\in X,\ n=1,\dots,N;\ N=1,2,\dots\Big\},$$

whenever F is a vector space of scalar-valued functions.

Definition 1.1.4. *A function $f: S\to X$ is* strongly measurable *if there exists a sequence of simple functions $f_n: S\to X$ such that* $\lim_{n\to\infty} f_n = f$ *pointwise on S.*

If we wish to emphasise the underlying σ-algebra we shall speak of a *strongly $\mathscr{A}$-measurable* function.

We shall see below that if X is separable, then an X-valued function f is strongly measurable if and only if it is measurable. The next example shows that the word 'separable' cannot be omitted from this statement.

Example 1.1.5. For any non-separable Banach space X, the identity map $I: X\to X$ is continuous and hence measurable, but it fails to be strongly measurable. Suppose, for a contradiction, that $I_n: X\to X$ is a sequence of simple Borel functions converging to I pointwise. Let V be the countable set of values taken by these functions. Then every $x\in X$ is the limit of a sequence in V, which implies that X is separable. By the same argument, I even fails to be strongly $\mathscr{P}(X)$-measurable (with $\mathscr{P}(X)$ the power set of X).

A function $f : S \to X$ is called *separably valued* if there exists a separable closed subspace $X_0 \subseteq X$ such that $f(s) \in X_0$ for all $s \in S$. The function f is called *weakly measurable* if

$$s \mapsto \langle f, x^* \rangle(s) := \langle f(s), x^* \rangle,$$

is measurable for all $x^* \in X^*$.

Theorem 1.1.6 (Pettis measurability theorem, first version). *Let $(S, \mathscr{A})$ be a measurable space and let Y be a weak* dense subspace of X^*. For a function $f : S \to X$ the following assertions are equivalent:*

(1) *f is strongly measurable;*
(2) *f is separably valued and weakly measurable;*
(3) *f is separably valued and $\langle f, x^* \rangle$ is measurable for all $x^* \in Y$.*

Moreover, if f takes its values in a closed linear subspace X_0 of X, then f is the pointwise limit of a sequence of X_0-valued simple functions.

Proof. (1)⇒(2): Let $(f_n)_{n \geqslant 1}$ be a sequence of simple functions converging to f pointwise and let X_1 be the closed subspace spanned by the countably many values taken by these functions. Then X_1 is separable and f takes its values in X_1. Furthermore, each $\langle f, x^* \rangle$ is measurable, being the pointwise limit of the measurable functions $\langle f_n, x^* \rangle$.

(2)⇒(3): This implication holds trivially.

(3)⇒(2): Let X_1 be a separable closed subspace of X with the property that $f(s) \in X_1$ for all $s \in S$, and let $j_1 : X_1 \subseteq X$ be the inclusion mapping. Then for all $x^* \in X^*$ we have $\langle f, x^* \rangle = \langle f, j_1^* x^* \rangle$, where on the left we regard f as an X-valued function and on the right as an X_1-valued function. Thus is suffices to prove that $\langle f, x_1^* \rangle$ is measurable for every $x_1^* \in X_1^*$.

Let Y_1 be the subspace in X_1^* consisting of all $x_1^* \in X_1^*$ for which $\langle f, x_1^* \rangle$ is measurable. Then $j_1^*(Y) \subseteq Y_1$. Since j_1^*, being an adjoint operator, is weak* continuous, $j_1^*(Y)$ is weak* dense in X_1^*. Therefore the same is true for Y_1. Also, Y_1 is weak* sequentially closed in X_1^*. Hence by a corollary to the Krein–Šmulian theorem (Corollary B.1.14), $Y_1 = X_1^*$.

(2)⇒(1): Choose a sequence $(x_n^*)_{n \geqslant 1}$ of unit vectors in X^* that is norming for a separable closed subspace X_1 of X where f takes its values. By the weak measurability of f, for each $x \in X_1$ the real-valued function

$$s \mapsto \|f(s) - x\| = \sup_{n \geqslant 1} |\langle f(s) - x, x_n^* \rangle|$$

is measurable. Let $(x_n)_{n \geqslant 1}$ be a dense sequence in X_1 with $x_1 = 0$.

Define the functions $\phi_n : X_1 \to \{x_1, \ldots, x_n\}$ as follows. For each $y \in X_1$ let $k(n, y)$ be the least integer $1 \leqslant k \leqslant n$ with the properties that $\|x_k\| \leqslant \|y\|$ and

$$\|y - x_k\| = \min_{1 \leqslant j \leqslant n} \|y - x_j\|,$$

and put $\phi_n(y) := x_{k(n,y)}$. Since $(x_n)_{n\geqslant 1}$ is dense in X_1, we obtain

$$\lim_{n\to\infty} \|\phi_n(y) - y\| = 0 \quad \text{and} \quad \|\phi_n(y)\| \leqslant \|y\| \qquad \forall y \in X_1.$$

Now define $f_n : S \to X$ by

$$f_n(s) := \phi_n(f(s)), \qquad s \in S.$$

Then for all $x \in X_1$, $\|f_n(x)\| \leqslant \|f(x)\|$. Moreover, for all $1 \leqslant k \leqslant n$ we have

$$\{f_n = x_k\} = \Big\{\|f - x_k\| = \min_{1\leqslant j\leqslant n} \|f - x_j\| < \min_{1\leqslant j<k} \|f - x_j\|\Big\}.$$

The set on the right hand side is in $\mathscr{A}$. Hence each f_n is simple, takes values in X_1, and for all $s \in S$ we have

$$\lim_{n\to\infty} \|f_n(s) - f(s)\| = \lim_{n\to\infty} \|\phi_n(f(s)) - f(s)\| = 0.$$

The final assertion follows from the fact that if X_0 has the stated properties, then $X_1 \subseteq X_0$. □

We state a number of simple corollaries. The first corollary follows from the proof of the implication (2)⇒(1).

Corollary 1.1.7. *If $f : S \to X$ is strongly measurable, then there exists a sequence of simple functions $(f_n)_{n\geqslant 1}$ such that*

$$\|f_n(x)\| \leqslant \|f(x)\| \quad \textit{and} \quad f_n(x) \to f(x) \quad \textit{for all } x \in X.$$

Corollary 1.1.8. *If $f : S \to X$ takes values in a closed subspace X_0 of X, then f is strongly measurable as a function with values in X if and only if f is strongly measurable as a function with values in X_0.*

Corollary 1.1.9. *The pointwise limit $f : S \to X$ of a sequence of strongly measurable functions $f_n : S \to X$ is strongly measurable.*

Proof. Each function f_n takes its values in a separable subspace of X. Then f takes its values in the closed linear span of these subspaces, which is again separable. The measurability of the functions $\langle f, x^*\rangle$ follows by noting that each $\langle f, x^*\rangle$ is the pointwise limit of the measurable functions $\langle f_n, x^*\rangle$. □

The next corollary gives the precise connection between measurability and strong measurability.

Corollary 1.1.10. *For a function $f : S \to X$, the following assertions are equivalent:*

(1) *f is strongly measurable;*
(2) *f is separably valued and measurable.*

Proof. (1)⇒(2): If f is strongly measurable, then f is weakly measurable and separably valued, say with values in a separable closed subspace X_0 of X. By the Hahn–Banach theorem, f is weakly measurable as an X_0-valued function, and by Corollary 1.1.2, this implies that f is measurable as an X_0-valued function. If $B \in \mathscr{B}(X)$, then $B_0 := B \cap X_0 \in \mathscr{B}(X_0)$, and

$$\{f \in B\} = \{f \in B_0\} \in \mathscr{A},$$

so that f is also measurable as an X-valued function.

(2)⇒(1): The functions $\langle f, x^* \rangle$ are measurable for all $x^* \in X^*$. The result now follows from the Pettis measurability theorem. □

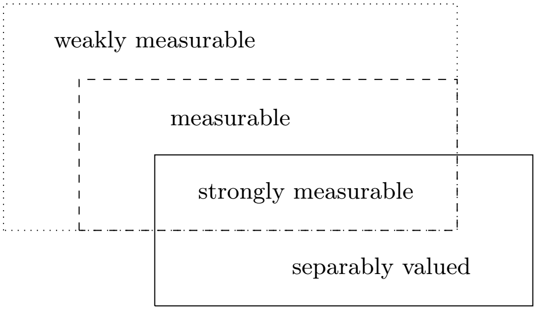

Fig. 1.1: The interrelations between different notions of measurability, as established in Theorem 1.1.6 and Corollary 1.1.10.

If $f : S \to X$ is strongly measurable and takes values in an open subset $O \subseteq X$, and $\phi : O \to Y$ is continuous, where Y is another Banach space, then $\phi \circ f$ is strongly measurable. In fact f is the pointwise limit of some simple f_n, and therefore also the pointwise limit of $\widetilde{f}_n := \mathbf{1}_{\{f_n \in O\}} f_n + \mathbf{1}_{\{f_n \in \complement O\}} x_0$, where x_0 is some fixed element of O. Then $\phi \circ \widetilde{f}_n$ is well defined, simple, and converges to $\phi \circ f$, which proves the claim.

More generally the following is true:

Corollary 1.1.11. *If $f : S \to X$ is strongly measurable and $\phi : X \to Y$ is measurable, where Y is another Banach space, then $\phi \circ f$ is strongly measurable.*

The proof of Corollary 1.1.11 uses the following topological fact.

Lemma 1.1.12. *Let E be a separable metric space and let F be a metric space. If $f : E \to F$ is measurable, then $f(E)$ is a separable subset of F.*

Proof. Suppose that $f(E)$ is non-separable. Then there exists an uncountable family of disjoint open sets $(O_i)_{i\in I}$ in F, each of which intersects $f(E)$. For every subset $I' \subseteq I$ we obtain an open set $O_{I'} := \bigcup_{i\in I'} O_i$ in F, hence a Borel set $f^{-1}(O_{I'})$ in E. If $I' \neq I''$, then also $O_{I'} \neq O_{I''}$, which shows that in E there are at least $2^{|I|}$ Borel sets. This is impossible since separable metric spaces have at most $2^{|\mathbb{N}|}$ Borel sets (see the Notes for a sketch of the proof). □

Proof of Corollary 1.1.11. It is clear that $\phi \circ f$ is measurable, and therefore by Corollary 1.1.10 it suffices to show that $\phi \circ f$ takes its values in a separable closed subspace of Y. The function f takes values in a separable closed subspace X_0 of X. Then $\phi \circ f$ takes its values in the subspace $\phi(X_0)$ of Y, which is separable by Lemma 1.1.12. □

1.1.b Functions on a measure space $(S, \mathscr{A}, \mu)$

Up to this point we have considered measurability properties of X-valued functions defined on a measurable space $(S, \mathscr{A})$. Next we consider functions defined on a measure space $(S, \mathscr{A}, \mu)$.

Definition 1.1.13. *A* μ-simple function *with values in X is a function of the form $f = \sum_{n=1}^N \mathbf{1}_{A_n} \otimes x_n$, where $x_n \in X$ and the sets $A_n \in \mathscr{A}$ satisfy $\mu(A_n) < \infty$.*

We say that a property holds *μ-almost everywhere* if there exists a μ-null set $N \in \mathscr{A}$ such that the property holds on the complement of N. Note that this definition makes no statement with regard to the validity of the property on N; the property may also hold on some subset of N, and this subset need not be in $\mathscr{A}$.

Definition 1.1.14. *A function $f : S \to X$ is* strongly μ-measurable *if there exists a sequence $(f_n)_{n\geqslant 1}$ of μ-simple functions converging to f μ-almost everywhere.*

When $X = \mathbb{K}$, we shall usually omit the prefix ‘strongly’. Thus, a function $f : S \to \mathbb{K}$ is *μ-measurable* if it is the μ-almost everywhere limit of a sequence if μ-simple functions $f_n : S \to \mathbb{K}$.

A measure μ is said to be *σ-finite* if it admits an *exhausting sequence*, i.e., an increasing sequence $S^{(1)} \subseteq S^{(2)} \subseteq \dots$ of sets in $\mathscr{A}$ of finite μ-measure such that $\bigcup_{n\geqslant 1} S^{(n)} = S$. The next result shows that strongly μ-measurable functions are μ-essentially supported on σ-finite measure spaces:

Proposition 1.1.15. *Suppose that $f : S \to X$ is strongly μ-measurable. Then we have a disjoint decomposition $S = S_0 \cup S_1$ with $S_0, S_1 \in \mathscr{A}$ such that:*

(i) *$f \equiv 0$ μ-almost everywhere on S_0;*
(ii) *μ is σ-finite on S_1.*

Proof. Suppose $f_n \to f$ μ-almost everywhere, with each f_n a μ-simple function, say $f_n = \sum_{m=1}^{N_n} \mathbf{1}_{A_{mn}} \otimes x_{mn}$ with $\mu(A_{mn}) < \infty$. The sets $S_1 = \bigcup_{n\geqslant 1} \bigcup_{m=1}^{N_n} A_{mn}$ and $S_0 = \complement S_1$ have the desired properties. $\square$

The next proposition relates the notions of 'strong measurability' and 'strong μ-measurability'.

Proposition 1.1.16. *Consider a function $f : S \to X$.*

(1) *If f is strongly μ-measurable, then f is μ-almost everywhere equal to a strongly measurable function.*
(2) *If μ is σ-finite and f is μ-almost everywhere equal to a strongly measurable function, then f is strongly μ-measurable.*

Proof. (1): Suppose that $f_n \to f$ pointwise outside the null set $N \in \mathscr{A}$, with each f_n μ-simple. Then we have $\lim_{n\to\infty} \mathbf{1}_{\complement N} f_n = \mathbf{1}_{\complement N} f$ pointwise on S, and since the functions $\mathbf{1}_{\complement N} f_n$ are simple, $\mathbf{1}_{\complement N} f$ is strongly measurable. Clearly, $f = \mathbf{1}_{\complement N} f$ μ-almost everywhere.

(2): Let $\widetilde{f}$ be a strongly measurable function and let $N \in \mathscr{A}$ be a μ-null set such that $f = \widetilde{f}$ on $\complement N$. If $(\widetilde{f}_n)_{n\geqslant 1}$ is a sequence of simple functions converging pointwise to $\widetilde{f}$, then $\lim_{n\to\infty} \widetilde{f}_n = f$ on $\complement N$, so $\lim_{n\to\infty} \widetilde{f}_n = f$ μ-almost everywhere. If $(S^{(n)})_{n\geqslant 1}$ is an exhausting sequence for μ, then the functions $f_n := \mathbf{1}_{S^{(n)}} \widetilde{f}_n$ are μ-simple and we have $\lim_{n\to\infty} f_n = f$ μ-almost everywhere. $\square$

Part (2) is wrong without the σ-finiteness assumption:

Example 1.1.17. The constant function $\mathbf{1}$ is always strongly measurable (as an indicator function); it is strongly μ-measurable if and only if μ is σ-finite.

Remark 1.1.18. As a consequence of Proposition 1.1.16(1), a separably valued and strongly μ-measurable function $f : S \to X$ is strongly $\mathscr{A}_\mu$-measurable, where $\mathscr{A}_\mu$ is the *completion* of $\mathscr{A}$ with respect to μ, i.e., the σ-algebra generated by $\mathscr{A}$ and the collection of all subsets of the μ-null sets in $\mathscr{A}$. The converse holds if μ is σ-finite.

Remark 1.1.19 (Pre-images with respect to strongly μ-measurable functions). If $f : S \to X$ is strongly μ-measurable and $\widetilde{f} : S \to X$ is strongly measurable such that $f = \widetilde{f}$ μ-almost everywhere, then by Corollary 1.1.10 the set

$$\{\widetilde{f} \in B\} := \{s \in S : \ \widetilde{f}(s) \in B\}$$

belongs to $\mathscr{A}$ for all Borel sets $B \in \mathscr{B}(X)$. The μ-measure of the set $\{\widetilde{f} \in B\}$ does not depend on the particular choice of measurable function $\widetilde{f}$. This justifies the notation

$$\mu\{f \in B\} := \mu\{\widetilde{f} \in B\}$$

which often use without further notice.

An X-valued function f is said to be *μ-essentially separably valued* if there exists a closed separable subspace X_0 of X such that $f(s) \in X_0$ for μ-almost all $s \in S$, and *weakly μ-measurable* if $\langle f, x^* \rangle$ is μ-measurable for all $x^* \in X^*$.

Theorem 1.1.20 (Pettis measurability theorem, second version). *For a function $f : S \to X$ the following assertions are equivalent:*

(1) *f is strongly μ-measurable;*
(2) *f is μ-essentially separably valued and weakly μ-measurable;*
(3) *f is μ-essentially separably valued and there exists a weak* dense subspace Y of X^* such that $\langle f, x^* \rangle$ is μ-measurable for all $x^* \in Y$.*

Moreover, if f takes its values μ-almost everywhere in a closed linear subspace X_0 of X, then f is the μ-almost everywhere pointwise limit of a sequence of X_0-valued simple functions.

Proof. The implications (1)⇒(2)⇔(3) are proved in the same way as in Theorem 1.1.6. For the implication (2)⇒(1) we have to be a bit more careful: the corresponding proof in Theorem 1.1.6 produces a sequence of simple functions, but not necessarily a sequence of μ-simple functions.

Let X_1 be a separable closed subspace in which f takes μ-almost all of its values, and let $(x_k^*)_{k\geqslant 1}$ be a norming sequence for X_1. The functions $g_k = \langle f, x_k^* \rangle$ are μ-measurable, and therefore by Proposition 1.1.15 we find decompositions $S = S_{k,0} \cup S_{k,1}$ such that $g_k \equiv 0$ μ-almost everywhere on $S_{k,0}$ and μ is σ-finite on $S_{k,1}$. Put $S_0 = \bigcap_{k\geqslant 1} S_{k,0}$ and $S_1 := \complement S_0$. Then $f \equiv 0$ μ-almost everywhere on S_0 and μ is σ-finite on S_1.

This argument shows that in the rest of the proof we may assume that μ is σ-finite. Then for all $x \in X$ the constant function $\mathbf{1}_S \otimes x$ is strongly μ-measurable. Letting $(x_j)_{j\geqslant 1}$ be a dense sequence in X_1, it follows that each of the functions $g_{jk} = \langle f - x_j, x_k^* \rangle$ is μ-measurable. Hence by Proposition 1.1.16 there is a μ-null set $N \in \mathscr{A}$ such that the functions $\mathbf{1}_{\complement N} g_{jk}$ are measurable. Replacing if necessary S by $\complement N$, we may therefore assume that each of the functions g_{jk} is measurable. Then also the functions $\|f - x_j\|$ are measurable.

Let f_n be the simple functions constructed in the proof of (2)⇒(1) of Theorem 1.1.6. These functions converge to f pointwise. If $(S^{(n)})_{n\geqslant 1}$ is an exhaustion for μ, we have $\mathbf{1}_{S^{(n)}} f_n \to f$ pointwise, and each of the functions $\mathbf{1}_{S^{(n)}} f_n$ is μ-simple. □

For completeness we list a number of corollaries which may be proved in the same way as their strongly measurable counterparts.

Corollary 1.1.21. *If $f : S \to X$ is strongly μ-measurable, there exists a sequence of μ-simple functions $(f_n)_{n\geqslant 1}$ such that*

$$\|f_n(x)\| \leqslant \|f(x)\| \quad \text{and} \quad f_n(x) \to f(x) \quad \text{for } \mu\text{-almost all } x \in X.$$

Corollary 1.1.22. *If $f : S \to X$ is strongly μ-measurable and takes values in a closed subspace X_0 of X almost everywhere, then f is strongly μ-measurable as a function with values in X_0.*

Corollary 1.1.23. *The μ-almost everywhere limit $f : S \to X$ of a sequence of strongly μ-measurable functions $f_n : S \to X$ is strongly μ-measurable.*

Corollary 1.1.24. *If $f : S \to X$ is strongly μ-measurable and $\phi : X \to Y$ is measurable, where Y is another Banach space, then $\phi \circ f$ is strongly μ-measurable provided at least one of the following two conditions is satisfied:*

(i) *μ is σ-finite;*
(ii) *$\phi(0) = 0$.*

Proof. If μ is σ-finite, the result follows by combining Corollary 1.1.11 and Proposition 1.1.16.

If $\phi(0) = 0$ and if $\lim_{n\to\infty} f_n = f$ is an almost everywhere approximation of f by μ-simple functions f_n, then f vanishes μ-almost everywhere outside the union A of the sets supporting the f_n. On A, μ restricts to a σ-finite measure; outside A we have $\phi \circ f = 0$ almost everywhere. Now the result follows by applying the previous case to $f|_A$, viewed as a strongly $\mu|_A$-measurable function on A. $\square$

The conditions (i) and (ii) cannot be omitted, even if $X = Y = \mathbb{K}$. Indeed, suppose that μ is non-σ-finite, let $f \equiv 0$ and $\phi(t) = 1$ for all $t \in \mathbb{K}$. The function $\mathbf{1} = \phi \circ f$ fails to be strongly μ-measurable.

The following result gives a convenient way to reduce proofs of vector-valued equalities to the scalar case. The corresponding version for strongly measurable functions is trivial.

Corollary 1.1.25. *If f and g are strongly μ-measurable X-valued functions which satisfy $\langle f, x^* \rangle = \langle g, x^* \rangle$ μ-almost everywhere for every $x^* \in Y$, where Y is a weak* dense linear subspace of X^*, then $f = g$ μ-almost everywhere.*

Proof. Both f and g take values in a separable closed subspace X_0 μ-almost everywhere, say outside the μ-null set N. Using Proposition B.1.11 we choose a sequence $(x_n^*)_{n\geqslant 1}$ in Y separating the points of X_0. Since $\langle f, x_n^* \rangle = \langle g, x_n^* \rangle$ outside a μ-null set N_n, we conclude that f and g agree outside the union of the μ-nulls set N and $\bigcup_{n\geqslant 1} N_n$. $\square$

The following example illustrates how the results above can be used to check strong measurability.

Example 1.1.26. Suppose X and Y are Banach spaces with X separable, and let $T : X \to Y$ be an injective bounded linear operator. If $f : S \to X$ is a function with the property that $T \circ f$ is strongly μ-measurable, then f is strongly μ-measurable. Indeed, f is separably valued by assumption, and for all $y^* \in Y^*$ the function $\langle f, T^* y^* \rangle$ is μ-measurable. The injectivity of T implies that T^* has weak* dense range, and therefore the result follows from the Pettis measurability theorem.

The example $S = (0, 1)$, $X = L^\infty(0, 1)$, $Y = L^2(0, 1)$, T the natural injection $f \mapsto f$, and $f(t) = \mathbf{1}_{(0,t)}$ shows that the separability assumption on X cannot be omitted.

1.1.c Operator-valued functions

Throughout these volumes there will be occasions to study properties of operator-valued functions. With respect to the uniform operator topology, the Banach space $\mathscr{L}(X,Y)$ is in general non-separable and because of this, few functions $f: S \to \mathscr{L}(X,Y)$ will be strongly measurable. To get a grasp of the situation, just consider the mapping $T: \mathbb{R} \to \mathscr{L}(L^2(\mathbb{R}))$, $t \mapsto T_t$, defined by

$$T_t f(u) = f(u+t), \quad u \in \mathbb{R}.$$

To see that this function fails to be strongly measurable with respect to the uniform operator topology of $\mathscr{L}(L^2(\mathbb{R}))$ we may argue as follows. For any two $s \neq t$ in $\mathbb{R}$ we note that

$$\|T_s - T_t\| = 2.$$

As a result, no matter how we choose the null set $N \subseteq \mathbb{R}$, the set $\{T_t : t \in \mathbb{R}\backslash N\}$ cannot be contained in a separable closed subspace of $\mathscr{L}(L^2(\mathbb{R}))$. Hence, by the Pettis measurability theorem, $t \mapsto T_t$ fails to be strongly measurable as an $\mathscr{L}(L^2(\mathbb{R}))$-valued function.

On the other hand, the orbits $t \mapsto T_t x$ enjoy many good properties, such as being continuous with respect to the norm of $L^2(\mathbb{R})$. This suggests the following definition.

Definition 1.1.27. *A function $f : S \to \mathscr{L}(X,Y)$ is called* strongly measurable *(respectively,* strongly μ-measurable*) if for all $x \in X$ the Y-valued function $fx : s \mapsto f(s)x$ is strongly measurable (respectively, strongly μ-measurable).*

It would be more accurate to refer to such functions as being *strongly (μ-)measurable with respect to the strong operator topology*, as opposed to those functions which are strongly measurable with respect to the uniform operator topology (for the definitions of these topologies we refer to Appendix B). The reader will agree that this terminology would be unnecessarily cumbersome. The slight ambiguity in our terminology is therefore taken for granted.

Proposition 1.1.28. *Let $(S,\mathscr{A})$ be a measurable space (respectively, $(S,\mathscr{A},\mu)$ a measure space) and let X and Y be Banach spaces. If $f : S \to X$ and $g : S \to \mathscr{L}(X,Y)$ are strongly (μ-)measurable, then $gf : S \to Y$ is strongly (μ-)measurable.*

Proof. By assumption there exists a sequence $(f_n)_{n\geqslant 1}$ of (μ-)simple functions converging pointwise to f (μ-almost everywhere). The functions gf_n are strongly (μ-)measurable and satisfy $\lim_{n\to\infty} gf_n \to gf$ pointwise (μ-almost everywhere). Corollary 1.1.9 (Corollary 1.1.23) now implies the strong (μ-)-measurability of gf. $\square$

Corollary 1.1.29. *Let X, Y, Z be Banach spaces. If $f : S \to \mathscr{L}(X,Y)$ and $g : S \to \mathscr{L}(Y,Z)$ are strongly (μ-)measurable, then $g \circ f : S \to \mathscr{L}(X,Z)$ is strongly (μ-)measurable.*

Proof. Let $x \in X$ be fixed. Since $s \mapsto f(s)x$ is strongly (μ-)measurable, the preceding proposition shows that $s \mapsto gf(s)x$ is strongly (μ-)measurable for all $x \in X$, so that $g \circ f$ is strongly (μ-)measurable. □

1.2 Integration

In this section we discuss the vector-valued extension of the Lebesgue integral, the so-called Bochner integral. At various places in this book we will also need its 'weak' companion, the Pettis integral.

1.2.a The Bochner integral

We fix a measure space $(S, \mathscr{A}, \mu)$. For a μ-simple function $f = \sum_{n=1}^{N} \mathbf{1}_{A_n} \otimes x_n$ we define

$$\int_S f \,\mathrm{d}\mu := \sum_{n=1}^{N} \mu(A_n) x_n.$$

It is routine to check that this definition does not depend on the particular representation of f and that $\| \int_S f \,\mathrm{d}\mu\| \leqslant \int_S \|f\| \,\mathrm{d}\mu$. If f and g are μ-simple, then $\int_S f \,\mathrm{d}\mu + \int_S g \,\mathrm{d}\mu = \int_S f + g \,\mathrm{d}\mu$.

Definition 1.2.1. *A strongly μ-measurable function $f : S \to X$ is* Bochner integrable *with respect to μ if there exists a sequence of μ-simple functions $f_n : S \to X$ such that*

$$\lim_{n\to\infty} \int_S \|f - f_n\| \,\mathrm{d}\mu = 0.$$

Note that $s \mapsto \|f(s) - f_n(s)\|$ is μ-measurable, so that this definition makes sense. From

$$\Big\| \int_S f_n \,\mathrm{d}\mu - \int_S f_m \,\mathrm{d}\mu \Big\| \leqslant \int_S \|f_n - f_m\| \,\mathrm{d}\mu \leqslant \int_S \|f_n - f\| \,\mathrm{d}\mu + \int_S \|f - f_m\| \,\mathrm{d}\mu$$

we see that the integrals $\int_S f_n \,\mathrm{d}\mu$ form a Cauchy sequence. By completeness, this sequence converges to an element of X. This limit is called the *Bochner integral* of f with respect to μ, notation

$$\int_S f \,\mathrm{d}\mu := \lim_{n\to\infty} \int_S f_n \,\mathrm{d}\mu.$$

It is routine to check that this definition does not depend on the choice of approximating sequence. When the reference measure μ is understood, we shall omit the phrase 'with respect to μ'.

If f is Bochner integrable and $f = g$ almost everywhere, then g is Bochner integrable and the Bochner integrals of f and g agree. In particular, in the definition of the Bochner integral it suffices that f be almost everywhere defined.

Proposition 1.2.2. *A strongly μ-measurable function $f : S \to X$ is Bochner integrable with respect to μ if and only if*

$$\int_S \|f\| \,\mathrm{d}\mu < \infty,$$

and in this case we have

$$\Big\| \int_S f \,\mathrm{d}\mu \Big\| \leqslant \int_S \|f\| \,\mathrm{d}\mu.$$

Proof. First let f be a strongly μ-measurable function satisfying $\int_S \|f\| \,\mathrm{d}\mu < \infty$. By Corollary 1.1.21 we may choose μ-simple functions f_n such that almost everywhere we have $\lim_{n\to\infty} f_n = f$ and $\|f_n\| \leqslant \|f\|$. Then, by dominated convergence,

$$\lim_{n\to\infty} \int_S \|f_n - f\| \,\mathrm{d}\mu = 0.$$

Conversely, let f be Bochner integrable and let the μ-simple functions f_n be as in Definition 1.2.1. Then for any large enough and fixed n,

$$\int_S \|f\| \,\mathrm{d}\mu \leqslant \int_S \|f - f_n\| \,\mathrm{d}\mu + \int_S \|f_n\| \,\mathrm{d}\mu < \infty.$$

Finally, the inequality $\| \int_S f \,\mathrm{d}\mu\| \leqslant \int_S \|f\| \,\mathrm{d}\mu$ is trivial for μ-simple functions, and the general case follows by approximation, using μ-simple functions as in Corollary 1.1.21. □

As a simple application of Proposition 1.2.2, note that if $f : S \to X$ is Bochner integrable, then for all $A \in \mathscr{A}$ the truncated function $\mathbf{1}_A f : S \to X$ is Bochner integrable and the restricted function $f|_A : A \to X$ is Bochner integrable with respect to the restricted measure $\mu|_A$. Moreover,

$$\int_S \mathbf{1}_A f \,\mathrm{d}\mu = \int_A f|_A \,\mathrm{d}\mu|_A.$$

Henceforth, both integrals will be denoted by $\int_A f \,\mathrm{d}\mu$.

An immediate consequence of Corollary 1.1.22 and Proposition 1.2.2 is the following result. A more refined result will be proved in Proposition 1.2.12 below.

Proposition 1.2.3. *Let $f : S \to X$ be Bochner integrable. If X_0 is a closed subspace of X such that $f(s) \in X_0$ for almost all $s \in S$, then f is Bochner integrable as an X_0-valued function. In particular, $\int_S f \,\mathrm{d}\mu \in X_0$.*

It is immediate from the definition of the Bochner integral that if $f : S \to X$ is Bochner integrable and T is a bounded linear operator from X into another Banach space Y, then $Tf : S \to Y$ is Bochner integrable and

$$T \int_S f \, d\mu = \int_S Tf \, d\mu. \tag{1.2}$$

In particular, for all $x^* \in E^*$ we have

$$\Big\langle \int_S f \, d\mu, x^* \Big\rangle = \int_S \langle f, x^* \rangle \, d\mu.$$

The identity (1.2) has a useful extension to the class of closed linear operators. A linear operator T defined on a linear subspace $\mathsf{D}(T) \subseteq X$ (the *domain* of T) and taking values in another Banach space Y, is said to be *closed* if the set

$$\mathsf{G}(T) := \{(x, Tx) : \ x \in \mathsf{D}(T)\}$$

(the *graph* of T) is a closed subspace of $X \times Y$. If T is closed, then $\mathsf{D}(T)$ is a Banach space with respect to the *graph norm*

$$\|x\|_{\mathsf{D}(T)} := \|x\| + \|Tx\|$$

and T is a bounded operator from $\mathsf{D}(T)$ to X. The *closed graph theorem* asserts that if $T : X \to Y$ is a closed linear operator with domain $\mathsf{D}(T) = X$, then T is bounded.

Theorem 1.2.4 (Hille). *Let $f : S \to X$ be Bochner integrable and let T be a closed linear operator with domain $\mathsf{D}(T)$ in X and with values in a Banach space Y. Suppose that f takes its values in $\mathsf{D}(T)$ almost everywhere and the almost everywhere defined function $Tf : S \to Y$ is Bochner integrable. Then f is Bochner integrable as a $\mathsf{D}(T)$-valued function, $\int_S f \, d\mu \in \mathsf{D}(T)$, and*

$$T \int_S f \, d\mu = \int_S Tf \, d\mu.$$

Proof. We begin with a simple observation which is a consequence of Proposition 1.2.2 and the fact that the coordinate mappings commute with Bochner integrals: if X_1 and X_2 are Banach spaces and $f_1 : S \to X_1$ and $f_2 : S \to X_2$ are Bochner integrable, then $f = (f_1, f_2) : S \to X_1 \times X_2$ is Bochner integrable and

$$\int_S f \, d\mu = \Big(\int_S f_1 \, d\mu, \ \int_S f_2 \, d\mu \Big).$$

Turning to the proof of the proposition, by the preceding observation the function $g : S \to X \times Y$, $g(s) := (f(s), Tf(s))$, is Bochner integrable. Moreover, since g takes its values in the graph $\mathsf{G}(T)$, Proposition 1.2.3 shows that (f, Tf) is Bochner integrable as a $\mathsf{G}(T)$-valued function. In particular, $\int_S g \, d\mu \in \mathsf{G}(T)$. On the other hand,

$$\int_S g \, d\mu = \Big(\int_S f \, d\mu, \int_S Tf \, d\mu \Big).$$

The identity in the statement of the theorem follows by combining these facts.

Finally, the mapping $x \mapsto (x, Tx)$ sets up an isomorphism of Banach spaces $\mathsf{D}(T) \simeq \mathsf{G}(T)$. Therefore the Bochner integrability of g as a $\mathsf{G}(T)$-valued function implies the Bochner integrability of f as a $\mathsf{D}(T)$-valued function. $\square$

It is implicit in the formulation of the theorem that the function $f : S \to \mathsf{D}(T)$ is strongly μ-measurable. This fact can also be deduced from Example 1.1.26.

Dominated convergence, substitution rule and Fubini's theorem

As a rule of thumb, results from the theory of Lebesgue integration carry over to the Bochner integral as long as there are no non-negativity assumptions involved. For example, as we next show, there are analogues of the dominated convergence theorem, the substitution rule and the Fubini theorem.

Proposition 1.2.5 (Dominated convergence theorem). *Let the functions $f_n : S \to X$ be Bochner integrable. If there exists a function $f : S \to X$ and a non-negative integrable function $g : S \to \mathbb{R}$ such that* $\lim_{n\to\infty} f_n = f$ *almost everywhere and $\|f_n\| \leqslant g$ almost everywhere, then f is Bochner integrable and we have*

$$\lim_{n\to\infty} \int_S \|f_n - f\| \, d\mu = 0.$$

In particular,

$$\lim_{n\to\infty} \int_S f_n \, d\mu = \int_S f \, d\mu.$$

Proof. We have $\|f_n - f\| \leqslant 2g$ almost everywhere, and therefore the result follows from the scalar dominated convergence theorem. $\square$

Proposition 1.2.6 (Substitution). *Let $(S, \mathscr{A}, \mu)$ be a measure space and let $(T, \mathscr{B})$ be a measurable space. Let $\phi : S \to T$ be measurable and let $f : T \to X$ be strongly measurable. Let $\nu = \mu \circ \phi^{-1}$ be the image measure of μ under ϕ. Then $f \circ \phi$ is Bochner integrable with respect to μ if and only if f is Bochner integrable with respect to ν, and in this situation*

$$\int_S f \circ \phi \, d\mu = \int_T f \, d\nu.$$

Proof. The function $f \circ \phi$ is measurable and has separable range, and therefore it is strongly measurable by the Pettis measurability theorem. For functions $g = \mathbf{1}_A$ with $A \in \mathscr{B}$ the above identity with f replaced by g follows from the definition of the image measure. By linearity it extends to simple functions g (both vector-valued and scalar valued). By monotone convergence we can then extend it to non-negative measurable functions g. In particular, this gives the identity

$$\int_S \|f \circ \phi\| \,\mathrm{d}\mu = \int_T \|f\| \,\mathrm{d}\nu,$$

which implies that $f \circ \phi$ is Bochner integrable if and only if f is Bochner integrable.

Now let $f : T \to X$ be Bochner integrable. By Corollary 1.1.7 we can find simple functions $f_n : T \to X$ such that $\|f_n\| \leqslant \|f\|$ and $f_n \to f$ pointwise. Then, by dominated convergence,

$$\int_S f \circ \phi \,\mathrm{d}\mu = \lim_{n\to\infty} \int_S f_n \circ \phi \,\mathrm{d}\mu = \lim_{n\to\infty} \int_T f_n \,\mathrm{d}\nu = \int_T f \,\mathrm{d}\nu,$$

where we applied the previously observed identity for simple functions. □

If $(S, \mathscr{A})$ and $(T, \mathscr{B})$ measurable spaces, we denote by $\mathscr{A} \times \mathscr{B}$ the *product σ-algebra*, i.e., the smallest σ-algebra in $S \times T$ containing all sets of the form $A \times B$ with $A \in \mathscr{A}$ and $B \in \mathscr{B}$. If μ and ν are σ-finite measures on $(S, \mathscr{A})$ and $(T, \mathscr{B})$, we denote by $\mu \times \nu$ the *product measure*, i.e., the unique σ-finite measure on $(S \times T, \mathscr{A} \times \mathscr{B})$ satisfying $\mu \times \nu(A \times B) = \mu(A)\nu(B)$ for all $A \in \mathscr{A}$ and $B \in \mathscr{B}$.

The Pettis measurability theorem and the Fubini theorem for scalar-valued functions imply the following simple observation. If $f : S \times T \to X$ is strongly measurable, then for all $s \in S$ the function $t \mapsto f(s,t)$ is strongly measurable, and for all $t \in T$ the function $s \mapsto f(s,t)$ is strongly measurable. Similarly, if $f : S \times T \to X$ is strongly $\mu \times \nu$-measurable, then for μ-almost all $s \in S$ the function $t \mapsto f(s,t)$ is strongly ν-measurable, and for ν-almost all $t \in T$ the function $s \mapsto f(s,t)$ is strongly μ-measurable. Concerning the Bochner integrability of these functions we have the following result.

Proposition 1.2.7 (Fubini's theorem). *Let $(S, \mathscr{A}, \mu)$ and $(T, \mathscr{B}, \nu)$ be σ-finite measure spaces and let $f : S \times T \to X$ be Bochner integrable.*

(1) *For almost all $s \in S$ the function $t \mapsto f(s,t)$ is Bochner integrable.*
(2) *For almost all $t \in T$ the function $s \mapsto f(s,t)$ is Bochner integrable.*
(3) *The functions $s \mapsto \int_T f(s,t) \,\mathrm{d}\nu(t)$ and $t \mapsto \int_S f(s,t) \,\mathrm{d}\mu(s)$ are Bochner integrable and*

$$\int_{S\times T} f(s,t) \,\mathrm{d}\mu{\times}\nu(s,t) = \int_T \int_S f(s,t) \,\mathrm{d}\mu(s) \,\mathrm{d}\nu(t) = \int_S \int_T f(s,t) \,\mathrm{d}\nu(t) \,\mathrm{d}\mu(s).$$

Proof. By the assumptions, the non-negative function $(s,t) \mapsto \|f(s,t)\|$ is integrable. Hence, by the scalar Fubini theorem, for almost all $s \in S$ the function $t \mapsto \|f(s,t)\|$ is integrable and for almost all $t \in T$ the function $s \mapsto \|f(s,t)\|$ is integrable. By Proposition 1.2.2 this proves (1) and (2). If we fix an arbitrary $x^* \in X^*$, by another application of the scalar Fubini theorem we have

$$\int_{S\times T} \langle f(s,t), x^*\rangle \,\mathrm{d}\mu \times \nu(s,t) = \int_T \int_S \langle f(s,t), x^*\rangle \,\mathrm{d}\mu(s) \,\mathrm{d}\nu(t)$$

$$= \int_S \int_T \langle f(s,t), x^* \rangle \, \mathrm{d}\nu(t) \, \mathrm{d}\mu(s).$$

By (repeated) use of the identity (1.2), the functional x^* may be taken out of each of the (repeated) integrals, and the desired identity follows by an application of the Hahn–Banach theorem. $\square$

Remark 1.2.8. When $(S, \mathscr{A}, \mu)$ and $(T, \mathscr{B}, \nu)$ are arbitrary measure spaces, not necessarily σ-finite, and $f : S \times T \to X$ is a strongly $(\mathscr{A} \times \mathscr{B})$-measurable function satisfying

$$\int_T \int_S \|f(s,t)\| \, \mathrm{d}\mu(s) \, \mathrm{d}\nu(t) < \infty \quad \text{or} \quad \int_S \int_T \|f(s,t)\| \, \mathrm{d}\nu(t) \, \mathrm{d}\mu(s) < \infty,$$

one may ask whether

$$\int_T \int_S f(s,t) \, \mathrm{d}\mu(s) \, \mathrm{d}\nu(t) = \int_S \int_T f(s,t) \, \mathrm{d}\nu(t) \, \mathrm{d}\mu(s).$$

This question has a negative answer, even for $(\mathscr{A} \times \mathscr{B})$-measurable indicator functions. For instance, one may take $S = (0,1)$, the unit interval with the Lebesgue measure on the Borel σ-algebra, $T = (0,1)$ with the counting measure on the power set, and $f = \mathbf{1}_{\{s=t\}}$. It is easy to see that f is $(\mathscr{A} \times \mathscr{B})$-measurable (the diagonal is the intersection of a countable family of rectangles) but it cannot be approximated by simple functions having the properties stated above (such an approximation would involve at most countably many points of T).

Convexity and integration

We continue with some aspects of the Bochner integral related to convexity and in particular we derive a vector-valued version of Jensen's inequality.

Lemma 1.2.9. *Let C be a closed convex set in a separable normed space X. Then there exists a sequence of affine functions of the form $\phi_i = \Re\langle \cdot, x_i^* \rangle - t_i$, with $x_i^* \in X^*$ and $t_i \in \mathbb{R}$, such that for all $x \in X$ one has*

$$x \in C \iff \forall i \ \phi_i(x) \geqslant 0.$$

Proof. The separability of X allows us to find a sequence $(x_i)_{i=1}^\infty$ in $\complement C$ which is dense in this set. Set $\delta_i := \operatorname{dist}(x_i, C) := \inf_{x \in C} \|x_i - x\|$ and note that $\delta_i > 0$. The open balls $B(x_i, \delta_i)$ are convex and disjoint from C. Therefore, by the Hahn–Banach separation theorem (Theorem B.1.3), there exist functionals $x_i^* \in X^*$ and real numbers t_i such that $\Re\langle x, x_i^* \rangle \geqslant t_i > \Re\langle y, x_i^* \rangle$ for all $x \in C$ and $y \in B(x_i, \delta_i)$. Thus $\phi_i(x) := \Re\langle x, x_i^* \rangle - t_i \geqslant 0$ for all $x \in C$ and $i = 1, 2, \dots$

It remains to check that for all $y \in \complement C$ we have $\phi_i(y) < 0$ for at least one i. To this end, denote $\delta := \operatorname{dist}(y, C)$ and pick an x_i from the chosen sequence satisfying $\|x_i - y\| < \frac{1}{2}\delta$. Then $\delta_i = \operatorname{dist}(x_i, C) > \frac{1}{2}\delta$, and therefore $y \in B(x_i, \delta_i)$. Hence $\phi_i(x_i) < 0$ and the proof is complete. $\square$

A function $\phi : X \to \mathbb{R}$ is called *lower semi-continuous* if

$$\phi(x_0) \leqslant \liminf_{x \to x_0} \phi(x) \qquad \forall x_0 \in X.$$

Lemma 1.2.10. *Let $\phi : X \to \mathbb{R}$ be a convex, lower semi-continuous function on a separable normed space X. Then there is a sequence of affine functions $\phi_i : X \to \mathbb{R}$ of the form $\phi_i = \Re\langle \cdot, y_i^* \rangle + a_i$, with $y_i^* \in X^*$ and $a_i \in \mathbb{R}$, such that*

$$\phi(x) = \sup_i \phi_i(x) \quad \forall x \in X.$$

Proof. The convexity of ϕ implies that the subset $C := \{(x,s) : \phi(x) \leqslant \Re s\}$ of $X \times \mathbb{K}$ is convex, and the lower semi-continuity of ϕ implies that C is closed. Lemma 1.2.9 implies the existence of $x_i^* \in X^*$, $s_i^* \in \mathbb{K}$, and $t_i \in \mathbb{R}$ such that for all $x \in X$ one has $(x,s) \in C$ if and only if

$$\Re\langle x, x_i^* \rangle + \Re(s \cdot s_i^*) \geqslant t_i, \qquad \forall\, i = 1, 2, \ldots. \tag{1.3}$$

For a fixed $x \in X$, the condition $(x,s) \in C$ (i.e., $\phi(x) \leqslant \Re s$) holds whenever $\Re s$ is large enough. It follows that we must have $\Re s_i^* > 0$ and $\Im s_i^* = 0$, for the other possibilities would lead to a contradiction with (1.3).

Dividing both sides by $s_i^* > 0$, we can rewrite (1.3) as

$$\Re s \geqslant -\Re\langle x, x_i^*/s_i^* \rangle + t_i/s_i^* =: \Re\langle x, y_i^* \rangle + a_i =: \phi_i(x), \qquad \forall\, i = 1, 2, \ldots,$$

or equivalently,

$$\Re s \geqslant \sup_i \phi_i(x).$$

But this holds if and only if $(x,s) \in C$ (by the properties of the ϕ_i as stated in Lemma 1.2.9), and this holds if and only if $\Re s \geqslant \phi(x)$ (by the definition of C). It follows that $\phi(x) = \sup_i \phi_i(x)$. $\square$

We are ready for the main inequality concerning convex functions:

Proposition 1.2.11 (Jensen's inequality). *Suppose that $\mu(S) = 1$. Let $f : S \to X$ be a Bochner integrable function and let $\phi : X \to \mathbb{R}$ be convex and lower semi-continuous. If $\phi \circ f$ is integrable, then*

$$\phi\Big(\int_S f \,\mathrm{d}\mu\Big) \leqslant \int_S \phi \circ f \,\mathrm{d}\mu.$$

Proof. By strong measurability, we may assume that X is separable. Let ϕ_i be the sequence of affine functions provided by Lemma 1.2.10. By the linearity of the integral, we have

$$\phi\Big(\int_S f \,\mathrm{d}\mu\Big) = \sup_i \phi_i\Big(\int_S f \,\mathrm{d}\mu\Big) = \sup_i \int_S \phi_i \circ f \,\mathrm{d}\mu$$

and, by the pointwise bound $\phi_i \circ f \leqslant \phi \circ f$ and the monotonicity of the integral, we have

$$\int_S \phi_i \circ f \,\mathrm{d}\mu \leqslant \int_S \phi \circ f \,\mathrm{d}\mu \qquad \text{for all } i.$$

□

In the next result, $\operatorname{conv} V$ denotes the *convex hull* of a subset $V \subseteq X$, i.e., the set of all finite linear combinations $\sum_{j=1}^k \lambda_j x_j$ with the real numbers $\lambda_j \geqslant 0$ satisfying $\sum_{j=1}^k \lambda_j = 1$ and $x_j \in V$ for $j = 1, \dots, k$. The closure of this set is denoted by $\overline{\operatorname{conv}}\, V$.

Proposition 1.2.12. *Suppose that $\mu(S) = 1$. If $f : S \to X$ is a Bochner integrable function, then*

$$\int_S f \,\mathrm{d}\mu \in \overline{\operatorname{conv}}\{f(s) : \ s \in S\}.$$

Proof. By strong measurability we may take X to be separable. Let $C := \overline{\operatorname{conv}}\{f(s) : \ s \in S\}$ denote the convex set on the right, and let ϕ_i be the affine functions provided by Lemma 1.2.9. Then

$$\phi_i\Big(\int_S f \,\mathrm{d}\mu\Big) = \int_S \phi_i(f) \,\mathrm{d}\mu \geqslant \int_S 0 \,\mathrm{d}\mu = 0$$

for all i, and therefore $\int_S f \,\mathrm{d}\mu \in C$. □

In the converse direction we have the following result.

Proposition 1.2.13. *Let $f : S \to X$ be Bochner integrable and let $C \subseteq X$ be closed and convex. If $\frac{1}{\mu(A)} \int_A f \,\mathrm{d}\mu \in C$ for all $A \in \mathscr{A}$ with $0 < \mu(A) < \infty$, then $f(s) \in C$ for almost all $s \in S$.*

Proof. By strong measurability we may assume X to be separable. Let ϕ_i be the affine functions provided by Lemma 1.2.9. The assumption of the proposition implies that

$$\frac{1}{\mu(A)} \int_A \phi_i \circ f \,\mathrm{d}\mu = \phi_i\Big(\frac{1}{\mu(A)} \int_A f \,\mathrm{d}\mu\Big) \geqslant 0$$

for all i and all $A \in \mathscr{A}$ with $0 < \mu(A) < \infty$. But this implies that $\phi_i \circ f \geqslant 0$ almost everywhere for all i, and thus $f \in C$ μ-almost everywhere. □

The next result gives a useful criterion for testing when two integrable functions are equal almost everywhere.

Proposition 1.2.14. *Let $\mathscr{C} \subseteq \mathscr{A}$ be a collection of sets, closed under taking finite intersections, containing S, and such that $\sigma(\mathscr{C}) = \mathscr{A}$. If $f : S \to X$ is Bochner integrable and satisfies*

$$\int_C f \,\mathrm{d}\mu = 0 \quad \forall C \in \mathscr{C},$$

then $f = 0$ almost everywhere.

Proof. By Corollary 1.1.25 it suffices to check that for all $x^* \in E^*$ we have $\langle f, x^* \rangle = 0$ almost everywhere. For each $x^* \in E^*$ the finite $\mathbb{K}$-valued measure $\nu(A) := \int_A \langle f, x^* \rangle \, d\mu$ satisfies $\nu(C) = 0$ for all $C \in \mathscr{C}$, and therefore $\nu = 0$ by Lemma A.1.3. This shows that $\int_A f \, d\mu = 0$ for every $A \in \mathscr{A}$, and an application of Proposition 1.2.13 with $C = \{0\}$ shows that $f = 0$ almost everywhere, as we claimed. □

1.2.b The Bochner spaces $L^p(S; X)$

Two strongly μ-measurable functions $f : S \to X$ and $g : S \to X$ will be called *equivalent* if $f(s) = g(s)$ for μ-almost all $s \in S$. This defines an equivalence relation on the set of all strongly μ-measurable functions from S to X. Following common practice we shall make no distinction between a function and its equivalence class.

Definition 1.2.15. *For $1 \leqslant p < \infty$ we define $L^p(S; X)$ as the linear space of all (equivalence classes of) strongly μ-measurable functions $f : S \to X$ for which*

$$\int_S \|f\|^p \, d\mu < \infty.$$

We define $L^\infty(S; X)$ as the linear space of all (equivalence classes of) strongly μ-measurable functions $f : S \to X$ for which there exists a real number $r \geqslant 0$ such that $\mu\{\|f\| > r\} = 0$.

Endowed with the norms

$$\|f\|_{L^p(S;X)} := \Big(\int_S \|f\|^p \, d\mu \Big)^{1/p}$$

and

$$\|f\|_{L^\infty(S;X)} := \inf \Big\{ r \geqslant 0 : \ \mu\{\|f\| > r\} = 0 \Big\},$$

the spaces $L^\infty(S; X)$, $1 \leqslant p \leqslant \infty$, are Banach spaces; the proofs for the scalar case carry over *verbatim*. In particular these proofs show that if $\lim_{n\to\infty} f_n = f$ in $L^p(S; X)$, then there exists a subsequence such that $\lim_{n\to\infty} f_{k_k} = f$ in X almost everywhere. Note that the elements of $L^1(S; X)$ are precisely the (equivalence classes of) Bochner integrable functions.

For $1 \leqslant p \leqslant \infty$ we write

$$L^p(S) := L^p(S; \mathbb{K}).$$

Note that a strongly μ-measurable function $s \mapsto f(s)$ belongs to $L^p(S; X)$ if and only if $s \mapsto \|f(s)\|$ belongs to $L^p(S)$. In the scalar-valued setting we will abbreviate 'strongly (μ-)measurable' to '(μ-)measurable'.

Remark 1.2.16. We have defined $L^\infty(S;X)$ as the space of essentially bounded strongly μ-measurable functions, identifying functions that are equal μ-almost everywhere. In the non-σ-finite case, this space should be carefully distinguished from the larger space of essentially bounded strongly measurable functions, identifying functions that are equal almost everywhere. Indeed, the constant functions $\mathbf{1} \otimes x$ are always bounded and strongly measurable, but they are strongly μ-measurable (and hence belong to $L^\infty(S;X)$) if and only if μ is σ-finite (cf. Example 1.1.17).

Our definition is prompted by the observation that no meaningful definition of the Bochner integral of a simple function f can be given if f is supported on sets of infinite measure. In the scalar-valued setting, this problem is usually circumvented by defining the Lebesgue integral first for non-negative simple functions, in which case the integral is allowed to take the value ∞. In the present vector-valued setting we are forced to working with μ-simple functions only. In view of this, our definition of $L^\infty(S;X)$ is the most natural one.

Occasionally we will write, when $\mathscr{F}$ is a sub-σ-algebra of $\mathscr{A}$,

$$L^p(S,\mathscr{F};X)$$

for the L^p-space with respect to the measure space $(S,\mathscr{F},\mu|_{\mathscr{F}})$. It coincides with the closed linear subspace of $L^p(S;X)$ consisting of all equivalence classes of functions with a representative that is strongly μ-measurable with respect to $\mathscr{F}$. We write $L^p(S,\mathscr{F}) := L^p(S,\mathscr{F};\mathbb{K})$.

Proposition 1.2.17. *For a strongly μ-measurable function $f : S \to X$ the following assertions are equivalent:*

(1) *$f \in L^\infty(S;X)$;*
(2) *$\langle f, x^*\rangle \in L^\infty(S)$ for all $x^* \in X^*$.*

Moreover,

$$\|f\|_{L^\infty(S;X)} = \sup_{\|x^*\|\leqslant 1} \|\langle f, x^*\rangle\|_{L^\infty(S)}.$$

We isolate an ingredient of the proof in a slightly more general form than immediately needed, since we shall have other use for it later:

Lemma 1.2.18. *Let $(S,\mathscr{A},\mu)$ be a measure space and let Y be a closed subspace of the dual X^* of a Banach space X. Suppose that $f : S \to X$ is a function with the property that $\langle f, x^*\rangle$ belongs to $L^1(S)$ for all $x^* \in Y$. Then*

$$T_f x^* := \langle f, x^*\rangle, \quad x^* \in Y$$

defines a bounded linear operator $T_f : Y \to L^1(S)$.

Proof. It is clear that T_f is well defined and linear. We claim that it is also closed, from which the boundedness follows by the closed graph theorem. Suppose that $\lim_{n\to\infty} x_n^* = x^*$ in Y and $\lim_{n\to\infty} T_f x_n^* = g$ in $L^p(S)$. By passing to a subsequence (if $p \in [1,\infty)$; this is not needed for $p = \infty$) we may assume that $\lim_{n\to\infty} T_f x_n^* = g$ almost everywhere on S. On the other hand, $\lim_{n\to\infty} T_f x_n^* = \lim_{n\to\infty}\langle f, x_n^*\rangle = \langle f, x^*\rangle$ pointwise on S. Therefore $g = \langle f, x^*\rangle$ and the claim is proved. □

Proof of Proposition 1.2.17. We only need to prove that (2) implies (1); the proof will also give the isometry. Without loss of generality we may assume X to be separable. Assuming (2), the mapping $x^* \mapsto \langle f, x^*\rangle$ from X^* to $L^\infty(S)$ is bounded by Lemma 1.2.18. Denote by C its norm. Select a norming sequence of norm one functionals $(x_n^*)_{n\geqslant 1}$ such that for almost all $s \in S$ we have $\|f(s)\| = \sup_{n\geqslant 1} |\langle f(s), x_n^*\rangle|$. If $\mu\{\|f\| > C\} = \mu(\bigcup_{n\geqslant 1}\{|\langle f(s), x_n^*\rangle| > C\})$ were strictly positive, then for some $n \geqslant 1$ we would have $\mu\{|\langle f(s), x_n^*\rangle| > C\} > 0$, a contradiction. □

Lemma 1.2.19 (Approximation by simple functions). *Let $1 \leqslant p \leqslant \infty$.*

(1) *The μ-simple functions are dense in $L^p(S;X)$, $1 \leqslant p < \infty$. In particular, the algebraic tensor product $L^p(S) \otimes X$ is dense in $L^p(S;X)$, $1 \leqslant p < \infty$.*
(2) *The μ-simple functions are dense in $L^\infty(S;X)$ with respect to convergence in measure. More precisely, if f is a function in $L^\infty(S;X)$ and $A \in \mathscr{A}$ has finite μ-measure, then for all $\varepsilon > 0$ there exist a μ-simple function $g : S \to X$ and a set $A' \in \mathscr{A}$ with $A' \subseteq A$ and $\mu(A \setminus A') < \varepsilon$ such that $\|g\|_\infty \leqslant \|f\|_\infty$ and*
$$\sup_{s\in A'} \|f(s) - g(s)\| < \varepsilon.$$

For a discussion of convergence in measure we refer to Appendix A.

Proof. Part (1) follows by arguing as in the proof of Proposition 1.2.2. For part (2) we may assume that $\mu(S) < \infty$. Fix $0 < \varepsilon < 1$. We may furthermore assume that X is separable and that $\|f\|_\infty = 1$. Let $(x_n)_{n\geqslant 1}$ be an ε-net in X. Put $A_n = \{s \in S : \|f(s) - x_n\| \leqslant \varepsilon/4\}$ and define $B_1 = A_1$, $B_{n+1} := A_{n+1} \setminus \bigcup_{j=1}^n B_j$. The functions $g_n := \sum_{j=1}^n \mathbf{1}_{B_n} \otimes x_j$ satisfy $\|g_n\|_\infty \leqslant 1 + \varepsilon/4$ and
$$\sup_{s\in\bigcup_{j=1}^n B_n} \|f(s) - g_n(s)\| \leqslant \varepsilon/4.$$

Therefore the functions $f_n := (1+\varepsilon/4)^{-1} g_n$ satisfy $\|f_n\|_\infty \leqslant 1$ and
$$\begin{aligned}\sup_{s\in\bigcup_{j=1}^n B_n} \|f(s) - f_n(s)\| &\leqslant \|f_n - g_n\|_\infty + \varepsilon/4 \\ &\leqslant \Big(1 - \frac{1}{1+\varepsilon/4}\Big)(1 - \varepsilon/4) + \varepsilon/4 = \varepsilon/2.\end{aligned}$$

In view of $\bigcup_{n\geqslant 1} B_n = \bigcup_{n\geqslant 1} A_n = S$ we have $\mu(\bigcup_{n\geqslant 1} B_n) \uparrow_n \mu(S)$. Hence for large enough n the functions $g = f_n$ have the desired properties. □

Remark 1.2.20. When μ is a finite measure and $\mathscr{A}'$ is an algebra of sets that generates $\mathscr{A}$, then the μ-simple functions supported on sets in $\mathscr{A}'$ are still dense in $L^p(S;X)$, $1 \leqslant p < \infty$. This useful improvement of (1) may be proved by using Lemma A.1.2 to approximate an arbitrary μ-simple function by μ-simple functions built on sets in $\mathscr{A}'$.

For example, every $f \in L^p(\mathbb{R}^d;X)$, $p \in [1,\infty)$, can be approximated in $L^p(\mathbb{R}^d;X)$ by functions of the form $\sum_{j=1}^n \mathbf{1}_{Q_j} \otimes x_j$, where $Q_j \subseteq \mathbb{R}^d$ are cubes with sides parallel to the coordinate axes.

Remark 1.2.21. It is sometimes useful to approximate functions in $L^\infty(S;X)$ by countably-valued μ-simple functions; see Lemma 2.1.4 in the next chapter.

As a quick application of part (1) and Proposition 1.2.2, let us give a proof of the continuous version of Minkowski's inequality. Of course, the inequality follows from its scalar counterpart applied to the function $s \mapsto \|f(s)\|$. But even the scalar version admits a neat proof in terms of the Bochner integral. It is this point that we wish to bring out.

Proposition 1.2.22 (Minkowski's inequality, continuous version). *Let $(S,\mathscr{A},\mu)$ and $(T,\mathscr{B},\nu)$ be measure spaces and let X be a Banach space. For all $1 \leqslant p \leqslant q < \infty$ the mapping*

$$\mathbf{1}_A \otimes (\mathbf{1}_B \otimes x) \mapsto \mathbf{1}_B \otimes (\mathbf{1}_A \otimes x)$$

extends uniquely to a contractive embedding

$$L^p(S;L^q(T;X)) \hookrightarrow L^q(T;L^p(S;X)).$$

Proof. For functions of the form $f = \sum_{m=1}^M \mathbf{1}_{A_m} \otimes (\sum_{n=1}^N \mathbf{1}_{B_n} \otimes x_{mn})$ with $\mu(A_m) < \infty$ and $\nu(B_n) < \infty$, the inequality in Proposition 1.2.2 gives

$$\begin{aligned}
\|f\|_{L^q(T,\nu;L^p(S;X))} &= \Big(\int_T \Big(\int_S \|f(x,y)\|^p \,\mathrm{d}\mu(x)\Big)^{q/p} \mathrm{d}\nu(y)\Big)^{1/q} \\
&= \Big\| \int_S \|f(x,\cdot)\|^p \,\mathrm{d}\mu(x)\Big\|_{L^{q/p}(T)}^{1/p} \\
&\leqslant \Big(\int_S \Big\| \|f(x,\cdot)\|^p \Big\|_{L^{q/p}(T)} \,\mathrm{d}\mu(x)\Big)^{1/p} \\
&= \Big(\int_S \Big(\int_T \|f(x,y)\|^q \,\mathrm{d}\nu(y)\Big)^{p/q} \mathrm{d}\mu(x)\Big)^{1/p} \\
&= \|f\|_{L^p(S;L^q(T;X))}.
\end{aligned}$$

Since these functions f are dense in $L^p(S;L^q(T;X))$ this gives the result. □

The case $p = q$ is of special interest:

Corollary 1.2.23. *Let $(S, \mathscr{A}, \mu)$ and $(T, \mathscr{B}, \nu)$ be measure spaces and let X be a Banach space. For all $1 \leqslant p < \infty$ the mapping $\mathbf{1}_A \otimes (\mathbf{1}_B \otimes x) \mapsto \mathbf{1}_B \otimes (\mathbf{1}_A \otimes x)$ induces an isometric isomorphism*

$$L^p(S; L^p(T; X)) \eqsim L^p(T; L^p(S; X)).$$

This is a version of Fubini's theorem that does not require any σ-finiteness assumptions (cf. the discussion in Remark 1.2.8).

In the σ-finite case one can say more:

Proposition 1.2.24. *Let $(S, \mathscr{A}, \mu)$ and $(T, \mathscr{B}, \nu)$ be σ-finite measure spaces and let $1 \leqslant p < \infty$. The mapping $\mathbf{1}_A \otimes (\mathbf{1}_B \otimes x) \mapsto (\mathbf{1}_A \otimes \mathbf{1}_B) \otimes x$ extends uniquely to an isometric isomorphism*

$$\iota : L^p(S; L^p(T; X)) \simeq L^p(S \times T; X).$$

If $\widetilde{f} : S \times T \to X$ is a strongly $(\mu \times \nu)$-measurable function in the equivalence class of $f \in L^p(S \times T; X)$, then for almost all $s \in S$ the function $\widetilde{f}(s, \cdot)$ belongs to $L^p(T; X)$, the function $F : s \mapsto \widetilde{f}(s, \cdot)$ defines an element of $L^p(S; L^p(T; X))$, and we have $\iota F = f$.

Proof. By the previous lemma and the density of linear combinations of μ-simple functions of the form $\mathbf{1}_A \otimes \mathbf{1}_B$ in $L^p(S \times T)$, linear combinations of μ-simple functions of the form $\mathbf{1}_A \otimes (\mathbf{1}_B \otimes x)$ and $(\mathbf{1}_A \otimes \mathbf{1}_B) \otimes x$ are dense in $L^p(S; L^p(T; X))$ and $L^p(S \times T; X)$, respectively; in each case we assume $\mu(A) < \infty$ and $\nu(B) < \infty$. On such functions, the mapping defined in the statement of the proposition is trivially seen to act isometrically.

To prove the final assertions let $f \in L^p(S \times T; X)$ be given and let $\widetilde{f}$ be a strongly $(\mu \times \nu)$-measurable representative. By the observation preceding Proposition 1.2.7, $s \mapsto \widetilde{f}(s, t)$ is strongly ν-measurable for almost all $s \in S$, and the Fubini theorem implies that $\|\widetilde{f}(s, \cdot)\|_X$ belongs to $L^p(T)$ for almost all $s \in S$. Hence $\widetilde{f}(s, \cdot) \in L^p(T; X)$ for almost all $s \in S$.

Next we check that the $L^p(T; X)$-valued function $F : s \mapsto \widetilde{f}(s, \cdot)$ is strongly μ-measurable. To this end let $f_n \to f$ in $L^p(S \times T; X)$ with each f_n a linear combination of $(\mu \times \nu)$-simple functions of the form $(\mathbf{1}_A \otimes \mathbf{1}_B) \otimes x$. Defining $F_n : s \mapsto f_n(s, \cdot)$, by the Fubini theorem we have $\|F_n(\cdot) - F(\cdot)\|_{L^p(T;X)} \to 0$ in $L^p(S)$. Therefore, after passing to a subsequence, we may assume that $\|F_n(s) - F(s)\|_{L^p(T;X)}$ for ν-almost all $s \in S$. Each $F_n : S \to L^p(T; X)$ is μ-simple, and the strong μ-measurability of F follows.

The identity $\iota F = f$ is again clear for linear combinations of $(\mu \times \nu)$-simple functions f of the above type, and the general case follows by density. $\square$

This proposition breaks down for $p = \infty$: the function $f(s, t) = \mathbf{1}_{\{0<s<t<1\}}$ belongs to $L^\infty((0, 1) \times (0, 1))$, but not to $L^\infty(0, 1; L^\infty(0, 1))$ (it fails to be essentially separably valued).

The next result states that Bochner integrals with values in a space $L^p(T; X)$ may be evaluated 'pointwise almost everywhere'.

Proposition 1.2.25 (Pointwise evaluation of L^p-valued integrals). *Let $(S,\mathscr{A},\mu)$ and $(T,\mathscr{B},\nu)$ be measure spaces, let X be a Banach space, and let $1\leqslant p<\infty$. If $F:S\to L^p(T;X)$ is Bochner integrable, there exists a strongly measurable function $f:S\times T\to X$ with the following properties:*

(i) *For μ-almost all $s\in S$, $f(s,\cdot)=F(s)$ in $L^p(T;X)$;*
(ii) *For ν-almost all $t\in T$, the function $s\mapsto f(s,t)$ is Bochner integrable and*
$$\Big(\int_S F\,\mathrm{d}\mu\Big)(t)=\int_S f(s,t)\,\mathrm{d}\mu(s).$$

If both μ and ν are σ-finite, the function f may be taken to be strongly $(\mu\times\nu)$-measurable and is essentially unique in the following sense: if $g:S\times T\to X$ is another strongly $(\mu\times\nu)$-measurable function satisfying (i)*, then $g=f$ $(\mu\times\nu)$-almost everywhere.*

Proof. We split the proof into four steps.

Step 1 – The strong μ-measurability allows us to make two reductions. Firstly, by Proposition 1.1.15 we may assume that μ is σ-finite. Secondly, we take advantage of the fact that f takes μ-almost all of its values in a separable closed subspace of $L^p(T;X)$. Proposition 1.2.29 below will show that such a subspace is always contained in a subspace of the form $L^p(T,\mathscr{B}',\nu|_{\mathscr{B}'};Y)$ for some sub-σ-algebra $\mathscr{B}'\subseteq\mathscr{B}$, with $\nu|_{\mathscr{B}'}$ σ-finite, and a separable closed subspace Y of X. Admitting this result for the moment, it follows that there is no loss of generality in assuming that also ν is σ-finite.

Step 2 – We first assume that ν is a *finite* measure. Then we may view F as a Bochner integrable function with values in $L^1(T;X)$.

Let f be a strongly measurable representative of ιF, where ι is the mapping of Proposition 1.2.24. By this proposition, f satisfies (i). The identity in (ii) is evidently true if F is a linear combination of functions of the form $\mathbf{1}_A\otimes(\mathbf{1}_B\otimes x)$ with $\mu(A)<\infty$ and $\nu(B)<\infty$. The general case follows by approximation: if $F_n\to F$ in $L^1(S;L^1(T;X))$, then $\int_S F\,\mathrm{d}\mu=\lim_{n\to\infty}\int_S F_n\,\mathrm{d}\mu$ in $L^1(T;X)$, and by passing to a subsequence we may assume that
$$\Big(\int_S F\,\mathrm{d}\mu\Big)(t)=\lim_{n\to\infty}\Big(\int_S F_n\,\mathrm{d}\mu\Big)(t)$$
for almost all $t\in T$. Also, $F_n\to F$ in $L^1(S;L^1(T;X))$ implies $f_n\to f$ in $L^1(S\times T;X)$ and hence, by the Fubini theorem and passing to a further subsequence, $f_n(\cdot,t)\to f(\cdot,t)$ in $L^1(S;X)$ for almost all $t\in T$. Hence, for almost all $t\in T$,
$$\begin{aligned}\Big(\int_S F\,\mathrm{d}\mu\Big)(t)&=\lim_{n\to\infty}\Big(\int_S F_n\,\mathrm{d}\mu\Big)(t)\\&=\lim_{n\to\infty}\int_S f_n(s,t)\,\mathrm{d}\mu(s)=\int_S f(s,t)\,\mathrm{d}\mu(s).\end{aligned}$$

Step 3 – If μ is σ-finite, we apply the above to the functions $F^{(j)} = \mathbf{1}_{T^{(j)}} F$, where $(T^{(j)})_{j\geqslant 1}$ is a disjoint decomposition of T by measurable sets of finite measure, and piece together the resulting functions f_j.

Step 4 – It remains to prove the uniqueness assertion. Suppose g is as stated in the proposition. Since $g(s,\cdot) = F(s)$ in $L^p(T;X)$ for almost all $s \in S$, for almost all $s \in S$ we have $g(s,t) = f(s,t)$ for almost all $t \in T$. Hence, by Fubini's theorem, $g(s,t) = f(s,t)$ for $(\mu \times \nu)$-almost all $(s,t) \in S \times T$. $\square$

Remark 1.2.26. If ν is σ-finite, the result also holds for $p = \infty$. Indeed, the finiteness of p was only needed in order to be able to apply Proposition 1.2.29.

We continue with a criterion for separability of Bochner spaces. We shall need the following terminology.

Definition 1.2.27. *A measure space $(S, \mathscr{A}, \mu)$ is called:*

(a) countably generated*, if there exists a sequence $(S_n)_{n\geqslant 1}$ in $\mathscr{A}$ which generates $\mathscr{A}$.*
(b) μ-countably generated*, if there exists a sequence $(S_n)_{n\geqslant 1}$ in $\mathscr{A}$, consisting of sets of finite μ-measure, which μ-essentially generates $\mathscr{A}$ in the sense that for all $A \in \mathscr{A}$ we can find a set A' in the σ-algebra generated by $(S_n)_{n\geqslant 1}$ such that $\mu(A \Delta A') = 0$.*
(c) purely infinite*, if $\mu(A) \in \{0, \infty\}$ for all $A \in \mathscr{A}$.*

Here, $A \Delta A' = (A \setminus A') \cup (A' \setminus A) = (A \cup A') \setminus (A \cap A')$ is the symmetric difference of A and A'.

Example 1.2.28. The Borel σ-algebra of $\mathbb{R}^d$ is countably generated (e.g., by all balls with rational radius centred at point with rational coordinates), but the Lebesgue σ-algebra of $\mathbb{R}^d$ (the completion of the Borel σ-algebra) is only λ-countably generated, where λ is the Lebesgue measure.

For $A \in \mathscr{A}$ we denote

$$\mathscr{A}|_A = \{A \cap B : \ B \in \mathscr{A}\} = \{B \in \mathscr{A} : \ B \subseteq A\}.$$

This is a σ-algebra in A. The restriction of μ to $\mathscr{A}|_A$ is denoted by $\mu|_A$.

Proposition 1.2.29 (Separability of Bochner spaces). *Let $(S, \mathscr{A}, \mu)$ be a measure space, let $1 \leqslant p < \infty$, and let X be a Banach space. If $\dim L^p(S;X) \geqslant 1$, the following assertions are equivalent:*

(1) *$L^p(S;X)$ is separable;*
(2) *X is separable and we have a disjoint decomposition $S = S_0 \cup S_1$ in $\mathscr{A}$ such that $(S_0, \mathscr{A}|_{S_0}, \mu|_{S_0})$ is purely infinite and $(S_1, \mathscr{A}|_{S_1}, \mu|_{S_1})$ is μ-countably generated.*

If these equivalent conditions hold, then $(S_1, \mathscr{A}|_{S_1}, \mu|_{S_1})$ is σ-finite and we have $L^p(S;X) = L^p(S_1;X)$ isometrically.

Proof. (1)⇒(2): By considering the sets $\{f\otimes x_0 : f \in L^p(S)\}$ with $x_0 \neq 0$ and $\{f_0 \otimes x : x \in X\}$ with $f_0 \neq 0$ we see that $L^p(S)$ and X are separable. Choose a dense sequence in $L^p(S)$ and approximate each element of this sequence by μ-simple functions. Denote by S_1 the union of the countably many sets in $\mathscr{A}$ of finite μ-measure involved in this process. Obviously, $(S_1, \mathscr{A}|_{S_1}, \mu|_{S_1})$ is σ-finite and μ-countably generated.

Let $S_0 = \complement S_1$. We claim that $(S_0, \mathscr{A}|_{S_0}, \mu|_{S_0})$ is purely infinite. For if $A_0 \in \mathscr{A}|_{S_0}$ with $\mu(A_0) < \infty$, the function $\mathbf{1}_{A_0} \in L^p(S)$ can be approximated μ-almost everywhere by the dense sequence of μ-simple functions we selected before. Since these functions are supported in S_1, then necessarily $\mu(A_0 \setminus S_1) = 0$. But by assumption $A_0 \subseteq S_0$, and therefore $\mu(A_0) = 0$.

(2)⇒(1): Let $S = S_0 \cup S_1$ as indicated. Then for all $f \in L^p(S;X)$ we have $f|_{S_0} = 0$ almost everywhere, and therefore the mapping $f \mapsto f|_{S_1}$ is an isometry from $L^p(S;X)$ onto $L^p(S_1;X)$.

It remains to be shown that if $(S_1, \mathscr{A}|_{S_1}, \mu|_{S_1})$ is μ-countably generated by the countable family $\mathscr{C} \subseteq \mathscr{A}$, with $\mu(C_n) < \infty$ for all $C_n \in \mathscr{C}$, then $L^p(S_1;X)$ is separable. Let $\mathscr{B}_1$ denote the algebra generated by $\mathscr{C}$; this algebra is countable. Then for all $A_1 \in \mathscr{A}|_{S_1}$ we can find $A \in \sigma(\mathscr{B}_1) = \sigma(\mathscr{C})$ such that $\mu(A_1 \Delta A) = 0$, and then $B \in \mathscr{B}_1$ such that $\mu(A_1 \Delta B) < \epsilon$, hence $\mu(A \Delta B) < \epsilon$. It then follows from Remark 1.2.20 that all μ-simple functions in $L^p(S_1)$ can be approximated, in the norm of $L^p(S_1)$, by linear combinations of the (countably many) indicators $\mathbf{1}_B$ with $B \in \mathscr{B}_1$. Hence, the $\mathbb{K}$-rational linear combinations of these indicators form a countable dense set U in $L^p(S_1)$. If V is a countable dense set in X, then $U \otimes V$ is a countable dense set in $L^p(S_1;X)$. □

Convolution and mollification on $\mathbb{R}^d$

We conclude this subsection with a discussion of some results for the Bochner spaces $L^p(\mathbb{R}^d;X)$ over the measure space $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d), \lambda)$, where λ is the Lebesgue measure. We always understand that $\mathbb{R}^d$ is equipped with this measure space structure, unless otherwise stated. The dimension $d \geqslant 1$ is taken arbitrary and fixed.

Lemma 1.2.30 (Young's inequality). *For $f \in L^p(\mathbb{R}^d;X)$ and $\phi \in L^1(\mathbb{R}^d)$, the convolution*

$$\phi * f(x) := \int_{\mathbb{R}^d} \phi(y) f(x-y)\,\mathrm{d}y$$

is well defined as a Bochner integral for almost every $x \in \mathbb{R}^d$, and

$$\|\phi * f\|_p \leqslant \|\phi\|_1 \|f\|_p.$$

Proof. By applications of Hölder's inequality and Fubini's theorem,

$$\|\phi * f\|_p^p \leqslant \|\phi\|_1^p \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} \|f(x-y)\| \frac{|\phi(y)|\,\mathrm{d}y}{\|\phi\|_1} \Big)^p \,\mathrm{d}x$$

$$\leqslant \|\phi\|_1^p \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \|f(x-y)\|^p \frac{|\phi(y)| \, \mathrm{d}y}{\|\phi\|_1} \, \mathrm{d}x$$
$$= \|\phi\|_1^p \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \|f(x-y)\|^p \, \mathrm{d}x \frac{|\phi(y)| \mathrm{d}y}{\|\phi\|_1} = \|\phi\|_1^p \|f\|_p^p.$$

These estimates also provide the existence of the Bochner integrals. □

The reader familiar with Lusin's theorem will notice that the following lemma holds more generally with $\mathbb{R}^d$ replaced by an arbitrary locally compact Hausdorff space equipped with a regular Borel measure μ. Since we shall not need this more general result we content ourselves with the present special form, which admits a much simpler proof.

Lemma 1.2.31. *For all $p \in [1, \infty)$ the space $C_c(\mathbb{R}^d; X)$ of compactly supported functions is dense in $L^p(\mathbb{R}^d; X)$.*

Proof. Since every $f \in L^p(\mathbb{R}^d; X)$ can be approximated by λ-simple functions, where λ denotes the Lebesgue measure on $\mathbb{R}^d$, it suffices to approximate indicators $\mathbf{1}_B$ for Borel sets $B \in \mathscr{B}(\mathbb{R}^d)$ of finite measure. Fixing such B along with an arbitrary $\varepsilon > 0$, by Lemma A.1.2 there exists a set $B' \in \mathscr{B}(\mathbb{R}^d)$, which is a finite union of cubes of the form $Q = [a_1, b_1) \times \cdots \times [a_d, b_d)$, such that the symmetric difference $B \Delta B' = (B \setminus B') \cup (B' \setminus B)$ has measure less than ε. Since the L^p-norm of the indicator function of $B \Delta B'$ is then less than $\varepsilon^{1/p}$, it suffices to approximate each indicator function $\mathbf{1}_Q$, with Q a cube in $\mathbb{R}^d$ with finite side-lengths, by a function in $C_c(\mathbb{R}^d)$. But this is straightforward. □

As an application we prove the following useful approximation result.

Proposition 1.2.32 (Mollifiers). *Let $f \in L^p(\mathbb{R}^d; X)$ for some $p \in [1, \infty)$, and $\phi \in L^1(\mathbb{R}^d)$. For $\varepsilon > 0$, denote $\phi_\varepsilon(y) := \varepsilon^{-d} \phi(\varepsilon^{-1} y)$. Then*

$$\phi_\varepsilon * f \to c_\phi f \ \text{ in } L^p(\mathbb{R}^d; X)$$

as $\varepsilon \downarrow 0$, where $c_\phi := \int_{\mathbb{R}^d} \phi(y) \, \mathrm{d}y$.

Proof. Since $\int_{\mathbb{R}^d} \phi_\varepsilon(y) \, \mathrm{d}y = \int_{\mathbb{R}^d} \phi(y) \, \mathrm{d}y = c_\phi$, we have

$$\begin{aligned} \phi_\varepsilon * f(x) - c_\phi f(x) &= \int_{\mathbb{R}^d} \phi_\varepsilon(y)[f(x-y) - f(x)] \, \mathrm{d}y \\ &= \int_{\mathbb{R}^d} \phi(y)[f(x - \varepsilon y) - f(x)] \, \mathrm{d}y, \end{aligned} \tag{1.4}$$

and taking $L^p(\mathbb{R}^d; X)$ norms,

$$\|\phi_\varepsilon * f - c_\phi f\|_p \leqslant \int_{\mathbb{R}^d} |\phi(y)| \|f(\cdot - \varepsilon y) - f\|_p \, \mathrm{d}y.$$

Since $\|f(\cdot - \varepsilon y) - f\|_p \leqslant 2\|f\|_p$ uniformly in ε and y, and $\phi \in L^1(\mathbb{R}^d)$, it suffices by dominated convergence to show that $f(\cdot - \varepsilon y) \to f$ in $L^p(\mathbb{R}^d; X)$ for each fixed $y \in \mathbb{R}^d$. But this is evident for $f \in C_c(\mathbb{R}^d; X)$ and follows in general by the density of such functions in $L^p(\mathbb{R}^d; X)$. □

The related almost everywhere convergence problem will be taken up in Section 2.3.b.

1.2.c The Pettis integral

Although the theory of Bochner integration is very satisfactory, the conditions for Bochner integrability are sometimes quite restrictive. In this subsection we briefly sketch a more general integral, the Pettis integral, which can be thought of as the 'weak analogue' (in the functional analytic sense) of the Bochner integral.

We begin by considering the following general setup. Let $(S, \mathscr{A}, \mu)$ be a measure space and let Y be a closed subspace of the dual X^* of a Banach space X. Suppose that $f : S \to X$ is a function with the property that $\langle f, x^* \rangle$ belongs to $L^1(S)$ for all $x^* \in Y$. By Lemma 1.2.18, such a function induces a bounded linear mapping $T_f : Y \to L^1(S)$ by putting

$$T_f x^* := \langle f, x^* \rangle, \quad x^* \in Y.$$

For each set $A \in \mathscr{A}$ we now define

$$\tau(X,Y)\text{-}\int_A f \, \mathrm{d}\mu := T_f^* \mathbf{1}_A.$$

Here we view the function $\mathbf{1}_A \in L^\infty(S)$ as an element in the dual of $L^1(S)$. We call $\tau(X,Y)\text{-}\int_A f \, \mathrm{d}\mu$ the $\tau(X,Y)$-*integral* of f over A. It is the unique element in Y^* which satisfies

$$\Big\langle x^*, \tau(X,Y)\text{-}\int_A f \, \mathrm{d}\mu \Big\rangle = \int_A \langle f, x^* \rangle \, \mathrm{d}\mu, \quad x^* \in Y.$$

Example 1.2.33. Let $f : S \to X$ be a function which is *weakly integrable* with respect to μ, i.e., $\langle f, x^* \rangle \in L^1(S)$ for all $x^* \in X^*$. The $\tau(X, X^*)$-integral of f is called the *weak integral* of f. Note that $\tau(X, X^*)\text{-}\int_A f \, \mathrm{d}\mu \in X^{**}$ for all $A \in \mathscr{A}$.

Example 1.2.34. Let $f : S \to X^*$ be a function which is *weak** μ*-integrable*, i.e., $\langle x, f \rangle \in L^1(S)$ for all $x \in X$. The $\tau(X^*, X)$-integral of a function $f : S \to X^*$ is called the *weak** *integral* of f. Note that $\tau(X^*, X)\text{-}\int_A f \, \mathrm{d}\mu \in X^*$ for all $A \in \mathscr{A}$.

In the context of Example 1.2.33 it is natural to ask for conditions which guarantee that the integrals $\tau(X, X^*)\text{-}\int_A f \, \mathrm{d}\mu$ belong to X (rather than just to X^{**}) for all $A \in \mathscr{A}$. When X is reflexive (see Appendix B) this is of course automatically the case, but in general this need not always happen; see Example 1.2.39. This prompts the following definition.

Definition 1.2.35. *A weakly μ-integrable function $f : S \to X$ is called* Pettis integrable *if the adjoint T_f^* of the operator $T_f : x^* \mapsto \langle f, x^* \rangle$ maps $L^\infty(S)$ into X.*

Here, $L^\infty(S)$ is identified isometrically with a norming closed subspace of the dual $(L^1(S))^*$. If μ is σ-finite, then $(L^1(S))^* = L^\infty(S)$. These facts will be taken for granted for the moment; their vector-valued extensions will be proved in the next section.

Proposition 1.2.36. *For a weakly integrable function $f : S \to X$ the following assertions are equivalent:*

(1) *f is Pettis integrable with respect to μ;*
(2) *for all $A \in \mathscr{A}$ there exists an element $x_A \in X$ such that for all $x^* \in X^*$ we have*

$$\langle x_A, x^* \rangle = \int_A \langle f, x^* \rangle \, \mathrm{d}\mu. \tag{1.5}$$

Under these equivalent conditions, we have

$$x_A = T_f^* \mathbf{1}_A =: (P)\text{-}\int_A f \, \mathrm{d}\mu,$$

and call it the Pettis integral *of f over A.*

Proof. (1)⇒(2): The elements $x_A = T_f^* \mathbf{1}_A \in X$ have the required properties.

(2)⇒(1): We proceed in three steps.

Step 1 – If μ is finite, this implication follows from the fact that $T_f^* g \in X$ for all μ-simple functions g, noting that these functions are dense in $L^\infty(S)$.

Step 2 – The σ-finite case may be deduced from Step 1 as follows. Choose an integrable function $w : S \to (0, 1)$; such a function is easily seen to exist since μ is σ-finite. The measure $\nu := w\mu$ is finite. Now suppose that (1.5) holds. From

$$\langle x_A, x^* \rangle = \int_A \langle f, x^* \rangle \, \mathrm{d}\mu = \int_A \langle f/w, x^* \rangle \, \mathrm{d}\nu$$

we see that (1.5) holds with f and μ replaced by f/w and ν. Moreover, for $g \in L^\infty(S, \mu) = L^\infty(S, \nu)$,

$$\langle x^*, (T_f^{(\mu)})^* g \rangle = \int_S \langle f, x^* \rangle g \, \mathrm{d}\mu = \int_S \langle f/w, x^* \rangle g \, \mathrm{d}\nu = \langle x^*, (T_{f/w}^{(\nu)})^* g \rangle,$$

where the superscripts refer to the measure in question. Hence, as elements of X^{**}, we have $(T_f^{(\mu)})^* g = (T_{f/w}^{(\nu)})^* g$. Since we already checked (2)⇒(1) for finite measures, it follows that $(T_{f/w}^{(\nu)})^* g \in X$, and therefore also $(T_f^{(\mu)})^* g \in X$. This proves the Pettis integrability of f with respect to μ.

Step 3 – To prove the result for general measure μ, possibly non-σ-finite, let again $g \in L^\infty(S)$ be arbitrary; we wish to show that $T_f^* g \in X$. Let $S = S_0 \cup S_1$ as in Proposition 1.1.15 (applied to g); thus $g \equiv 0$ μ-almost everywhere on S_0 and μ is σ-finite on S_1. By what we have just proved for the σ-finite case, applied to the measure space $(S_1, \mathscr{A}|_{S_1}, \mu|_{S_1})$, it follows that $T^*_{f|_{S_1}} g|_{S_1} \in X$ and

$$\langle x^*, T^*_{f|_{S_1}} g|_{S_1}\rangle = \int_{S_1} \langle f|_{S_1}, x^*\rangle g|_{S_1} \,\mathrm{d}\mu|_{S_1} = \int_S \langle f, x^*\rangle g \,\mathrm{d}\mu = \langle x^*, T^*_f g\rangle$$

for all $x^* \in X^*$. Hence, $T^*_f g = T^*_{f|_{S_1}} g|_{S_1} \in X$. $\square$

Clearly, every Bochner integrable function is Pettis integrable and the integrals agree on every set $A \in \mathscr{A}$; in this case we have $T^*_f \mathbf{1}_A = \int_A f \,\mathrm{d}\mu$. An example of a Pettis integrable function that is not Bochner integrable will be given below.

The following result gives a sufficient condition for the Pettis integrability of a strongly μ-measurable function.

Theorem 1.2.37 (Pettis). *Let $1 < p \leqslant \infty$ and $1 \leqslant q < \infty$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Let $f : S \to X$ be a strongly μ-measurable function satisfying $\langle f, x^*\rangle \in L^p(S)$ for all $x^* \in X^*$. Then for all $\phi \in L^q(S)$ the function*

$$s \mapsto \phi(s) f(s)$$

is Pettis integrable.

Proof. We prove the theorem for $1 < q < \infty$, the case $q = 1$ being an obvious consequence of Proposition 1.2.17 (by which we actually obtain that $\phi f \in L^1(S; X)$ in that case). By Proposition 1.1.15 we may assume that μ is σ-finite. Let $(S^{(n)})_{n\geqslant 1}$ be an exhausting sequence.

By Lemma 1.2.18, the linear mapping $T_f : X^* \to L^p(S)$, $T_f x^* := \langle f, x^*\rangle$ is bounded, and we obtain

$$\int_S |\langle f, x^*\rangle|^p \,\mathrm{d}\mu \leqslant \|T_f\|^p \|x^*\|^p, \qquad \forall x^* \in X^*.$$

The strong μ-measurability of f implies that the function $s \mapsto \|f(s)\|$ is μ-measurable. Let $A \in \mathscr{A}$ be given, and put $A_n := A \cap \{\|f\| \leqslant n\} \cap S^{(n)}$. By Proposition 1.2.2, the integrals $x_n := \int_{A_n} \phi f \,\mathrm{d}\mu$ exist as a Bochner integrals in X. We claim that $(x_n)_{n\geqslant 1}$ is a Cauchy sequence.

Let $\varepsilon > 0$ be arbitrary and fixed and choose $N \geqslant 1$ so large that

$$\int_{S\setminus S^{(N)}} |\phi|^q \,\mathrm{d}\mu < \varepsilon^q.$$

For all $n \geqslant m \geqslant N$ and $x^* \in X^*$ we have, noting that $A_n \backslash A_m \subseteq S \setminus S^{(N)}$ and using Hölder's inequality,

$$|\langle x_n - x_m, x^*\rangle| \leqslant \int_{A_n\backslash A_m} |\phi||\langle f, x^*\rangle| \,\mathrm{d}\mu$$

$$\leqslant \int_{S\setminus S^{(N)}} |\phi||\langle f, x^*\rangle| \, \mathrm{d}\mu$$
$$\leqslant \Big(\int_{S\setminus S^{(N)}} |\phi|^q \, \mathrm{d}\mu\Big)^{1/q} \Big(\int_S |\langle f, x^*\rangle|^p \, \mathrm{d}\mu\Big)^{1/p}$$
$$\leqslant \varepsilon \|T_f\| \|x^*\|.$$

Taking the supremum over all $x^* \in X^*$ with $\|x^*\| \leqslant 1$ and then using that $\varepsilon > 0$ was arbitrary, the claim follows. Hence the limit

$$x_A := \lim_{n\to\infty} \int_{A_n} \phi f \, \mathrm{d}\mu$$

exists in X. Clearly,

$$\langle x_A, x^*\rangle = \lim_{n\to\infty} \int_{A_n} \phi\langle f, x^*\rangle \, \mathrm{d}\mu = \int_A \phi\langle f, x^*\rangle \, \mathrm{d}\mu$$

for all $x^* \in X^*$. This proves that f is Pettis integrable. □

Corollary 1.2.38. *Let $(S, \mathscr{A}, \mu)$ be a finite measure space and let $p > 1$. If $f : S \to X$ is strongly measurable and satisfies $\langle f, x^*\rangle \in L^p(S)$ for all $x^* \in X^*$, then f is Pettis integrable.*

For $p = 1$ and $q = \infty$, Theorem 1.2.37 and Corollary 1.2.38 break down:

Example 1.2.39. Let $(A_n)_{n\geqslant 1}$ be a sequence of disjoint intervals of positive measure $|A_n|$ in the interval $(0, 1)$ and define $f : (0, 1) \to c_0$ by $f = \sum_{n\geqslant 1} \mathbf{1}_{A_n} \otimes e_n/|A_n|$, where $(e_n)_{n\geqslant 1}$ is the standard unit basis of c_0. Then f is strongly measurable and weakly integrable. Denoting by $x^{**} \in l^\infty = c_0^{**}$ the weak integral of f, for all $n \geqslant 1$ we have

$$\langle e_n^*, x^{**}\rangle = \int_0^1 \langle f(t), e_n^*\rangle \, \mathrm{d}t = 1.$$

This shows that $x^{**} = \mathbf{1}$, the constant one vector in l^∞. This vector does not belong to c_0 and as a consequence f fails to be Pettis integrable.

This example also explains the need of considering arbitrary sets $A \in \mathscr{A}$ in the definition of the Pettis integral and not just the set S. In fact, let $A := (0, 2)$, let $f : (0, 1) \to c_0$ be the function of Example 1.2.39 and define $g : S \to c_0$ by $g(t) := f(t)$ for $t \in (0, 1)$ and $g(t) := -f(t - 1)$ for $t \in (1, 2)$. Then the weak integral $\tau(c_0, l^1)$-$\int_S g(t) \, \mathrm{d}t$ equals 0 and therefore belongs to c_0, whereas both $\tau(c_0, l^1)$-$\int_0^1 g(t) \, \mathrm{d}t$ and $\tau(c_0, l^1)$-$\int_1^2 g(t) \, \mathrm{d}t$ belong to $l^\infty \setminus c_0$.

We shall prove next that Example 1.2.39 is essentially the only example of a strongly measurable function which is weakly integrable but not Pettis integrable. We need the following classical result from Banach space theory.

Theorem 1.2.40 (Bessaga and Pełczyński). *Let $(x_n)_{n\geqslant 1}$ be a sequence in a Banach space X satisfying $\inf_{n\geqslant 1}\|x_n\|>0$. If there exists a finite constant $C\geqslant 0$ with the property that*

$$\Big\|\sum_{j=1}^{k}\epsilon_j x_j\Big\|\leqslant C$$

for all $k\geqslant 1$ and all signs $\epsilon_1,\dots,\epsilon_k\in\{-1,1\}$, then the closed linear span of $(x_n)_{n\geqslant 1}$ contains a subspace isomorphic to c_0.

Proof. The conditions imply that $\|x_n\|\leqslant 2C$ for all n. By a simple convexity argument, there is no loss of generality if we make the normalising assumption $\|x_n\|=1$ for all $n\geqslant 1$.

Let $(b_1,\dots,b_n)\in\mathbb{K}^n$ satisfy $\max_{1\leqslant j\leqslant n}|b_j|\leqslant 1$. If the b_j are real, then $(b_1,\dots,b_n)$ is a convex combination of the 2^n elements of the form $(\epsilon_1,\dots,\epsilon_n)$ with $\epsilon_j\in\{-1,1\}$. As a consequence,

$$\Big\|\sum_{j=1}^{n}b_j x_j\Big\|\leqslant C.$$

If the b_j are complex we consider real and imaginary parts separately and obtain the same estimate, but with constant $2C$.

After this preliminary remark we define, for $k=1,2,3,\dots$,

$$a_k:=\sup\Big\{\Big\|\sum_{j=k}^{l}b_j x_j\Big\|:\ l\geqslant k,\ |b_j|\leqslant 1\text{ for all }k\leqslant j\leqslant l\Big\}.$$

Note that $1\leqslant a_k\leqslant C$ (in the case of real scalars) and $1\leqslant a_k\leqslant 2C$ (in the case of complex scalars) for all $k\geqslant 1$ and that $a_1\geqslant a_2\geqslant\dots$. Set $a:=\lim_{k\to\infty}a_k$ and pick $k_1\geqslant 1$ such that

$$a_{k_1}\leqslant\frac{4}{3}a. \tag{1.6}$$

Choose $k_2>k_1$ and scalars $b_{k_1},\dots,b_{k_2-1}$ of modulus $\leqslant 1$ such that

$$\Big\|\sum_{j=k_1}^{k_2-1}b_j x_j\Big\|>\frac{3}{4}a.$$

Continuing this way we arrive at a sequence $k_1<k_2<\dots$ and vectors $y_i:=\sum_{j=k_i}^{k_{i+1}-1}b_j x_j$ such that $\|y_i\|>\frac{3}{4}a$ and all $|b_j|\leqslant 1$. For any $n\geqslant 1$ and scalars $c_1,\dots,c_n$, using (1.6) we find

$$\Big\|\sum_{i=1}^{n}c_i y_i\Big\|=\Big\|\sum_{i=1}^{n}\sum_{j=k_i}^{k_{i+1}-1}c_i b_j x_j\Big\|\leqslant\frac{4}{3}a\max_{\substack{1\leqslant i\leqslant n\\ k_1\leqslant j\leqslant k_{n+1}-1}}|c_i b_j|\leqslant\frac{4}{3}a\max_{1\leqslant i\leqslant n}|c_i|.$$

On the other hand if $\max_{1\leqslant i\leqslant n}|c_i| = |c_{i_0}|$, then putting

$$\widetilde{c}_i = \begin{cases} c_i & \text{if } i \neq i_0, \\ -c_i & \text{if } i = i_0, \end{cases}$$

we obtain

$$\begin{aligned}\Big\|\sum_{i=1}^n c_i y_i\Big\| &\geqslant 2\|c_{i_0}y_{i_0}\| - \Big\|\sum_{i=1}^n \widetilde{c}_i y_i\Big\| \\ &\geqslant \frac{3}{2}a|c_{i_0}| - \Big\|\sum_{i=1}^n \widetilde{c}_i y_i\Big\| \\ &\geqslant \frac{3}{2}a|c_{i_0}| - \frac{4}{3}a\max_{1\leqslant i\leqslant n}|c_i| = \frac{1}{6}a\max_{1\leqslant i\leqslant n}|c_i|.\end{aligned}$$

These inequalities prove that the closed linear span of $(y_i)_{i\geqslant 1}$ is isomorphic to c_0. □

Proposition 1.2.41. *Suppose that $(S,\mathscr{A},\mu)$ is a measure space and that X does not contain a closed subspace isomorphic to c_0. If $f : S \to X$ is strongly μ-measurable and weakly integrable, then f is Pettis integrable.*

Proof. Arguing by contradiction, suppose that f is strongly μ-measurable and weakly integrable, but not Pettis integrable.

By Proposition 1.1.15 we may assume that μ is σ-finite. Let $(S^{(n)})_{n\geqslant 1}$ be an exhausting sequence and put $A_n := \{\|f\| \leqslant n\} \cap S^{(n)}$. Let $g \in L^\infty(S)$. If the sequence $(x_n^g)_{n\geqslant 1}$ defined by the Bochner integrals $x_n^g := \int_{A_n} gf\,\mathrm{d}\mu$ converges to some limit x^g in X, then for all $x^* \in X^*$ we have

$$\langle x^g, x^*\rangle = \lim_{n\to\infty}\int_{A_n} g\langle f, x^*\rangle\,\mathrm{d}\mu = \int_S g\langle f, x^*\rangle\,\mathrm{d}\mu = \langle x^*, T_f^* g\rangle,$$

where $T_f^* : L^\infty \to X^{**}$ is the adjoint of the operator $T_f : X^* \to L^1(S)$, $T_f x^* = \langle f, x^*\rangle$. It follows that $T_f^* g = x^g \in X$.

Since by assumption f fails to be Pettis integrable, this argument shows that there must be a function $g \in L^\infty(S)$ such that the corresponding sequence $(x_n^g)_{n\geqslant 1}$ fails to converge in X. We may of course assume that $\|g\|_\infty = 1$. Pick $\varepsilon > 0$ and indices $m_1 < n_1 < m_2 < n_2 < \dots$ such that

$$\|x_{n_j}^g - x_{m_j}^g\| \geqslant \varepsilon, \qquad j = 1, 2, \dots$$

Put $y_j := x_{n_j}^g - x_{m_j}^g = \int_{A_{n_j}\setminus A_{m_j}} gf\,\mathrm{d}\mu$. Then $\|y_j\| \geqslant \varepsilon$ and for all finite sequences of scalars $a = (a_1, \dots, a_k)$ and $x^* \in X^*$,

$$\Big|\Big\langle \sum_{j=1}^k a_j y_j, x^*\Big\rangle\Big| = \Big|\sum_{j=1}^k \int_{A_{n_j}\setminus A_{m_j}} a_j g\langle f, x^*\rangle\,\mathrm{d}\mu\Big|$$

$$\leqslant \|a\|_\infty \sum_{j=1}^{k} \int_{A_{n_j}\setminus A_{m_j}} |g||\langle f, x^*\rangle| \,\mathrm{d}\mu \leqslant \|a\|_\infty \|T_f\|\|x^*\|,$$

where $\|a\|_\infty := \max_{1\leqslant j\leqslant k} |a_j|$. It follows that

$$\Big\|\sum_{j=1}^{k} a_j y_j\Big\| \leqslant \|a\|_\infty \|T_f\|.$$

Now the Bessaga–Pełczyński theorem implies that X contains a closed subspace isomorphic to c_0. □

We leave it to the reader to check that the proofs of Propositions 1.2.14 and 1.2.11–1.2.13 carry over *verbatim* to the Pettis integral. For reasons of completeness we give the precise formulations.

Proposition 1.2.42. *Let $\mathscr{C} \subseteq \mathscr{A}$ be a collection of sets, closed under taking finite intersections, containing S, and such that $\sigma(\mathscr{C}) = \mathscr{A}$. If $f : S \to X$ is Pettis integrable and satisfies*

$$(P)\text{-}\int_C f \,\mathrm{d}\mu = 0 \quad \forall C \in \mathscr{C},$$

then $f = 0$ μ-almost everywhere.

Proposition 1.2.43 (Jensen's inequality). *Suppose that $\mu(S) = 1$. Let $f : S \to X$ be a Pettis integrable function, let $\phi : X \to \mathbb{R}$ be convex and lower semi-continuous, and assume that $\phi \circ f$ is integrable. Then*

$$\phi\Big((P)\text{-}\int_S f \,\mathrm{d}\mu\Big) \leqslant \int_S \phi \circ f \,\mathrm{d}\mu.$$

Proposition 1.2.44. *Suppose that $\mu(S) = 1$. If $f : S \to X$ is a Pettis integrable function, then*

$$(P)\text{-}\int_S f \,\mathrm{d}\mu \in \overline{\mathrm{conv}}\{f(s) : \ s \in S\}.$$

Proposition 1.2.45. *Let $f : S \to X$ be Pettis integrable and let $C \subseteq X$ be closed and convex. If $\frac{1}{\mu(A)}(P)\text{-}\int_A f \,\mathrm{d}\mu \in C$ for all $A \in \mathscr{A}$ with $0 < \mu(A) < \infty$, then $f(s) \in C$ for almost all $s \in S$.*

1.3 Duality of Bochner spaces

One of the cornerstones of the theory of classical Lebesgue spaces is the fact that $L^{p'}(S)$ provides representations of all bounded linear functionals on $L^p(S)$ when $p \in [1, \infty)$, $\frac{1}{p} + \frac{1}{p'} = 1$, and $(S, \mathscr{A}, \mu)$ is σ-finite. When it comes to the Bochner spaces $L^p(S;X)$, the extent to which this duality remains valid turns out to depend on the Banach space X, and leads to another example (cf. Proposition 1.2.41) of the interplay of analysis and Banach space theory, a recurrent topic throughout these volumes.

We begin with partial duality results of an elementary nature, which is valid independently of the Banach space X.

1.3.a Elementary duality results

If $1 \leqslant p, q \leqslant \infty$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, then by Hölder's inequality every function $g \in L^q(S, X^*)$ defines a linear functional $\phi_g \in (L^p(S;X))^*$ by the formula

$$\langle f, \phi_g \rangle := \int_S \langle f(s), g(s) \rangle \, \mathrm{d}\mu(s),$$

and we have

$$\|\phi_g\|_{(L^p(S;X))^*} \leqslant \|g\|_{L^q(S;X^*)}.$$

Proposition 1.3.1. *Let $1 \leqslant p \leqslant \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Let Y be a norming closed subspace of X^*. The mapping $g \mapsto \phi_g$ defines an isometry of $L^q(S;Y)$ onto a closed subspace of $(L^p(S;X))^*$ which is norming for $L^p(S;X)$.*

Proof. Fix $g \in L^q(S;Y)$. We have already noted that

$$\|\phi_g\| \leqslant \|g\|_{L^q(S;X^*)} = \|g\|_{L^q(S;Y)}.$$

To check that $\|\phi_g\| \geqslant \|g\|_{L^q(S;Y)}$ we may assume that $\|g\|_{L^q(S;Y)} = 1$. We then have to prove that $\|\phi_g\| \geqslant 1$.

First let $1 < p < \infty$. If $\lim_{n\to\infty} g_n = g$ in $L^q(S;Y)$, then given $\varepsilon > 0$ we have

$$\|\phi_g\| \geqslant \|\phi_{g_n}\| - \|\phi_{g-g_n}\| \geqslant \|\phi_{g_n}\| - \|g - g_n\|_{L^q(S;Y)} \geqslant \|\phi_{g_n}\| - \varepsilon$$

for large enough n. It therefore suffices to prove the inequality $\|\phi_g\| \geqslant 1$ for all μ-simple functions g satisfying $\|g\|_{L^q(S;Y)} = 1$. Fix such a function, say $g = \sum_{n=1}^N \mathbf{1}_{A_n} \otimes x_n^*$ with disjoint sets $A_n \in \mathscr{A}$ satisfying $0 < \mu(A_n) < \infty$ and all $x_n^* \in Y$ non-zero.

Fix an arbitrary $\varepsilon > 0$ and choose norm one elements $x_n \in X$ such that $\langle x_n, x_n^* \rangle \geqslant (1-\varepsilon)\|x_n^*\|$. Define $f := \sum_{n=1}^N \mathbf{1}_{A_n} \otimes \|x_n^*\|^{q-1} x_n$. Then

$$\|f\|_{L^p(S;X)}^p = \sum_{n=1}^N \mu(A_n)\|x_n^*\|^{p(q-1)} = \sum_{n=1}^N \mu(A_n)\|x_n^*\|^q = \|g\|_{L^q(S;Y)}^q = 1$$

and

$$\langle f, \phi_g \rangle = \sum_{n=1}^N \mu(A_n)\|x_n^*\|^{q-1}\langle x_n, x_n^* \rangle \geqslant (1-\varepsilon)\sum_{n=1}^N \mu(A_n)\|x_n^*\|^q = 1 - \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this shows that $\|\phi_g\| \geqslant 1$.

For $p = \infty$ we proceed in the same way but take $f := \sum_{n=1}^N \mathbf{1}_{A_n} \otimes x_n$. Then $\|f\|_{L^\infty(S;X)} = 1$ and

$$\langle f, \phi_g \rangle = \sum_{n=1}^N \mu(A_n)\langle x_n, x_n^* \rangle \geqslant (1-\varepsilon) \sum_{n=1}^N \mu(A_n)\|x_n^*\| = 1 - \varepsilon.$$

Next let $p = 1$. Let $\varepsilon > 0$ be fixed and put

$$A_\varepsilon := \{s \in S : \ \|g(s)\| > 1 - \varepsilon\}.$$

This set has strictly positive (possibly infinite) measure. Since g is strongly μ-measurable, an approximation argument shows that A_ε contains a subset B_ε of finite measure. By covering the range of $g|_{B_\varepsilon}$, which may be assumed to be separable, with countably many balls of radius ε, we find an $x^* \in X^*$ such that the set

$$B_{\varepsilon,x^*} := B_\varepsilon \cap \{s \in S : \ \|g(s) - x^*\| < \varepsilon\}$$

has positive measure. Then $\|x^*\| \geqslant 1 - 2\varepsilon$. Let $x \in X$ be a norm one vector such that $\langle x, x^* \rangle \geqslant \|x^*\| - \varepsilon$. With $f := \mathbf{1}_{B_{\varepsilon,x^*}} \otimes x/\mu(B_{\varepsilon,x^*})$ we have $\|f\|_{L^1(S;X)} = 1$ and

$$\begin{aligned} |\langle f, \phi_g \rangle| &= \frac{1}{\mu(B_{\varepsilon,x^*})}\Big| \int_{B_{\varepsilon,x^*}} \langle x, g \rangle \, \mathrm{d}\mu \Big| \\ &\geqslant \frac{1}{\mu(B_{\varepsilon,x^*})} \int_{B_{\varepsilon,x^*}} \langle x, x^* \rangle \, \mathrm{d}\mu - \varepsilon \geqslant \|x^*\| - 2\varepsilon \geqslant 1 - 4\varepsilon. \end{aligned}$$

To check that $L^q(S;Y)$ is norming for $L^p(S;X)$ we just have to note that by what we have already proved, $L^p(S;X)$ is norming for $L^q(S;Y)$ whenever $\frac{1}{p} + \frac{1}{q} = 1$ and Y is norming for X. The result follows from this by reversing the roles of p and q and of X and Y, identifying X isometrically with a closed subspace of Y^* which is norming for Y. □

The μ-simple functions in $L^\infty(S;X)$ generally fail to be dense in $L^\infty(S;X)$ (cf. Lemma 1.2.19). Nevertheless the above proof shows the following fact:

Corollary 1.3.2. *If the subspace $Y \subseteq X^*$ is norming form X, then the μ-simple functions in $L^\infty(S;Y)$ are norming for $L^1(S;X)$.*

In general it is not true that $(L^p(S;X))^* = L^q(S;X^*)$, $1 \leqslant p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. As we will show below, a sufficient condition in order that this relation be correct for all σ-finite measures spaces $(S, \mathscr{A}, \mu)$ is that X^* should have the so-called Radon–Nikodým property. However, for atomic measure spaces this result holds without any conditions on X or its dual. Recall that a set $A \in \mathscr{A}$ is called an *atom* if $\mu(A) > 0$ and $A = A_0 \cup A_1$ with disjoint $A_0, A_1 \in \mathscr{A}$ implies that $\mu(A_0) = 0$ or $\mu(A_1) = 0$. The measure space $(S, \mathscr{A}, \mu)$ is called *atomic* if $\mathscr{A}$ is generated by its atoms.

Proposition 1.3.3. *Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, and let $(S, \mathscr{A}, \mu)$ be an atomic measure space. Then for any Banach space X we have*

$$(L^p(S;X))^* = L^q(S;X^*)$$

isometrically. In particular, for any set I we have $(\ell^p(I;X))^* = \ell^q(I;X^*)$ isometrically. If μ is σ-finite the same result holds for $p = 1$.

Here, for $1 \leqslant r < \infty$, the norm of $\ell^r(X)$ is defined as

$$\|x\|_{\ell^r(X)}^r = \sup_{J \subseteq I \text{ finite}} \sum_{j \in J} \|x_j\|^r.$$

Proof. Let a functional $\phi \in (L^p(S;X))^*$ be given; we must prove that it can be represented as an element of $L^q(S;X^*)$ of norm $\leqslant \|\phi\|$. In combination with Proposition 1.3.1 this will give the result.

Let A denote the set of all atoms in $\mathscr{A}$. For each $A \in \mathsf{A}$ the linear mapping $x \mapsto \langle \mathbf{1}_A \otimes x, \phi \rangle$ is bounded and therefore defines an element $x_A^* \in X^*$. Clearly we have $\|x_A^*\| \leqslant \|\phi\|$. This concludes the proof for $p = 1$ in case μ is σ-finite, for then A is countable and ϕ is represented by the function $g = \sum_{A \in \mathsf{A}} \mathbf{1}_A \otimes x_A^*$ in $L^\infty(S;X^*)$. Note that in the absence of σ-finiteness g would fail to be strongly μ-measurable.

Suppose now that $1 < p < \infty$. We must show that $\sum_{A \in \mathsf{A}} \|x_A^*\|^q < \infty$, for then ϕ is represented by the function $g = \sum_{A \in \mathsf{A}} \mathbf{1}_A \otimes x_A^*$ in $L^q(S;X^*)$. Fix $N \geqslant 1$ and select atoms $A_1, \dots, A_N \in \mathsf{A}$. We have

$$\Big(\sum_{n=1}^N \|x_{A_n}^*\|^q \Big)^{1/q} = \sup_{\|c\|_{\ell_N^p} \leqslant 1} \sum_{n=1}^N |c_n| \, \|x_{A_n}^*\|.$$

Fix $\varepsilon > 0$ and choose elements $x_{A_n} \in X$ of norm one such that $\langle x_{A_n}, x_{A_n}^* \rangle \geqslant \|x_{A_n}^*\| - \varepsilon/N$. If $\|c\|_{\ell_N^p} \leqslant 1$, then $\psi_c := \sum_{n=1}^N |c_n| \mathbf{1}_{A_n}$ defines an element of $L^p(S;X)$ of norm $\leqslant 1$ and

$$\sum_{n=1}^N |c_n| \, \|x_{A_n}^*\| \leqslant \varepsilon + \sum_{n=1}^N |c_n| \langle x_{A_n}, x_{A_n}^* \rangle = \varepsilon + \langle \psi_c, \phi \rangle \leqslant \varepsilon + \|\phi\|.$$

Since $\varepsilon > 0$ was arbitrary, combination of the above inequalities gives

$$\sum_{n=1}^N \|x_{A_n}^*\|^q \leqslant \|\phi\|^q.$$

Since $N \geqslant 1$ and $A_1, \dots, A_N \in \mathsf{A}$ were chosen arbitrarily, this concludes the proof. $\square$

1.3.b Duality and the Radon–Nikodým property

In Proposition 1.3.1 it has been established that if $1 \leqslant p \leqslant \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, then every function $g \in L^q(S; X^*)$ determines a functional $\phi_g \in (L^p(S; X))^*$ by the formula

$$\langle f, \phi_g \rangle = \int_S \langle f(s), g(s) \rangle \, \mathrm{d}\mu(s).$$

Moreover, $\|\phi_g\|_{(L^p(S;X))^*} = \|g\|_{L^p(S;X^*)}$. It is natural to ask whether the mapping $g \mapsto \phi_g$ actually establishes an isometric isomorphism of Banach spaces $L^q(S; X^*) \simeq (L^p(S; X))^*$. In the scalar-valued case, this question has an affirmative answer for $1 \leqslant p < \infty$ if μ is σ-finite. The proof depends on the *Radon–Nikodým theorem*, which asserts that absolutely continuous measures of bounded variation have an L^1-density. This leads to the question as to whether X-valued measures of bounded variation have an L^1-density.

Let $(S, \mathscr{A}, \mu)$ be a measure space. A mapping $F : \mathscr{A} \to X$ is called an *X-valued measure* on $(S, \mathscr{A})$ if for all disjoint unions $A = \bigcup_{n \geqslant 1} A_n$ in $\mathscr{A}$ we have $F(A) = \sum_{n \geqslant 1} F(A_n)$ with convergence in the norm of X. Since we may permute the order of the A_n, the convergence of this sum is unconditional.

The *variation* of an X-valued measure F is the mapping $\|F\| : \mathscr{A} \to [0, \infty]$ defined by

$$\|F\|(A) = \sup_{\pi} \sum_{B \in \pi} \|F(B)\|,$$

where the supremum is taken over all finite disjoint partitions π of A.

Definition 1.3.4. *Let $(S, \mathscr{A}, \mu)$ be a measure space and $F : \mathscr{A} \to X$ an X-valued measure. We say that:*

(a) *F has* bounded variation *if $\|F\|(S) < \infty$,*
(b) *F is* bounded by μ *on $A_0 \subseteq S$ if there is a constant $C < \infty$ such that for all $A \in \mathscr{A}$ with $A \subseteq A_0$ we have $\|F\|(A) \leqslant C\mu(A)$,*
(c) *F is* absolutely continuous *with respect to μ if for all $A \in \mathscr{A}$ with $\mu(A) = 0$ we have $F(A) = 0$.*

The reader may recall from classical measure theory that every complex-valued measure has bounded variation (we shall not use this fact here), but this need not be true in the vector-valued setting, as is shown by the following easy example.

Example 1.3.5. Consider $F : \mathscr{P}(\mathbb{Z}_+) \to \ell^p$, $p \in (1, \infty]$, given by $F(A) := \sum_{j \in A} j^{-1} e_j$, where e_j is the jth unit vector. It is easy to check that this is an ℓ^p-valued measure. Its variation is given by $\|F\|(A) = \sum_{j \in A} j^{-1}$, so in particular $\|F\|(\mathbb{Z}_+) = \infty$, and F does not have bounded variation.

Lemma 1.3.6. *If F has bounded variation, then $\|F\|$ is a finite measure.*

Proof. Let $A_1, A_2, \dots \in \mathscr{A}$ be disjoint, and let A be their union. We need to show that $\|F\|(A) = \sum_{n\geqslant 1} \|F\|(A_n)$.

Consider the disjoint decomposition $A = A_1 \cup \dots \cup A_N \cup (\bigcup_{n\geqslant N+1} A_N)$. If $A_n = A_{n,1} \cup \dots \cup A_{n,M_n}$ are disjoint decompositions, we see that

$$\|F\|(A) \geqslant \sum_{n=1}^{N} \sum_{m=1}^{M_n} \|F(A_{n,m})\| + \Big\|F\Big(\bigcup_{n\geqslant N+1} A_N\Big)\Big\| \geqslant \sum_{n=1}^{N} \sum_{m=1}^{M_n} \|F(A_{n,m})\|$$

and therefore $\|F\|(A) \geqslant \sum_{n\geqslant 1} \|F\|(A_n)$. On the other hand, suppose that $A = B_1 \cup \dots \cup B_M$ is a disjoint decomposition and fix $\varepsilon > 0$. Using the countable additivity of F we may choose N so large that $\|F(\bigcup_{n\geqslant N+1}(A_n \cap B_m))\| < \varepsilon/M$ for all $m = 1, \dots, M$. Then

$$\begin{aligned}\sum_{m=1}^{M} \|F(B_m)\| &= \sum_{m=1}^{M} \|F(\bigcup_{n\geqslant 1}(A_n \cap B_m))\| \leqslant \varepsilon + \sum_{m=1}^{M} \|F(\bigcup_{n=1}^{N}(A_n \cap B_m))\| \\ &\leqslant \varepsilon + \sum_{n=1}^{N}\sum_{m=1}^{M} \|F(A_n \cap B_m)\| \leqslant \varepsilon + \sum_{n=1}^{N} \|F\|(A_n).\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, it follows that $\|F\|(A) \leqslant \sum_{n\geqslant 1} \|F\|(A_n)$. $\square$

Lemma 1.3.7. *For an X-valued measure $F : \mathscr{A} \to X$ of bounded variation the following assertions are equivalent:*

(1) *F is absolutely continuous with respect to μ;*
(2) *$\|F\|$ is absolutely continuous with respect to μ;*
(3) *for all $\varepsilon > 0$ there is a $\delta > 0$ such that for all $A \in \mathscr{A}$ with $\mu(A) < \delta$ we have $\|F(A)\| < \varepsilon$;*
(4) *for all $\varepsilon > 0$ there is a $\delta > 0$ such that for all $A \in \mathscr{A}$ with $\mu(A) < \delta$ we have $\|F\|(A) < \varepsilon$;*

Proof. (2)⇒(1) and (4)⇒(3) are trivial.

(3)⇒(2): Suppose $A \in \mathscr{A}$ satisfies $\mu(A) = 0$. If $A = A_1 \cup \dots \cup A_n$, then $\mu(A_j) = 0$ for $j = 1, \dots, n$ and therefore $\sum_{j=1}^{n} \|F(A_j)\| = 0$. It follows that $\|F\|(A) = 0$.

(1)⇒(4): Suppose (4) is false. Then there is an $\varepsilon > 0$ for which we can find sets $A_n \in \mathscr{A}$ of measure $\mu(A_n) < 2^{-n}$ such that $\|F\|(A_n) \geqslant \varepsilon$. Put $B_j := \bigcup_{n\geqslant j} A_n$ and $B := \bigcap_{j\geqslant 1} B_j$. Then $B_1 \supseteq B_2 \supseteq \dots$ and $\mu(B_j) < 2^{-j+1}$, so $\mu(B) = 0$. On the other hand, the countable additivity and finiteness of $\|F\|$ imply

$$\|F\|(B) = \lim_{j\to\infty} \|F\|(B_j) \geqslant \liminf_{n\to\infty} \|F\|(A_n) \geqslant \varepsilon.$$

It follows that $\|F(B')\| > 0$ for some $B' \in \mathscr{A}$ with $B' \subseteq B$; for this set we also have $\mu(B') = 0$. $\square$

The following lemma provides prominent examples of absolutely continuous X-valued measures of bounded variation:

Lemma 1.3.8. *Let $\phi \in L^1(S;X)$ be given and define $F : \mathscr{A} \to X$ by*

$$F(A) = \int_A \phi \,\mathrm{d}\mu, \quad A \in \mathscr{A}.$$

Then F is an X-valued measure of bounded variation, and absolutely continuous with respect to μ. Furthermore, for all $A \in \mathscr{A}$ we have

$$\|F\|(A) = \int_A \|\phi\| \,\mathrm{d}\mu.$$

If moreover $\phi \in L^\infty(S;X)$, then F is bounded by μ.

Proof. The restriction of μ to $\{\|\phi\| > 0\}$ being σ-finite, there is no loss of generality in assuming that μ is σ-finite.

The trivial estimate $\|F(A)\| \leqslant \int_A \|\phi\| \,\mathrm{d}\mu$ and the dominated convergence theorem imply that F is countably additive. It also implies that if $A = A_1 \cup \dots \cup A_N$ is a finite disjoint partition of A, then

$$\sum_{n=1}^N \|F(A_n)\| \leqslant \sum_{n=1}^N \int_{A_n} \|\phi\| \,\mathrm{d}\mu = \int_A \|\phi\| \,\mathrm{d}\mu.$$

Taking the supremum over all such partitions we obtain the inequality

$$\|F\|(A) \leqslant \int_A \|\phi\| \,\mathrm{d}\mu.$$

To prove the converse inequality we may assume that X is a real Banach space. By redefining ϕ on a null set, we may assume that ϕ is strongly measurable. Since ϕ is separably valued, we may further assume that X is separable. Fix $\varepsilon > 0$ and choose a countable symmetric norming set $C = \{x_n^* : n \geqslant 1\}$ of norm one vectors in X^*. For $n \geqslant 1$ put

$$A_n := \big\{ s \in S : \; \|\phi(s)\| \leqslant \langle \phi(s), x_n^* \rangle + \varepsilon \big\}.$$

These sets are measurable and we have $\bigcup_{n\geqslant 1} A_n = S$ by the symmetry of C. Put $B_1 := A_1$ and $B_{n+1} := A_{n+1} \setminus \bigcup_{m=1}^n B_n$ for $n \geqslant 1$. The sets B_n are measurable and disjoint and $\bigcup_{n\geqslant 1} B_n = S$.

Let $A \in \mathscr{A}$ be any set of finite measure. Then,

$$\begin{aligned}\int_A \|\phi\| \,\mathrm{d}\mu &= \sum_{n\geqslant 1} \int_{A\cap B_n} \|\phi\| \,\mathrm{d}\mu \\ &\leqslant \sum_{n\geqslant 1} \int_{A\cap B_n} \langle \phi, x_n^* \rangle \,\mathrm{d}\mu + \varepsilon\mu(A)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{n\geqslant 1} \langle F(A\cap B_n), x_n^*\rangle + \varepsilon\mu(A) \\
&\leqslant \sum_{n\geqslant 1} \|F(A\cap B_n)\| + \varepsilon\mu(A).
\end{aligned}
$$

Note that the right-hand side sum is convergent, since it is dominated by $\|F\|(A) + \varepsilon\mu(A)$. Hence for all $N \geqslant 1$ large enough,

$$
\int_A \|\phi\| \,\mathrm{d}\mu \leqslant \sum_{n=1}^{N} \|F(A\cap B_n)\| + \varepsilon(1+\mu(A)) \leqslant \|F\|(A) + \varepsilon(1+\mu(A)).
$$

Since $\varepsilon > 0$ was arbitrary, this proves the desired inequality for sets $A \in \mathscr{A}$ of finite measure. This inequality extends to arbitrary $A \in \mathscr{A}$ by taking intersections with sets from an exhausting sequence $(S^{(n)})_{n\geqslant 1}$. Indeed, by the case just proved we know that

$$
\int_{A\cap S^{(n)}} \|\phi\| \,\mathrm{d}\mu \leqslant \|F\|(A\cap S^{(n)}) \leqslant \|F\|(A),
$$

and the desired result follows by letting $n \to \infty$.

The absolute continuity of F with respect to μ follows from the absolute continuity of the non-negative measure $\|\phi\| \,\mathrm{d}\mu$ with respect to μ. □

Inquiring about the converse of Lemma 1.3.8 leads to the following notion, which turns out to provide a necessary and sufficient condition for the duality $(L^p(S;X))^* = L^q(S;X^*)$.

Definition 1.3.9. *A Banach space X is said to have the* Radon–Nikodým property (RNP) *with respect to a measure space $(S,\mathscr{A},\mu)$ if, for every X-valued measure F of bounded variation on $(S,\mathscr{A})$ that is absolutely continuous with respect to μ, there exists a function $\phi \in L^1(S;X)$ such that*

$$
F(A) = \int_A \phi \,\mathrm{d}\mu, \quad A \in \mathscr{A}.
$$

With this notion at hand, the characterisation of duality reads as follows:

Theorem 1.3.10. *Let $(S,\mathscr{A},\mu)$ be a σ-finite measure space, X be a Banach space, and let $1 \leqslant p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. The following assertions are equivalent:*

(1) *X^* has the RNP with respect to $(S,\mathscr{A},\mu)$;*
(2) *the mapping $\phi \mapsto g_\phi$ establishes an isometric isomorphism of Banach spaces*
$$
L^q(S;X^*) \simeq (L^p(S;X))^*.
$$

In order to avoid the use of the scalar version of the Radon–Nikodým theorem in the proof of this theorem we insert a technical lemma of elementary nature:

Lemma 1.3.11. *Let μ and ν be σ-finite positive measures on $(S, \mathscr{A})$, and ν be absolutely continuous with respect to μ. Then there is an exhausting sequence of sets $(S^{(n)})_{n\geqslant 1}$ such that μ and ν are finite on each $S^{(n)}$, and ν is bounded by μ on each $S^{(n)}$.*

Admitting the scalar-valued Radon–Nikodým theorem, this lemma would immediately follow by taking $S^{(n)} := S_n \cap \{\phi \leqslant n\}$, where $(S_n)_{n\geqslant 1}$ is an exhausting sequence of sets of finite μ-measure, and ϕ is the Radon–Nikodým density of ν with respect to μ. Otherwise, we can argue in a self-contained manner with the help of the following:

Lemma 1.3.12. *Let (P) be a property of measurable sets of a space $(S, \mathscr{A}, \mu)$ of finite measure such that:*

(i) *every set of measure zero has (P),*
(ii) *every set of positive measure has a subset of positive measure with (P),*
(iii) *every countable union of disjoint sets with (P) has (P).*

Then every measurable set has (P).

Proof. Let A be a set of positive measure. We choose disjoint subsets $A_n \subseteq A$ with (P) inductively as follows. Suppose that $A_1, \ldots, A_{n-1}$ are already chosen, which is vacuous if $n = 1$. If $C_n := A \setminus \bigcup_{k=1}^{n-1} A_k$ has measure zero, then it has (P), and also $A = C_n \cup \bigcup_{k=1}^{n-1} A_k$ has (P), and we are done. Otherwise, we can find some $A_n \subseteq C_n$ of positive measure with (P). Moreover, among different possible choices of A_n, we choose a quasi-maximal one, so that any other $A' \subseteq C_n$ with (P) satisfies $\mu(A') \leqslant 2\mu(A_n)$. Continuing this way, unless the process stops at some step with $\mu(C_n) = 0$, we find a disjoint sequence $A_n \subseteq A$ with (P). By disjointness and $\mu(S) < \infty$, we have $\mu(A_n) \to 0$.

Suppose for contradiction that $C := A \setminus \bigcup_{n=0}^{\infty} A_n$ has positive measure. Then there is a subset $A' \subseteq C$ of positive measure with (P). But then $\mu(A') > 2\mu(A_n)$ for large enough n, contradicting the quasi-maximal choice of A_n. Thus $\mu(C) = 0$, hence C has (P), and thus $A = C \cup \bigcup_{n=1}^{\infty} A_n$ has (P). $\square$

Proof of Lemma 1.3.11. By σ-finiteness, we can find an exhausting sequence on which μ and ν are finite. Studying each set in this sequence separately, we may assume that μ and ν are finite on all S.

We claim that every set of positive measure has a subset of positive measure, on which ν is bounded by μ. Suppose for contradiction that this fails for a set A and fix some $r > \nu(A)/\mu(A)$; whence every subset A' of A with positive measure has a subset A'' of positive measure such that $\nu(A'') \geqslant r\mu(A'')$. Let us denote this property of A'' by (P). On the space A of finite measure, this property (P) satisfies the conditions in Lemma 1.3.12, and hence by the lemma, every subset $A' \subseteq A$ satisfies $\nu(A') \geqslant r\mu(A')$. But this cannot be for $A' = A$ by the choice of r. This contradiction proves the claim.

Now let (P') be the property that A admits an exhausting sequence $(A^{(n)})_{n\geqslant 1}$ such that ν is bounded by μ on each $A^{(n)}$. If ν is bounded by

μ on A, then clearly (P') holds with the exhausting sequence $A^{(n)} \equiv A$. Thus, the previously establishes claim shows that every set of positive measure has a subset of positive measure with (P'). The other conditions of Lemma 1.3.12 are easily verified for $(P'$, and hence every measurable $A \subseteq S$ has (P'). In particular, S itself has (P'), which completes the proof. (Recall that we had already reduced considerations to finite measures in the beginning.) $\square$

Proof of theorem 1.3.10. (1)$\Rightarrow$(2): Let $\Lambda \in (L^p(S;X))^*$ be given. We must show that Λ can be represented as ϕ_g for some function $g \in L^q(S;X^*)$. Once this has been established, Proposition 1.3.1 shows that $\|\Lambda\| = \|g\|$.

Let $(S^{(n)})_{n\geqslant 1}$ be an exhausting sequence. Define $F_n : \mathscr{A} \to X^*$ by

$$\langle x, F_n(A)\rangle := \langle \mathbf{1}_{A\cap S^{(n)}} \otimes x, \Lambda\rangle, \quad A \in \mathscr{A},\ x \in X.$$

Then, for a disjoint union $A = \bigcup_{j=1}^k A_j$ and vectors $x_j \in X$ of norm one, we compute

$$\begin{aligned}\Big|\sum_{j=1}^k \langle x_j, F_n(A_j)\rangle\Big| &= \Big|\Big\langle \sum_{j=1}^k \mathbf{1}_{A_j\cap S^{(n)}} \otimes x_j, \Lambda\Big\rangle\Big| \\ &\leqslant \Big(\sum_{j=1}^k \|x_j\|^p \mu(A_j\cap S^{(n)})\Big)^{1/p}\|\Lambda\| = \mu(A\cap S^{(n)})^{1/p}\|\Lambda\|.\end{aligned}$$

Taking the supremum over all disjoint unions $A = \bigcup_{j=1}^k A_j$ and all $x_j \in X$ of norm one, we find that

$$\|F_n\|(A) \leqslant \mu(A\cap S^{(n)})^{1/p}\|\Lambda\|.$$

In particular, each F_n has bounded variation and is absolutely continuous with respect to μ. Since X^* has the RNP with respect to $(S,\mathscr{A},\mu)$, the X^*-valued measure F_n has a Radon–Nikodým derivative $g_n \in L^1(S;X^*)$ such that $F_n(A) = \int_A g_n \,\mathrm{d}\mu$ for all $A \in \mathscr{A}$. Since $F_n(A) = F_n(A\cap S^{(n)})$, we see that $\int_A g_n \,\mathrm{d}\mu = 0$ for all $A \in \mathscr{A}$ disjoint from $S^{(n)}$, from which it follows that ϕ_n vanishes identically outside $S^{(n)}$.

The functions g_n are consistent in the sense that $g_n|_{S^{(m)}} = g_m|_{S^{(m)}}$ whenever $m \leqslant n$. Therefore they uniquely define a strongly measurable function $g : S \to X^*$ and we have

$$\langle \mathbf{1}_{A\cap S^{(n)}} \otimes x, \Lambda\rangle = \langle x, F_n(A)\rangle = \int_A \langle x, g_n\rangle \,\mathrm{d}\mu = \langle \mathbf{1}_{A\cap S^{(n)}} \otimes x, g\rangle.$$

It remains to be shown that $g \in L^q(S;X^*)$ and that g represents Λ. To this end, observe that the mapping

$$f \mapsto \int_{A_n} \langle f, g\rangle \,\mathrm{d}\mu$$

with $A_n := S^{(n)} \cap \{\|g\| \leqslant n\}$, is a bounded linear functional on $L^p(S;X)$. This functional coincides with Λ for all μ-simple functions f supported in A_n, and by density for all $f \in L^p(A_n;X)$. Thus the bounded function $g|_{A_n}$, considered as an element of $L^q(A_n;X^*)$, represents the linear functional $\Lambda|_{L^p(A_n;X)}$, and by Proposition 1.3.1 we have

$$\|g\|_{L^q(A_n;X^*)} = \sup_{\|f\|_{L^p(A_n;X)} \leqslant 1} |\langle f,g\rangle| = \|\Lambda\|_{(L^p(A_n;X))^*} \leqslant \|\Lambda\|.$$

Since $A_n \uparrow S$ as $n \to \infty$, by monotone convergence theorem it follows that $\|g\|_{L^q(S;X^*)} \leqslant \|\Lambda\|$. Then, with the help of Hölder's inequality we can apply the dominated convergence theorem to the identity $\langle \mathbf{1}_{A_n} f, \Lambda\rangle = \langle \mathbf{1}_{A_n} f, g\rangle$ to show that we actually have $\langle f, \Lambda\rangle = \langle f, g\rangle$ for all $f \in L^p(S;X)$. This means that the function $g \in L^q(S;X^*)$ represents Λ.

(2)$\Rightarrow$(1): Let F be an X^*-valued measure of bounded variation, absolutely continuous with respect to μ. Let $(S^{(n)})_{n\geqslant 1}$ be an exhausting sequence for S such that μ and $\|F\|$ are finite on each $S^{(n)}$, and F is bounded by μ on $S^{(n)}$, say with constant C_n; this is provided by Lemma 1.3.11 applied to $\nu = \|F\|$.

On μ-simple functions $f = \sum_{j=1}^k \mathbf{1}_{A_j} \otimes x_j$, we define the functional

$$\Lambda_n(f) := \sum_{j=1}^k \langle x_j, F(A_j \cap S^{(n)})\rangle.$$

It is easy to see that this is well defined and linear, and

$$\begin{aligned}|\Lambda_n(f)| &\leqslant \sum_{j=1}^k \|x_j\|_X \|F(A_j \cap S^{(n)})\|_{X^*} \leqslant C_n \sum_{j=1}^k \|x_j\|_X \mu(A_j \cap S^{(n)}) \\ &= C_n \|f\|_{L^1(S^{(n)};X)} \leqslant C_n \mu(S^{(n)})^{1/p'} \|f\|_{L^p(S;X)},\end{aligned}$$

so that Λ_n extends to a bounded linear functional on $L^p(S;X)$. By assumption, it is represented by some $g_n \in L^{p'}(S;X^*)$, so that in particular

$$\langle x, F(A\cap S^{(n)})\rangle = \Lambda_n(\mathbf{1}_A \otimes x) = \int_S \langle \mathbf{1}_A \otimes x, g_n\rangle \,\mathrm{d}\mu = \Big\langle x, \int_{A\cap S^{(n)}} g \,\mathrm{d}\mu\Big\rangle, \quad (1.7)$$

where we have defined $g(s) := g_n(s)$ for $s \in S^{(n)}$: from the first two equalities in (1.7), it easily follows that g_n is supported on $S^{(n)}$, and g_n and g_m must agree almost everywhere on $S^{(n)} \cap S^{(m)}$, so this is an unambiguous definition of a μ-measurable function g.

It remains to check that $g \in L^1(S;X^*)$ and that it represents F on every $A \in \mathscr{A}$. From (1.7) and Lemma 1.3.8, it follows that

$$\|F\|(S^{(n)}) = \int_{S^{(n)}} \|g\| \,\mathrm{d}\mu.$$

Since $(S^{(n)})_{n\geqslant 1}$ is an exhausting sequence, this implies the same identity with S in place of $S^{(n)}$. Then we can also exhaust S by $S^{(n)}$ in (1.7) with the help of dominated convergence on the right, to conclude the proof. $\square$

Interestingly, for $1 < p < \infty$ the σ-finiteness condition can be dropped:

Corollary 1.3.13. *Let $(S, \mathscr{A}, \mu)$ be a measure space and let $1 < p < \infty$. If X^* has the RNP, the mapping $g \mapsto \phi_g$ from $L^q(S; X^*)$ to $(L^p(S; X))^*$, $\frac{1}{p} + \frac{1}{q} = 1$, defined by*

$$\langle f, \phi_g \rangle := \int_S \langle f, g \rangle \, \mathrm{d}\mu, \quad f \in L^p(S; X),$$

establishes an isometric isomorphism of Banach spaces

$$L^q(S; X^*) \simeq (L^p(S; X))^*.$$

Proof. We have already seen that, in the context of an arbitrary measure space $(S, \mathscr{A}, \mu)$, the mapping $g \mapsto \phi_g$ is well defined an isometric.

In the converse direction let $\Lambda \in (L^p(S; X))^*$ of norm one be given. Choose a sequence $(f_n)_{n \geqslant 1}$ of norm one functions in $L^p(S; X)$ such that $\langle f_n, \Lambda \rangle \geqslant 0$ and $\|\Lambda\| = \sup_{n \geqslant 1} \langle f_n, \Lambda \rangle$. By Proposition 1.1.15 we have a decomposition $S = S_0 \cup S_1$ with $S_0, S_1 \in \mathscr{A}$ such that $f_n = 0$ almost everywhere on S_0 for each $n \geqslant 0$ and μ is σ-finite on S_1.

Let Λ_1 denote the restriction of Λ to the closed subspace $L^p(S_1, \mathscr{A}|_{S_1}; X)$ of $L^p(S_1; X)$. Since X^* has the RNP with respect to the σ-finite measure space $(S_1, \mathscr{A}|_{S_1}, \mu|_{\mathscr{A}|_{S_1}})$, by Theorem 1.3.10 there is a function $g_1 \in L^q(S_1, \mathscr{A}|_{S_1}; X^*)$ such that $\Lambda_1 = \phi_{g_1}$. Let $g \in L^q(S; X^*)$ be the zero extension of g_1. We claim that $\Lambda = \phi_g$.

It clearly suffices to prove that Λ vanishes on $L^p(S_0, \mathscr{A}|_{S_0}; X)$. Suppose the contrary. Then there is a norm one function $f_0 \in L^p(S_0, \mathscr{A}|_{S_0}; X)$ such that $\delta_0 := \langle f_0, \Lambda \rangle > 0$. Then, for all $\lambda_0, \lambda_1 > 0$ with subject to the condition $\lambda_0^p + \lambda_1^p = 1$ and all $n \geqslant 1$,

$$\|\lambda_0 f_0 + \lambda_1 f_n\|_p^p = \lambda_0^p + \lambda_1^p = 1$$

and

$$\langle \lambda_0 f_0 + \lambda_1 f_n, \Lambda \rangle = \lambda_0 \delta_0 + \lambda_1 (1 - \varepsilon_n),$$

where we defined $\varepsilon_n \geqslant 0$ through the identity $\langle f_n, \Lambda \rangle = 1 - \varepsilon_n$. Since $\lim_{n \to \infty} \varepsilon_n = 0$ we infer that

$$\langle \lambda_0 f_0 + \lambda_1 f_n, \Lambda \rangle \geqslant \sup_{0 < \lambda < 1} [(1 - \lambda^p)^{1/p} \delta_0 + \lambda],$$

which by elementary calculus is seen to be greater than one, regardless the choice of $\delta_0 > 0$. This contracts the assumption $\|\Lambda\| = 1$. $\square$

For $p = 1$ the σ-finiteness assumption cannot be dropped:

Example 1.3.14. Let $S = (0, 1)$ with the counting measure ν on its power set. The mapping $\Lambda : L^1((0,1), \nu) \to \mathbb{R}$, defined for non-negative f by $\Lambda(f) := \sum_{t \in (0,1)} f(t)$ (the sum being defined as the supremum of all sums over finite subsets of $(0, 1)$), extends to a bounded functional on $L^1((0, 1), \nu)$ which cannot be represented by a function $g \in L^\infty((0, 1), \nu)$: the only candidate for g is the constant function $\mathbf{1}$, but this functions is not ν-measurable.

This failure is not an artifact of the fact that we have defined $L^\infty(S)$ in terms of μ-measurability (as opposed to measurability). To see this, consider a measure space consisting of a singleton $S = \{s\}$ with measure $\mu(\{s\}) = \infty$. Then $L^1(S) = \{0\} = L^\infty(S)$; on the other hand, the function $\mathbf{1}_{\{s\}}$ is bounded and measurable (but not μ-measurable).

1.3.c More about the Radon–Nikodým property

In the previous subsection, we introduced the Radon–Nikodým property with the goal of characterising the duality of the Bochner spaces. We now pursue a further study of this interesting property on its own right. In particular, this will allow us to present concrete examples of Banach space with (and without) the Radon–Nikodým property, and therefore with (and without) the natural duality of the related Bochner spaces.

A main workhorse of this subsection is the interplay of the Radon–Nikodým property with the following operator-theoretic notion:

Definition 1.3.15. *A bounded operator $T : L^1(S) \to X$ is* representable *if there exists a function $\phi \in L^\infty(S; X)$ such that*

$$Tf = \int_S f\phi \,\mathrm{d}\mu, \quad f \in L^1(S).$$

Note that if $T : L^1(S) \to X$ is represented by the function $\phi \in L^\infty(S; X)$, then

$$\begin{aligned}\|T\|_{\mathscr{L}(L^1(S),X)} &= \sup_{\substack{\|f\|_{L^1(S)}\leqslant 1\\ \|x^*\|\leqslant 1}} \Big| \int_S f\langle \phi, x^*\rangle \,\mathrm{d}\mu\Big| \\ &= \sup_{\|x^*\|\leqslant 1} \|\langle \phi, x^*\rangle\|_{L^\infty(S)} = \|\phi\|_{L^\infty(S;X)},\end{aligned}$$

the last identity being a consequence of Proposition 1.2.17.

This definition is linked with the Radon–Nikodým property as follows:

Theorem 1.3.16. *Let $(S, \mathscr{A}, \mu)$ be a σ-finite measure space. For a Banach space X the following assertions are equivalent:*

(1) *X has the RNP with respect to $(S, \mathscr{A}, \mu)$;*
(2) *every bounded linear operator $T : L^1(S) \to X$ is representable.*

Proof. (1)⇒(2): Suppose X has the RNP with respect to $(S, \mathscr{A}, \mu)$ and let $T : L^1(S) \to X$ be a bounded linear operator. Let $(S^{(n)})_{n\geqslant 1}$ be an exhausting sequence for $(S, \mathscr{A}, \mu)$.

Define $F_n : \mathscr{A} \to X$ by $F_n(A) := T(\mathbf{1}_{S^{(n)}\cap A})$. The estimate

$$\|F_n(A)\| \leqslant \|T\|\mu(S^{(n)} \cap A)$$

implies that F_n is an X-valued measure of bounded variation which is absolutely continuous with respect to μ. Since X is assumed to have the RNP, F_n has a density $\phi_n \in L^1(S;X)$. Then for all $A \in \mathscr{A}$,

$$T(\mathbf{1}_{S^{(n)}\cap A}) = F_n(A) = \int_A \phi_n \,\mathrm{d}\mu. \tag{1.8}$$

In particular for all $A \in \mathscr{A}$ disjoint from $S^{(n)}$ we see that $\int_A \phi_n \,\mathrm{d}\mu = 0$, from which it follows that ϕ_n vanishes identically outside $S^{(n)}$.

The functions ϕ_n are consistent in the sense that $\phi_n|_{S^{(m)}} = \phi_m|_{S^{(m)}}$ whenever $m \leqslant n$, and therefore we obtain a well-defined function $\phi : S \to X$. We finish the proof by showing that $\phi \in L^\infty(S;X)$ with $\|\phi\|_{L^\infty(S;X)} \leqslant \|T\|$ and that ϕ represents T.

Fix an arbitrary $\varepsilon > 0$ and put $S_\varepsilon^{(n)} := \{\|\phi_n\| \geqslant \|T\|+\varepsilon\}$. Then $S_\varepsilon^{(n)} \subseteq S^{(n)}$ up to a null set, and by Lemma 1.3.8 we have

$$\|F_n\|(S_\varepsilon^{(n)}) = \int_{S_\varepsilon^{(n)}} \|\phi_n\| \,\mathrm{d}\mu \geqslant (\|T\| + \varepsilon)\mu(S_\varepsilon^{(n)}).$$

On the other hand, for any partition $S_\varepsilon^{(n)} = B_1^{(n)} \cup \ldots \cup B_k^{(n)}$ we have

$$\sum_{j=1}^k \|F_n(B_j^{(n)})\| = \sum_{j=1}^k \|T\mathbf{1}_{B_j^{(n)}}\| \leqslant \|T\| \sum_{j=1}^k \mu(B_j^{(n)}) = \|T\|\mu(S_\varepsilon^{(n)})$$

and therefore $\|F_n\|(S_\varepsilon^{(n)}) \leqslant \|T\|\mu(S_\varepsilon^{(n)})$. Comparing these estimates we conclude that $\mu(S_\varepsilon^{(n)}) = 0$. Noting that $\bigcup_{n\geqslant 1} S_\varepsilon^{(n)} = \{\|\phi\| \geqslant \|T\| + \varepsilon\}$, we find that $\mu\{\|\phi\| \geqslant \|T\| + \varepsilon\} = 0$. Since $\varepsilon > 0$ was arbitrary this proves that $\mu\{\|\phi\| > \|T\|\} = 0$.

It remains to be shown that ϕ represents T. From (1.8), the defining properties of ϕ_n and ϕ, and an approximation argument we deduce that for all $f \in L^1(S)$,

$$T(\mathbf{1}_{S^{(n)}} f) = \int_S f\phi_n \,\mathrm{d}\mu = \int_S \mathbf{1}_{S^{(n)}} f\phi \,\mathrm{d}\mu.$$

Now we use dominated convergence to pass to the limit $n \to \infty$.

(2)⇒(1): Suppose that every operator $T : L^1(S) \to X$ is representable, and let $F : \mathscr{A} \to X$ be a measure of bounded variation which is absolutely continuous with respect to μ.

The non-negative measure $\|F\|$ is finite and absolutely continuous with respect to μ. Let $(A_i)_{i\geqslant 1}$ be a countable family of pairwise disjoint $\|F\|/\mu$-finite sets whose union equals S up to a μ-null set, as in Lemma 1.3.11. For $n = 1, 2, \ldots$ set

$$I_n := \{i \geqslant 1 : \ \|F\|(B) \leqslant n\mu(B) \text{ for all } B \in \mathscr{A} \text{ with } B \subseteq A_i\}$$

and set $J_1 := I_1$ and $J_{n+1} := I_{n+1} \setminus I_n$ for $n \geqslant 1$. Let $B_n := \bigcup_{i\in J_n} A_i$ and note that $\|F\|(B) \leqslant n\mu(B)$ for all $B \in \mathscr{A}$ with $B \subseteq B_n$, by the countable additivity of $\|F\|$ and μ.

For μ-simple functions $f = \sum_{j=1}^k c_j \mathbf{1}_{A_j}$ with the sets $A_j \in \mathscr{A}$ disjoint, define

$$T_n f := \sum_{j=1}^k c_j F(B_n \cap A_j).$$

It is easy to see that this definition does not depend on the representation of f and that T_n is linear. From

$$\begin{aligned}\|T_n f\| &\leqslant \sum_{j=1}^k |c_j| \|F\|(B_n \cap A_j) \\ &\leqslant n \sum_{j=1}^k |c_j| \mu(B_n \cap A_j) = n \big\| f|_{B_n} \big\|_1\end{aligned}$$

we see that T_n extends uniquely to a bounded operator from $L^1(S) \to X$. By assumption there exists a function $\phi_n \in L^\infty(S;X)$ representing T_n. If $A \in \mathscr{A}$, then

$$F(B_n \cap A) = T_n \mathbf{1}_A = \int_A \phi_n \, \mathrm{d}\mu.$$

By considering sets A which are disjoint from B_n we see that ϕ_n vanishes almost everywhere outside B_n. Set $\phi := \sum_{n\geqslant 1} \mathbf{1}_{B_n} \phi_n$; by the observation just made this sum converges pointwise almost everywhere for trivial reasons. From

$$\|F\|(B_n) = \int_{B_n} \|\phi\| \, \mathrm{d}\mu$$

it follows that $\phi \in L^1(S;X)$ and $\|\phi\|_1 = \|F\|(S)$. By the countable additivity of F and dominated convergence we obtain, for all $A \in \mathscr{A}$,

$$F(A) = \sum_{n\geqslant 1} F(B_n \cap A) = \sum_{n\geqslant 1} \int_{B_n \cap A} \phi \, \mathrm{d}\mu = \int_A \phi \, \mathrm{d}\mu.$$

□

As another tool in our investigation of the Radon–Nikodým property, we will use some elementary properties of *conditional expectations*, whose thorough study will be undertaken later in Section 2.6. In order to be self-contained at this stage, we make an *ad hoc* definition of the conditional expectation operator $E_{\mathscr{B}} : L^2(S) \to L^2(S)$ as the orthogonal projection of $L^2(S)$ onto its closed subspace $L^2(S, \mathscr{B})$; for all $B \in \mathscr{B}$ of finite measure and $f \in L^2(S)$ it satisfies

$$\int_B f \, \mathrm{d}\mu = \int_B E_{\mathscr{B}} f \, \mathrm{d}\mu,$$

simply using that $\mathbf{1}_B \in L^2(S,\mathscr{B})$ and $f - E_{\mathscr{B}}f \perp L^2(S,\mathscr{B})$. If $f \geqslant 0$ in $L^2(S)$, then $\int_B E_{\mathscr{B}}f \,\mathrm{d}\mu = \int_B f \,\mathrm{d}\mu \geqslant 0$ for all $B \in \mathscr{B}$, and therefore $E_{\mathscr{B}}f \geqslant 0$ in $L^2(S,\mathscr{B})$. If $f \in L^1(S) \cap L^2(S)$, then $f^{\pm} \geqslant 0$ in $L^2(S)$ and therefore

$$|E_{\mathscr{B}}f| \leqslant |E_{\mathscr{B}}f^+| + |E_{\mathscr{B}}f^-| = E_{\mathscr{B}}f^+ + E_{\mathscr{B}}f^- = E_{\mathscr{B}}|f|.$$

Integrating over S, we obtain that $E_{\mathscr{B}}f \in L^1(S,\mathscr{B})$ and $\|E_{\mathscr{B}}f\|_1 \leqslant \|f\|_1$. Finally, if $(S,\mathscr{A},\mu)$ is σ-finite, then $(L^1(S))^* = L^\infty(S)$ and $(L^1(S,\mathscr{B}))^* = L^\infty(S,\mathscr{B})$ isometrically, and for $f \in L^\infty(S) \cap L^2(S)$ we have $E_{\mathscr{B}}f = E_{\mathscr{B}}^*f \in L^\infty(S,\mathscr{B})$. It thus makes sense to define $E_{\mathscr{B}}f := E_{\mathscr{B}}^*f$ for $f \in L^\infty(S)$.

We can now prove the following stability result:

Lemma 1.3.17. *Suppose that X has the RNP with respect to a σ-finite measure space $(S,\mathscr{A},\mu)$, and let $\mathscr{B}$ be a sub-σ-algebra of $\mathscr{A}$. Then X has the RNP with respect to $(S,\mathscr{B},\mu|_{\mathscr{B}})$.*

Proof. Let $T : L^1(S,\mathscr{B}) \to X$ be a bounded operator. Let $E_{\mathscr{B}} : L^1(S) \to L^1(S,\mathscr{B})$ denote the conditional expectation operator. The operator $U := T \circ E_{\mathscr{B}} : L^1(S) \to X$ is bounded, and therefore representable since X has the RNP with respect to S, say by the function $\phi \in L^\infty(S;X)$. The operator $E_{\mathscr{B}}$ restricts to a bounded and positive operator from $L^\infty(S)$ to $L^\infty(S,\mathscr{B})$ and therefore by Theorem 2.1.7 it has a unique extension to a bounded and weak* continuous operator (in the sense of Theorem 2.1.7) from $L^\infty(S;X)$ to $L^\infty(S,\mathscr{B};X)$ (borrowing this result from the next section does not introduce any circularity in the argument). We claim that T is represented by $E_{\mathscr{B}}\phi$. Indeed, for all $B \in \mathscr{B}$ with $\mu(B) < \infty$ we have

$$T\mathbf{1}_B = U\mathbf{1}_B = \int_S \mathbf{1}_B\phi \,\mathrm{d}\mu = \int_S E_{\mathscr{B}}\mathbf{1}_B\phi \,\mathrm{d}\mu = \int_S \mathbf{1}_B E_{\mathscr{B}}\phi \,\mathrm{d}\mu.$$

In this computation, the final identity is immediate if we replace ϕ by an X-valued μ-simple function, and the identity for ϕ follows from it by approximation and dominated convergence. Hence $Tf = \int_S f E_{\mathscr{B}}\phi \,\mathrm{d}\mu$ for all μ-simple functions f, and the general case follows by approximation. □

The next result shows that the Radon–Nikodým property is *separably determined.*

Theorem 1.3.18. *Let $(S,\mathscr{A},\mu)$ be a σ-finite measure space. For a Banach space X the following assertions are equivalent:*

(1) *X has the RNP with respect to $(S,\mathscr{A},\mu)$;*
(2) *every closed separable subspace has the RNP with respect to $(S,\mathscr{A},\mu)$.*

For the proof we need the following lemma.

Lemma 1.3.19. *Let $(S,\mathscr{A},\mu)$ be a σ-finite measure space. If $\phi \in L^1(S;X)$, then the set $\{\int_A \phi \,\mathrm{d}\mu : A \in \mathscr{A}\}$ is relatively compact in X.*

Proof. Given $\varepsilon > 0$, let $\psi = \sum_{n=1}^{N} \mathbf{1}_{B_n} \otimes x_n$ be a μ-simple function such that $\|\phi - \psi\|_1 < \varepsilon$. Then for all $A \in \mathscr{A}$,

$$\int_A \phi \,\mathrm{d}\mu = \sum_{n=1}^{N} \mu(A \cap B_n) x_n + \int_A (\phi - \psi) \,\mathrm{d}\mu \in K_\varepsilon + \varepsilon \overline{B_X},$$

where

$$K_\varepsilon = \Big\{ \sum_{n=1}^{N} c_n x_n : c_n \geqslant 0, \sum_{n=1}^{N} |c_n| \, \|x_n\| \leqslant \|\phi\|_1 + \varepsilon \Big\}$$

is compact as a closed subset of $(\|\phi\|_1 + \varepsilon)\overline{B_Y}$, where $Y = \operatorname{span}\{x_1, \dots, x_N\}$ is a finite-dimensional Banach space. This proves that $\{\int_A \phi \,\mathrm{d}\mu : A \in \mathscr{A}\}$ is totally bounded, hence relatively compact. □

We shall actually use this in the form:

Lemma 1.3.20. *Let $(S, \mathscr{A}, \mu)$ be a finite measure space. If $T : L^1(S) \to X$ is representable, then the set $\{T\mathbf{1}_A : A \in \mathscr{A}\}$ is relatively compact in X.*

Proof. Since T is representable, we have $T\mathbf{1}_A = \int_A \phi \,\mathrm{d}\mu$ for some $\phi \in L^\infty(S; X) \subseteq L^1(S; X)$, using the finiteness of the measure space in the last step. Thus Lemma 1.3.19 applies. □

Proof of Theorem 1.3.18. (1)⇒(2): Suppose that X has the RNP with respect to $(S, \mathscr{A}, \mu)$, and let Y be a closed subspace. For any Y-valued measure F of bounded variation on $\mathscr{A}$ that is absolutely continuous with respect to μ, the RNP of X gives us a function $\phi \in L^1(S; X)$ such that

$$F(A) = \int_A \phi \,\mathrm{d}\mu, \quad A \in \mathscr{A}.$$

As F takes values in Y, we see that $\int_A \phi \,\mathrm{d}\mu$ belongs to Y for all $S \in \mathscr{A}$. Therefore ϕ takes values in Y almost everywhere by Proposition 1.2.13, which shows the RNP of Y.

(2)⇒(1): Suppose that every closed separable subspace of X has the RNP with respect to $(S, \mathscr{A}, \mu)$ and let $T : L^1(S) \to X$ be a bounded operator. We claim that T takes its values in a separable closed subspace Y of X. Once this has been proved, it follows from the assumed RNP of Y and Theorem 1.3.16 that T is representable as a bounded operator from $L^1(S)$ to Y, and hence as a bounded operator from $L^1(S)$ to X. Another application of Theorem 1.3.16 finishes the proof.

Fix an exhausting sequence $(S^{(n)})_{n \geqslant 1}$. Suppose, for the moment, that for some $n \geqslant 1$ the set $\{T\mathbf{1}_{A \cap S^{(n)}} : A \in \mathscr{A}\}$ fails to be relatively compact. Then we can find sets $A_j \in \mathscr{A}$ such that the sequence $(T\mathbf{1}_{A_j \cap S^{(n)}})_{j \geqslant 1}$ does not have a convergent subsequence in X. Let $\mathscr{A}^{(n)}$ denote the σ-algebra generated by the sets $A_j \cap S^{(n)}$, $j \geqslant 1$. By Proposition 1.2.29 the space $L^1(S, \mathscr{A}^{(n)})$ is separable, and therefore the restriction of T to this space has its range in a separable closed subspace $X^{(n)}$ of X. Since $X^{(n)}$ has the RNP with respect to $(S, \mathscr{A}, \mu)$ by the assumption of the theorem, and therefore with respect to $(S, \mathscr{A}^{(n)}, \mu|_{\mathscr{A}^{(n)}})$ by Lemma 1.3.17, this restricted operator is representable. But then Lemma 1.3.20 implies that the set $\{T\mathbf{1}_{A_j \cap S^{(n)}} : j \geqslant 1\}$ is relatively compact. In particular, it has a convergent subsequence. This contradiction proves that each of the sets $\{T\mathbf{1}_{A \cap S^{(n)}} : A \in \mathscr{A}\}$ is relatively compact.

Now we are ready to prove the claim. As relatively compact sets are separable, by what we just proved there exists a separable closed subspace Y of X containing $T\mathbf{1}_{A \cap S^{(n)}}$ for all $n \geqslant 1$ and $A \in \mathscr{A}$. By approximation, Y contains $T\mathbf{1}_A$ for all $A \in \mathscr{A}$ with $\mu(A) < \infty$. By another approximation argument, this implies that Y contains Tf for all $f \in L^1(S)$. □

We are finally in a position to give non-trivial examples of spaces with the RNP. Recall that a Banach space X is *reflexive* if the natural embedding of X into its bi-dual X^{**} is surjective (see Appendix B).

Theorem 1.3.21. *Each of the following conditions implies that X has the RNP with respect to every σ-finite measure space:*

(1) *X is a separable dual space;*
(2) *X is reflexive.*

Proof. Let $(S, \mathscr{A}, \mu)$ be a σ-finite measure space.

(1): By assumption there is a separable Banach space Y such that $X = Y^*$. We will check that every bounded operator $T : L^1(S) \to Y^*$ is representable.

For every $y \in Y$,

$$\langle y, T \rangle f := \langle y, Tf \rangle$$

defines a bounded linear functional $\langle y, T \rangle \in (L^1(S))^*$ of norm $\|\langle y, T \rangle\| \leqslant \|T\|\|y\|$. In view of $(L^1(S))^* = L^\infty(S)$ there exists a unique function $\phi_y \in L^\infty(S)$ of norm

$$\|\phi_y\|_\infty \leqslant \|T\|\|y\| \tag{1.9}$$

such that

$$\langle y, T \rangle f = \int_S \phi_y f \, \mathrm{d}\mu, \quad f \in L^1(S). \tag{1.10}$$

The separability of Y^* implies the separability of Y by Proposition B.1.9. Choose a countable dense sequence $(y_n)_{n \geqslant 1}$ in Y. If $y = \sum_{n=1}^N q_n y_n$ with the scalars $q_n \in \mathbb{Q}$ (if $\mathbb{K} = \mathbb{R}$) or $q_n \in \mathbb{Q} + i\mathbb{Q}$ (if $\mathbb{K} = \mathbb{C}$), then

$$\langle y, T \rangle f = \sum_{n=1}^N q_n \langle y_n, T \rangle f = \int_S \sum_{n=1}^N q_n \phi_{y_n} f \, \mathrm{d}\mu, \quad f \in L^1(S). \tag{1.11}$$

Comparing (1.10) and (1.11) we see that

$$\phi_y(s) = \sum_{n=1}^N q_n \phi_{y_n}(s) \tag{1.12}$$

for almost all $s \in S$. Since there are only countably many elements y of the above form, there exists a null set $N \subseteq S$ such that (1.12) holds for all $s \in S \setminus N$ and all $y = \sum_{n=1}^{N} q_n y_n$ of the above form. Now (1.9) and (1.12) imply that for almost all $s \in S$ the mapping $y \mapsto \phi_y(s)$ extends to a linear functional $\phi(s) \in Y^*$ of norm $\|\phi(s)\| \leqslant \|T\|$. From $\langle y, \phi(s)\rangle = \phi_y(s)$ we see that the resulting function $s \mapsto \phi(s)$ is weak* measurable. Because Y^* is separable, we may apply the Pettis measurability theorem to conclude that $s \mapsto \phi(s)$ is strongly measurable. It follows that $\phi \in L^\infty(S; Y^*)$ with norm $\|\phi\|_\infty \leqslant \|T\|$. The proof is concluded by noting that for all $y \in Y$,

$$\langle y, Tf\rangle = \langle y, T\rangle f = \int_S \langle y, \phi\rangle f \, \mathrm{d}\mu,$$

so T is represented by ϕ.

(2): Every separable closed subspace of X is a separable dual space and therefore has the RNP by (1). Now (2) follows by an appeal to Theorem 1.3.18. □

Corollary 1.3.22. *Let $(S, \mathscr{A}, \mu)$ be a σ-finite measure space and let X be reflexive or X^* be separable. Then for all $1 \leqslant p < \infty$ we have an isometric isomorphism*

$$(L^p(S; X))^* = L^q(S; X^*), \quad \tfrac{1}{p} + \tfrac{1}{q} = 1.$$

The σ-finiteness assumption is redundant for $1 < p < \infty$.

Example 1.3.23. Theorem 1.3.21 shows that the spaces l^p $(1 \leqslant p < \infty)$ and $L^p(S)$ $(1 < p < \infty)$ have the RNP.

We proceed with some examples of Banach space which fail to have the RNP.

Example 1.3.24. The spaces c_0, ℓ^∞, $C[0,1]$, $L^\infty(0,1)$ fail the RNP (with respect to the unit interval $([0,1], \mathscr{B}([0,1]), \lambda)$, with λ the Lebesgue measure). To see that c_0 fails the RNP consider the operator $T : L^1(0,1) \to c_0$ given by

$$(Tf)_n := \int_0^1 f(t) \sin(2\pi n t) \, \mathrm{d}t.$$

Suppose, for a contradiction, that T is representable by a function $\phi \in L^\infty(0,1; c_0)$. This would mean that for all indices n we have

$$\int_0^1 f(t) \sin(2\pi n t) \, \mathrm{d}t = (Tf)_n = \int_0^1 f(t) \phi_n(t) \, \mathrm{d}t$$

for all $f \in L^1(0,1)$, and therefore

$$\phi_n(t) = \sin(2\pi n t)$$

for almost all $t \in (0,1)$. For all irrational $t \in (0,1)$ we have $\phi(t) \in \ell^\infty \setminus c_0$. This is the desired contradiction, since it precludes ϕ from being in $L^\infty(0,1; c_0)$.

Since c_0 is contained in ℓ^∞, in $C[0,1]$, and in $L^\infty(0,1)$ as a closed subspace (for $C[0,1]$ is this seen by noting that the closed linear span of any normalised sequence of positive functions with disjoint supports is isometrically isomorphic to c_0), it follows that these spaces also fail the RNP.

Example 1.3.25. The space $L^1(0,1)$ fails the RNP. To see this we show that the identity operator $I : L^1(0,1) \to L^1(0,1)$ fails to be representable. Otherwise, there would exist a function $\phi \in L^\infty(0,1;L^1(0,1))$ such that

$$f = \int_0^1 f(t)\phi(t)\,\mathrm{d}t, \quad f \in L^1(0,1).$$

In particular, for all $a \in (0,1)$ and small $h > 0$,

$$\frac{1}{2h}\mathbf{1}_{(a-h,a+h)} = \frac{1}{2h}\int_{a-h}^{a+h} \phi(t)\,\mathrm{d}t.$$

By the Lebesgue differentiation theorem 2.3.4, for almost all $a \in (0,1)$

$$\Big\|\frac{1}{2h}\int_{a-h}^{a+h} \phi(t)\,\mathrm{d}t - \phi(a)\Big\|_{L^1(0,1)} \leqslant \frac{1}{2h}\int_{a-h}^{a+h} \|\phi(t) - \phi(a)\|_{L^1(0,1)}\,\mathrm{d}t.$$

converges to zero as $h \to 0$. But clearly, $\frac{1}{2h}\mathbf{1}_{(a-h,a+h)}$ fails to converges in $L^1(0,1)$ for all $a \in (0,1)$.

Universality of the unit interval $[0,1]$

The definition of the RNP as given in Definition 1.3.9 is somewhat unsatisfactory from a geometric point of view in that it does not only depend on X but also contains an explicit reference to the measure space $(S,\mathscr{A},\mu)$. This issue is fixed by the following:

Theorem 1.3.26. *For a Banach space X the following assertions are equivalent:*

(1) *X has the RNP with respect to the unit interval* $[0,1]$*;*
(2) *X has the RNP with respect to every σ-finite measure space* $(S,\mathscr{A},\mu)$*.*

This leads to the following addendum to Definition 1.3.9.

Definition 1.3.27. *A Banach space X* has the Radon–Nikodým property *if it has the Radon–Nikodým property with respect to* $[0,1]$.

A finite measure space $(S,\mathscr{A},\mu)$ is called *divisible* if for all real numbers $0 < t < 1$ and all sets $A \in \mathscr{A}$ there exist two disjoint sets $A_0, A_1 \in \mathscr{A}$ satisfying $A_0 \subseteq A$, $A_1 \subseteq A$, and

$$\mu(A_0) = (1-t)\mu(A), \quad \mu(A_1) = t\mu(A).$$

A σ-finite measure is divisible if and only if it is non-atomic. This result, which is proved in Appendix A, will not need this fact here.

Lemma 1.3.28. *Let $(S, \mathscr{A}, \mu)$ be a divisible and countably generated probability space. Let I be the set of all finite strings of 0's and 1's. There exists a family $(A_{\epsilon_1\dots\epsilon_n})_{\epsilon_1\dots\epsilon_n\in I}$ of sets in $\mathscr{A}$ with the following properties:*

(i) *for all $\epsilon_1 \dots \epsilon_n \in I$ we have*

$$\mu(A_{\epsilon_1\dots\epsilon_n}) = 2^{-n};$$

(ii) *for all $\epsilon_1 \dots \epsilon_n \in I$ we have a disjoint union*

$$A_{\epsilon_1\dots\epsilon_n} = A_{\epsilon_1\dots\epsilon_n 0} \cup A_{\epsilon_1\dots\epsilon_n 1};$$

(iii) *the family $(A_{\epsilon_1\dots\epsilon_n})_{\epsilon_1\dots\epsilon_n\in I}$ generates the σ-algebra $\mathscr{A}$.*

Proof. The proof is based on an intuitively simple construction. We only present an outline; the rigorous details are somewhat tedious and can be safely left to the reader.

Let $(A_n)_{n\geqslant 1}$ be a generating sequence in $\mathscr{A}$. We may assume that the sets A_n have non-zero measure and satisfy $\bigcup_{n\geqslant 1} A_n = S$. Using the divisibility property inductively, we may write $A_1 = \bigcup_{m\geqslant 1} A_{1m}$, where the sets $A_{1m} \in \mathscr{A}$ are disjoint and have measure $2^{-k_{1m}}$ for some integers $k_{1m} \geqslant 0$. Let us call these sets the *sets of the first generation* and write $\mathscr{G}_1 = \{A_{1m} : m \geqslant 1\}$.

Next we consider the sets $G \cap A_2$ for sets $G \in \mathscr{G}_1$. Using the divisibility property again, each of these intersections may be written as a disjoint union of sets in $\mathscr{A}$ of measure 2^{-k} for suitable integers $k \geqslant 0$. The resulting sets will be called the *sets of the second generation*. The collection of all sets of the second generation is denoted by $\mathscr{G}_2$.

Proceeding inductively we obtain generations $\mathscr{G}_n$ for each $n = 1, 2, \dots$. By their very construction, if A and A' are sets from $\mathscr{G}_m$ and $\mathscr{G}_n$ with $m \leqslant n$, respectively, then either $A' \subseteq A$ or $A \cap A' = \varnothing$.

From this it is easy to see that $\mathscr{G} := \bigcup_{n\geqslant 1} \mathscr{G}_n$ may be completed (using divisibility) to a 'dyadic' collection $\mathscr{H}$ containing, for each $n = 0, 1, \dots,$ precisely 2^n sets of measure 2^{-n} in such a way that each set of measure 2^{-n} in $\mathscr{H}$ is the union of two disjoint sets of measure 2^{-n-1} in $\mathscr{H}$. The σ-algebra generated by $\mathscr{H}$ contains the sets A_n and therefore it equals $\mathscr{A}$. It is clear that these sets can be label ed so as to have the desired properties. □

Let $(S, \mathscr{A}, \mu)$ be a finite measure space. We define an equivalence relation in $\mathscr{A}$ by declaring two sets A_1 and A_2 to be *μ-equivalent* if $\mu(A_1 \Delta A_2) = 0$, where $A_1 \Delta A_2 = (A_1 \cup A_2) \setminus (A_1 \cap A_2) = (A_1 \cup A_2) \setminus (A_1 \cap A_2)$ is the symmetric difference of A_1 and A_2. Denoting the equivalence class of a set A by $[A]$, the set

$$[\mathscr{A}] := \{[A] : A \in \mathscr{A}\}$$

is a metric space with respect to the distance function

$$\mathrm{dist}([A_1], [A_2]) = \mu(A_1 \Delta A_2).$$

The resulting metric space $[\mathscr{A}]$ is called the *measure algebra* associated with $(S, \mathscr{A}, \mu)$. It is easy to see that the set operations pass on to $[\mathscr{A}]$ by putting

$$[A_1] \cap [A_2] := [A_1 \cap A_2], \quad [A_1] \cup [A_2] := [A_1 \cup A_2], \quad \complement[A] := [\complement A].$$

We may regard μ as a set function on $[\mathscr{A}]$ by putting

$$\mu([A]) := \mu(A).$$

Lemma 1.3.29. *Let $(S, \mathscr{A}, \mu)$ and $(T, \mathscr{B}, \nu)$ be divisible and countably generated probability spaces. There exists a distance preserving bijection $j : [\mathscr{A}] \to [\mathscr{B}]$ preserving intersections, unions, and complements, and satisfying*

$$\mu(A) = \nu(B) \quad \textit{whenever} \quad j([A]) = [B]$$

with $A \in \mathscr{A}$ and $B \in \mathscr{B}$. As a consequence, j induces an isometric isomorphism of Banach spaces

$$j : L^1(S) \simeq L^1(T)$$

by defining $j(\mathbf{1}_A) := \mathbf{1}_B$ whenever $j([A]) = [B]$.

Proof. Choose families $(A_{\epsilon_1 \dots \epsilon_n})_{\epsilon_1 \dots \epsilon_n \in I}$ in $\mathscr{A}$ and $(B_{\epsilon_1 \dots \epsilon_n})_{\epsilon_1 \dots \epsilon_n \in I}$ in $\mathscr{B}$ with the properties stated in the lemma. Now define

$$j([A_{\epsilon_1 \dots \epsilon_n}]) := [B_{\epsilon_1 \dots \epsilon_n}]$$

and extend this mapping to finite unions of disjoint sets in the natural manner. By property (3) of the preceding lemma, the collection of all such unions of sets of the form $A_{\epsilon_1 \dots \epsilon_n}$ defines a dense set in the measure algebra associated with $(S, \mathscr{A}, \mu)$, and a similar assertion holds for $(T, \mathscr{B}, \nu)$. Moreover, j is an isometry between these dense subsets. It follows that j has a unique extension to an isometry between the full measure algebras, whence the result follows. □

Proof of Theorem 1.3.26. We only need to prove that (1) implies (2), the converse being trivial.

Step 1 – We prove that if X has the RNP with respect to $[0, 1]$, then X has the RNP with respect to every *divisible, countably generated, finite* measure space.

Let $(S, \mathscr{A}, \mu)$ be such a measure space. Evidently we may assume that $\mu(S) = 1$. Let $j : [\mathscr{A}] \to [\mathscr{B}]$ be the mapping of Lemma 1.3.29, where $\mathscr{B}$ is the Borel σ-algebra of $[0, 1]$, and let $j : L^1(S) \to L^1(0, 1)$ be the induced isometric isomorphism.

Suppose that $T : L^1(S) \to X$ is a bounded operator. Then the operator $\widetilde{T} : L^1(0, 1) \to X$ defined by $\widetilde{T} = T \circ j^{-1}$ is bounded and therefore representable by a function $\phi \in L^\infty(0, 1; X)$. Let $\psi \in L^\infty(S; X)$ be defined by $\psi = k\phi$, where $k : L^\infty(0, 1; X) \to L^\infty(S; X)$ is the unique bounded operator such that

$$k\Big(\sum_{n\geqslant 1} \mathbf{1}_{B_n} \otimes x_n\Big) = \sum_{n\geqslant 1} \mathbf{1}_{A_n} \otimes x_n$$

for bounded sequences $(x_n)_{n\geqslant 1}$ in X and sequences of disjoint sets $(A_n)_{n\geqslant 1}$ and $(B_n)_{n\geqslant 1}$ such that $j([A_n]) = [B_n]$ for all $n \geqslant 1$. For all $f \in L^1(S)$,

$$Tf = (\widetilde{T} \circ j)f = \int_0^1 (jf)(t)\phi(t)\,\mathrm{d}t = \int_S fk\phi\,\mathrm{d}\mu = \int_S f\psi\,\mathrm{d}\mu,$$

where the middle identity is checked by approximating f and ϕ by countably-valued functions and using the properties of j and k. This shows that T is represented by ψ. From Theorem 1.3.10 we conclude that X has the RNP with respect to $(S, \mathscr{A}, \mu)$.

Step 2 – We prove that if X has the RNP with respect to every divisible, countably generated, finite measure space, then X has the RNP with respect to every *countably generated and finite* measure space.

Let $(S, \mathscr{A}, \mu)$ be a countably generated and finite measure space. The measure space $(\widetilde{S}, \widetilde{\mathscr{A}}, \widetilde{\mu})$ is countably generated, finite, and divisible, where

$$\widetilde{S} = S \times [0,1], \quad \widetilde{\mathscr{A}} = \mathscr{A} \times \mathscr{B}([0,1]),\ \widetilde{\mu} = \mu \times \lambda$$

with λ the Lebesgue measure on $[0,1]$. Suppose that $T : L^1(S) \to X$ is a bounded operator. Define $r : L^1(\widetilde{S}) \to L^1(S)$ by

$$(rf)(s) = \int_0^1 f(s,t)\,\mathrm{d}t$$

and define $\widetilde{T} : L^1(\widetilde{S}) \to X$ by $\widetilde{T} := T \circ r$. Since X has the RNP with respect to $(\widetilde{S}, \widetilde{\mathscr{A}}, \widetilde{\mu})$, this operator is represented by a function $\phi \in L^\infty(\widetilde{S}; X)$. For $f \in L^1(S)$ put $\widetilde{f}(s,t) := f(s)$. Then $\widetilde{f} \in L^1(\widetilde{S})$, $r\widetilde{f} = f$, and

$$\begin{aligned} Tf = Tr\widetilde{f} = \widetilde{T}\widetilde{f} &= \int_0^1 \int_S \widetilde{f}(s,t)\phi(s,t)\,\mathrm{d}\mu(s)\,\mathrm{d}t \\ &= \int_S f(s) \int_0^1 \phi(s,t)\,\mathrm{d}t\,\mathrm{d}\mu(s) = \int_S f(s)\psi(s)\,\mathrm{d}\mu(s), \end{aligned}$$

where the function $\psi \in L^\infty(S; X)$ is defined by $\psi(s) = \int_0^1 \phi(s,t)\,\mathrm{d}t$. It follows that T is represented by ψ.

Step 3 – We prove that if X has the RNP with respect to every countably generated and finite measure space, then X has the RNP with respect to every *finite* measure space.

Let $(S, \mathscr{A}, \mu)$ be a finite measure space. In order to show that X has the RNP with respect to $(S, \mathscr{A}, \mu)$, thanks to Theorem 1.3.18 we may assume that X is separable. In view of Proposition B.1.10 this allows us to pick a sequence $(x_n^*)_{n\geqslant 1}$ in the unit ball B_{X^*} which is norming for X.

Let $T : L^1(S) \to X$ be a bounded operator and denote by $\widetilde{T} : L^2(S) \to X$ its restriction. Choose a measurable representative for each function $\widetilde{T}^* x_n^*$. By approximation with μ-simple functions one sees that the σ-algebra $\mathscr{B}$ generated by these representatives is countably generated. Let $E_{\mathscr{B}}$ be the orthogonal projection in $L^2(S)$ onto $L^2(S, \mathscr{B})$. For any $f \in L^2(S)$ we find

$$\langle \widetilde{T} f, x_n^* \rangle = \langle f, \widetilde{T}^* x_n^* \rangle = \langle f, E_{\mathscr{B}}^* \widetilde{T}^* x_n^* \rangle = \langle E_{\mathscr{B}} f, \widetilde{T}^* x_n^* \rangle = \langle \widetilde{T} E_{\mathscr{B}} f, x_n^* \rangle.$$

By Corollary 1.1.25 this implies that $\widetilde{T} f = \widetilde{T} E_{\mathscr{B}} f$ for all $f \in L^2(S)$.

Since X has the RNP with respect to $(S, \mathscr{B}, \mu|_{\mathscr{B}})$, the operator $T|_{L^1(S,\mathscr{B};X)}$ is representable by a function $\phi \in L^1(S, \mathscr{B}; X)$. For $f \in L^2(S)$ we then obtain, using the properties of conditional expectations that were explained preceding the proof of Lemma 1.3.17,

$$Tf = \widetilde{T} f = \widetilde{T} E_{\mathscr{B}} f = \int_S \phi(E_{\mathscr{B}} f) \,\mathrm{d}\mu = \int_S \phi f \,\mathrm{d}\mu,$$

where $E_{\mathscr{B}}$ denotes the extension of the conditional expectation to $L^1(S; X)$ whose existence is guaranteed by Theorem 2.1.3 (borrowing this result does not introduce any circularity). By density, the identity $Tf = \int_S \phi f \,\mathrm{d}\mu$ extends to functions $f \in L^1(S)$. This shows that ϕ also represents T.

Step 4 – To conclude the proof we note that if X has the RNP with respect to every finite measure space, then X has the RNP with respect to every σ-*finite* measure space. This is an easy application of Theorem 1.3.10, noting that we may write $S = \bigcup_{n \geqslant 1} S^{(n)}$, where $(S^{(n)})_{n \geqslant 1}$ is an exhausting sequence. By restricting a given operator $T : L^1(S) \to X$ to the closed subspaces $L^1(S^{(n)})$ we obtain representing functions $\phi_n \in L^\infty(S^{(n)}; X)$ satisfying

$$\|\phi_n\|_{L^\infty(S^{(n)};X)} = \|T|_{L^1(S^{(n)})}\| \leqslant \|T\|.$$

It is easy to check that the functions ϕ_n satisfy $\phi_n|_{S^{(m)}} = \phi_m$ whenever $m \leqslant n$. It follows that $\phi(s) := \phi_n(s)$ for $s \in S^{(n)}$ defines a function $\phi \in L^\infty(S; X)$ and $\|\phi\|_{L^\infty(S;X)} \leqslant \|T\|$. This function represents T on the dense linear span in $L^1(S)$ of the subspaces $L^1(S_n)$. Therefore it represents T on $L^1(S)$. □

1.4 Notes

Discussions of measurability and integration in Banach spaces can be found in the monographs Bogachev [1998], Diestel and Uhl [1977], Dinculeanu [2000], Dunford and Schwartz [1958], and Vakhania, Tarieladze, and Chobanyan [1987].

Section 1.1

The Pettis measurability theorem goes back to Pettis [1938], where Bochner [1933b] is cited for having introduced the notion of strong measurability.

Without proof we mention the following more general version of the Pettis measurability theorem valid for functions taking values in a metric space; see Vakhania, Tarieladze, and Chobanyan [1987, Propositions I.1.9 and I.1.10]. Strong measurability of such functions is defined is the same way as in the main text, namely in terms of pointwise approximation by simple functions.

Proposition 1.4.1. *Let $(S, \mathscr{A})$ be a measurable space, X be a complete metric space, Γ be a set of real-valued continuous functions defined on X and separating the points of X and let $f : S \to X$ be a function with separable range. Then the following conditions are equivalent:*

(1) *f is measurable;*
(2) *f is strongly measurable;*
(3) *$\phi \circ f$ is measurable for all $\phi \in \Gamma$.*

The next result is a partial converse to Proposition 1.1.1.

Proposition 1.4.2. *Let X be a Banach space. If $\sigma(X^*) = \mathscr{B}(X)$, then X^* contains a countable set which separates the points of X.*

Proof. As a preliminary step, let us call a set $A \in \sigma(X^*)$ *countably determined* if there is a countable set $Y \subseteq X^*$ such that $A \in \sigma(Y)$. The countably determined sets are easily seen to form a σ-algebra. Clearly, this σ-algebra contains every set of the form

$$\big\{x \in X : \ \big(\langle x, x_1^*\rangle, \dots, \langle x, x_n^*\rangle\big) \in B\big\} \tag{1.13}$$

with $B \in \mathscr{B}(\mathbb{K}^n)$, and therefore it is in fact equal to $\sigma(X^*)$. Thus every set in $\sigma(X^*)$ is countably determined.

Now suppose that no countable set in X^* separates the points of X. Let $A \in \sigma(X^*)$ be an arbitrary non-empty set and choose a countable set $Y \subseteq X^*$ such that $A \in \sigma(Y)$. By assumption, there exists a non-zero vector $x_0 \in X$ such that $\langle x_0, x^*\rangle = 0$ for all $x^* \in Y$. We claim that if $x \in A$, then also $x + cx_0 \in A$ for all $c \in \mathbb{K}$. Indeed, this is obvious if A is of the form (1.13) with $x_1^*, \dots, x_n^* \in Y$, and these sets generate $\sigma(Y)$. The general case then follows from the observation that the sets having the property described by the claim (with the given Y) form a σ-algebra.

The claim implies that every non-empty set in $\sigma(X^*)$ is unbounded. In particular $\sigma(X^*)$ must be a proper subset of $\mathscr{B}(X)$. □

We refer to Vakhania, Tarieladze, and Chobanyan [1987] for more on this. As an immediate corollary one finds that if X is reflexive, then $\sigma(X^*) = \mathscr{B}(X)$ implies that X is separable, an observation that is credited to S. Okada in Fremlin [1980]. In this paper it is shown that without reflexivity the converse does not hold: the non-separable Banach space $\ell^1(\aleph_1)$, with $\aleph_1$ the first uncountable cardinal number, also satisfies $\sigma((\ell^1(\aleph_1))^*) = \mathscr{B}(\ell^1(\aleph_1))$.

The following example from Edgar [1977, 1979] shows that the inclusion $\sigma(X^*) \subseteq \mathscr{B}(X)$ may indeed be proper; see also Talagrand [1978, 1984].

Example 1.4.3. For any uncountable set I we have $\sigma((\ell^2(I))^*) \subsetneq \mathscr{B}(\ell^2(I))$. Indeed, identifying $\ell^2(I)$ and its dual via the Riesz representation theorem, every $x^* \in (\ell^2(I))^*$ has countable support in I. Hence if $Y \subseteq (\ell^2(I))^*$ is countable, it is countably supported, say in the countable set J. But then for any $i \in \complement J$ we have $\langle \mathbf{1}_{\{i\}}, x^* \rangle = 0$ for all $x^* \in Y$. This argument shows that $(\ell^2(I))^*$ contains no countable set separating the points of $\ell^2(I)$.

Further references on measure theory on Banach spaces and more general topological vector spaces include Bogachev [2007b] and Schwartz [1973].

The Borel σ-algebra of a separable metric space E contains at most ω_1 sets, where ω_1 is the smallest uncountable ordinal. This fact was used in Lemma 1.1.12 and can be proved by transfinite induction. The idea is to prove that $\mathscr{B}(E) = \bigcup_{\alpha<\omega_1} \mathscr{G}_\alpha$, where $\mathscr{G}_0$ is the (countable) family of all open sets (every open set is a countable union of balls with rational radius and centre from a countable dense set), and $\mathscr{G}_\alpha$ consists of all countable intersections of sets from $\bigcup_{\beta<\alpha} \mathscr{G}_\beta$ if α is an odd ordinal, all countable unions of sets from $\bigcup_{\beta<\alpha} \mathscr{G}_\beta$ if α is an even ordinal, and $G_\alpha = \bigcup_{\beta<\alpha} \mathscr{G}_\beta$ if α is a limit ordinal; an odd (even) ordinal is one of the form $\lambda + n$ with λ a limit ordinal and n an odd (even) integer. The details can be found in Hewitt and Stromberg [1975, Theorem 10.23].

Section 1.2

The Bochner integral was introduced in Bochner [1933b] and Dunford [1935]. Early expositions are contained in Dunford and Schwartz [1958] and Hille and Phillips [1957]. A more recent presentation is offered in Diestel and Uhl [1977].

Theorem 1.2.4 is from Hille and Phillips [1957].

In the context of scalar-valued functions, the recognition of convexity in general and Jensen's inequality (Proposition 1.2.11) in particular goes back to the work of Jensen [1906].

The Pettis integral was introduced in Pettis [1938], which also contains Theorem 1.2.37. Modern treatments can be found in Diestel and Uhl [1977], Van Dulst [1989], Musiał [2002], Talagrand [1984]. Dilworth and Girardi [1993] show that in any infinite-dimensional Banach space X there exists a strongly measurable Pettis integrable function that fails to be Bochner integrable. Further examples concerning the difference between the Pettis integral and Bochner integral may be found in Dilworth and Girardi [1995]. Theorem 1.2.40 was proved in Bessaga and Pełczyński [1958]; our presentation follows Arendt, Batty, Hieber, and Neubrander [2011].

Proposition 1.3.3 is due to Lions and Peetre [1964].

$L^1(S;X)$ as a projective tensor product

The *projective tensor product* of two Banach spaces X and Y is the completion $X \widehat{\otimes} Y$ of the algebraic tensor product $X \otimes Y$ with respect to the norm

$$\|u\|_{X\widehat{\otimes}Y} := \inf\Big\{\sum_{n=1}^N \|x_n\|\|y_n\| : \ u = \sum_{n=1}^N x_n\otimes y_n\Big\}.$$

Now let $(S,\mathscr{A},\mu)$ be a measure space and X be a Banach space.

Proposition 1.4.4. *The identity mapping on $L^1(S)\otimes X$ extends to an isometry*

$$L^1(S;X) \eqsim L^1(S)\,\widehat{\otimes}\,X.$$

Proof. Fix a μ-simple function $f = \sum_{n=1}^N \mathbf{1}_{A_n}\otimes x_n$, where $x_n\in X$ and the sets A_n are disjoint and of finite measure, and suppose that $f = \sum_{m=1}^M g_m\otimes y_m$ is another representation with $g_m\in L^1(S)$ and $y_m\in X$. Then,

$$\begin{aligned}
\Big\|\sum_{n=1}^N \mathbf{1}_{A_n}\otimes x_n\Big\|_{L^1(S;X)} &= \sum_{n=1}^N \Big\|\mathbf{1}_{A_n}\sum_{m=1}^M g_m\otimes y_m\Big\|_{L^1(S;X)} \\
&\leqslant \sum_{n=1}^N\sum_{m=1}^M \|\mathbf{1}_{A_n}g_m\otimes y_m\|_{L^1(S;X)} \\
&= \sum_{m=1}^M\sum_{n=1}^N \|\mathbf{1}_{A_n}g_m\|_{L^1(S)}\|y_m\| \leqslant \sum_{m=1}^M \|g_m\|_{L^1(S)}\|y_m\|.
\end{aligned}$$

This proves that $\|f\|_{L^1(S;X)} \leqslant \|f\|_{L^1(S)\widehat{\otimes}X}$. The reverse inequality is even easier:

$$\|f\|_{L^1(S)\widehat{\otimes}X} \leqslant \sum_{n=1}^N \|\mathbf{1}_{A_n}\|_{L^1(S)}\|x_n\| = \Big\|\sum_{n=1}^N \mathbf{1}_{A_n}\otimes x_n\Big\|_{L^1(S;X)} = \|f\|_{L^1(S;X)}.$$

This proves the equality of the projective tensor norm and the L^1-norm on the algebraic tensor product $L^1(S)\otimes X$. Since this space is dense in both $L^1(S)\,\widehat{\otimes}\,X$ and $L^1(S;X)$, the proof is complete. □

For a fuller development of this topic we refer to Diestel and Uhl [1977, Chapter 8], Defant and Floret [1993], and Diestel, Fourie, and Swart [2008].

Spaces of weakly measurable functions

One may define the space $L^1_{\rm w}(S;X)$ as the linear space of all functions $f: S\to X$ such that $\langle f,x^*\rangle$ is μ-measurable and belongs to $L^1(S)$ for all $x^*\in X^*$. Endowed with the norm

$$\|f\|_{L^1_{\rm w}(S;X)} := \sup_{\|x^*\|\leqslant 1} \|\langle f,x^*\rangle\|_{L^1(S)}$$

(which is finite by Lemma 1.2.18) this is a normed linear space. The space $P^1(S;X)$ of all Pettis integrable functions is easily seen to be a closed subspace of $L^1_{\mathrm{w}}(S;X)$. Indeed, if $f_n \to f$ in $L^1_{\mathrm{w}}(S;X)$ with each f_n in $P^1(S;X)$, then for all $A \in \mathscr{A}$ we have

$$\begin{aligned}\Big\|\tau(X,X^*)\text{-}\int_A f\,\mathrm{d}\mu - \tau(X,X^*)\text{-}\int_A f_n\,\mathrm{d}\mu\Big\|_{X^{**}} &= \sup_{\|x^*\|\leqslant 1}\Big|\int_A \langle f-f_n, x^*\rangle\,\mathrm{d}\mu\Big| \\ &\leqslant \|f-f_n\|_{L^1_{\mathrm{w}}(S;X)},\end{aligned}$$

and therefore the $\tau(X,X^*)$-integral of f over A (see Example 1.2.33) belongs to X. This means that $f \in P^1(S;X)$ as was to be shown.

By a theorem of Thomas [1976], $P^1(S;X)$ (and hence $L^1_{\mathrm{w}}(S;X)$) fails to be complete whenever X is infinite-dimensional. Alternative proofs of this result may be found in Dilworth and Girardi [1993] and Janicka and Kalton [1977]. For infinite-dimensional Hilbert spaces H, the non-completeness of $P^1(0,1;H)$ follows from the following elementary example due to Birkhoff [1935] (see also Pietsch [2007, 5.1.2.12]):

Example 1.4.5. Let H be a separable infinite-dimensional Hilbert space with a doubly-indexed orthonormal basis $(h_{k,n})_{n\geqslant 0,\,0\leqslant k\leqslant 2^n-1}$. Set

$$f_n := \sum_{k=0}^{2^n-1} \mathbf{1}_{I_{k,n}} \otimes h_{k,n}$$

with $I_{k,n} := (k2^{-n},(k+1)2^{-n})$. Then $f_n \in P^1(0,1;H)$ and by the Cauchy–Schwarz inequality,

$$\|f\|_{L^1_{\mathrm{w}}(S;H)} = \sup_{\|h\|\leqslant 1} 2^{-n}\sum_{k=0}^{2^n-1} |(h_{k,n}|h)| \leqslant 2^{-n/2}.$$

If we had convergence $\sum_{n\geqslant 0} f_n = f$ in $L^1_{\mathrm{w}}(S;H)$, then for almost all $t \in (0,1)$ it would follow that $(h_{k,n}|f(t)) = \mathbf{1}_{I_{k,n}}(t)$ for almost all $t \in (0,1)$. Since every $t \in (0,1)$ belongs to infinitely many $I_{k,n}$, this would mean that almost every value of f has infinitely many coordinates equal to 1, which is impossible.

Measurability of operator-valued functions

Proposition 1.1.28 and its corollary are due (under some additional restrictions) to Schlüchtermann [1995]; see also Badrikian, Johnson, and Yoo [1995], Johnson [1993].

The aim of the next proposition is to describe strong measurability of operator-valued functions in terms of the measurability of inverse images of open sets in the strong operator topology, thereby providing an analogue of Corollary 1.1.10. Let us denote by $\sigma_{\mathrm{so}}(X,Y)$ the σ-algebra generated by the strong operator topology of $\mathscr{L}(X,Y)$. This σ-algebra is generated by the open sets in this topology of the form

$$V_{T_0,x_0,\varepsilon_0} := \{T \in \mathscr{L}(X,Y):\ \|(T-T_0)x_0\| < \varepsilon_0\}$$

with $T_0 \in \mathscr{L}(X,Y)$, $x_0 \in X$, and $\varepsilon_0 > 0$.

Proposition 1.4.6. *For a function $f : S \to \mathscr{L}(X,Y)$ the following assertions are equivalent:*

(1) *f is $\sigma_{\rm so}(X,Y)$-measurable;*
(2) *$fx : S \to Y$ is measurable for all $x \in X$.*

The proof is an easy exercise in using the identity

$$f^{-1}(V_{T_0,x_0,\varepsilon_0}) = \{\xi \in S : \ \|(f(\xi) - T_0)x_0\| < \varepsilon_0\}$$

and is left to the reader. Combining this result with Corollary 1.1.10 we obtain the following extension of it to operator-valued functions.

Corollary 1.4.7. *For a function $f : S \to \mathscr{L}(X,Y)$ the following assertions are equivalent:*

(1) *$f : S \to \mathscr{L}(X,Y)$ is strongly measurable (in the sense of Definition 1.1.27);*
(2) *fx is separably valued for all $x \in X$ and f is $\sigma_{\rm so}(X,Y)$-measurable.*

The 'weak' analogue of this result holds equally well: a function $f : S \to \mathscr{L}(X,Y)$ is weakly measurable if and only if fx is separably valued for all $x \in X$ and f is $\sigma_{\rm wo}(X,Y)$-measurable. Here, f is said to be *weakly measurable* if $s \mapsto \langle fx, x^*\rangle$ is measurable for all $x \in X$ and $x^* \in X^*$, and $\sigma_{\rm wo}(X,Y)$ denotes the σ-algebra generated by the weak operator topology on $\mathscr{L}(X,Y)$. We refer the interested reader to Hille and Phillips [1957, Theorem 3.5.5].

Further results along these lines can be found in Blasco and Van Neerven [2010] and Fourie [2010].

Integrability of operator-valued functions

Next we turn to the problem of integration of strongly μ-measurable functions $f : S \to \mathscr{L}(X,Y)$ (in the sense of Definition 1.1.27).

For $1 \leqslant p \leqslant \infty$ let $L^p_{\rm so}(S;\mathscr{L}(X,Y))$ denote the space of all strongly μ-measurable functions $f : S \to \mathscr{L}(X,Y)$ with the property that for all $x \in X$ the function fx belongs to $L^p(S;Y)$; we identify f and g when $fx = gx$ in $L^p(S;Y)$ for all $x \in X$. For instance, if $\mathscr{T} \subseteq \mathscr{L}(X,Y)$ is bounded and $f : S \to \mathscr{L}(X,Y)$ is strongly μ-measurable and takes values in $\mathscr{T}$ almost everywhere, then $f \in L^\infty_{\rm so}(S;\mathscr{L}(X,Y))$.

For every $f \in L^p_{\rm so}(S;\mathscr{L}(X,Y))$, the operator $x \mapsto fx$ is closed from X into $L^p(S;Y)$. Hence, by the closed graph theorem, it is bounded. With respect to the norm

$$\|f\|_{L^p_{\rm so}(S;\mathscr{L}(X,Y))} := \sup_{\|x\| \leqslant 1} \|fx\|_{L^p(S;Y)},$$

$L^p_{\rm so}(S;\mathscr{L}(X,Y))$ is a normed space. In general it fails to be complete. Indeed, if X is an infinite-dimensional separable reflexive Banach space, it is immediate from the definitions that

$$L^1_{\rm so}(S;\mathscr{L}(X,\mathbb{K})) = L^1_{\rm w}(S;X^*),$$

and we have seen above that the latter space fails to be complete.

Section 1.3

The standard reference for the Radon–Nikodým property is Diestel and Uhl [1977]; they provide references to the pioneering papers of Bochner, Dunford, Pettis, and others, who referred to this property as "property (D)". Our presentation is somewhat different and also covers the σ-finite case. The idea to give a direct proof of Theorem 1.3.10 via Lemma 1.3.11, thus avoiding a reference to the scalar version of the Radon–Nikodým theorem, seems to be new. The proof of Theorem 1.3.16 follows the one in Diestel and Uhl [1977], except for the use of Lemma 1.3.11. The proof of the first part of Theorem 1.3.21 follows Arendt, Batty, Hieber, and Neubrander [2011].

The Radon–Nikodým property plays an important role in infinite-dimensional convexity, in particular in connection with the Krein–Milman theorem on extreme points and Choquet-type integral representation theorems. For this material we refer to Benyamini and Lindenstrauss [2000] and the literature quoted therein, in particular the lecture notes of Bourgin [1983] and the memoir by Ghoussoub, Godefroy, Maurey, and Schachermayer [1987].

Duality of vector-valued Hardy spaces

Theorem 1.3.10 has an analogue for Hardy spaces due to Blasco [1988]. In order to state the result we need the following terminology.

Definition 1.4.8. *A function $a \in L^2(\mathbb{T};X)$ is said to be an* atom *if it is either constant or there exists an interval $I \subseteq \mathbb{T}$ such that*

(i) $\operatorname{supp}(a) \subseteq I$*;*
(ii) $\int_{\mathbb{T}} f(t)\,\mathrm{d}(t) = 0$*;*
(iii) $\int_{\mathbb{T}} \|f(t)\|^2\,\mathrm{d}(t) \leqslant 2\pi/|I|$*.*

The *atomic Hardy space* space $H^1_{\mathrm{at}}(\mathbb{T};X)$ is defined as the space of all strongly measurable $f:\mathbb{T}\to X$ that can be represented an ℓ^1-sum of atoms, i.e.,

$$f = \sum_{n\geqslant 1} \lambda_n a_n$$

with each a_n an atom and $\sum_{n\geqslant 1}|\lambda_n| < \infty$. This space is a Banach space with respect to the norm

$$\|f\|_{H^1_{\mathrm{at}}(\mathbb{T};X)} := \inf\Big\{\sum_{n\geqslant 1}|\lambda_n| : \ f = \sum_{n\geqslant 1}\lambda_n a_n, \ \text{each } a_n \text{ an atom}\Big\}.$$

There are various alternative and equivalent ways to introduce this space; this topic will be taken up in the Notes of Chapter 5.

It is a celebrated result in harmonic analysis, due to Fefferman [1971], that the dual of $H^1_{\mathrm{at}}(\mathbb{T}) := H^1_{\mathrm{at}}(\mathbb{T};\mathbb{K})$ equals $\mathrm{BMO}(\mathbb{T}) := \mathrm{BMO}(\mathbb{T};\mathbb{K})$, the space of functions of *bounded mean oscillation*. We refer to Stein [1993] for a comprehensive account and references to the literature. In the vector-valued context, the space $\mathrm{BMO}(\mathbb{T};X)$ is defined as the space of all strongly measurable functions $f:\mathbb{T}\to X$ for which

$$\|f\|_{\mathrm{BMO}(\mathbb{T};X)} := \|f_{\mathbb{T}}\| + \sup_I \frac{1}{|I|}\int_I \|f(t)-f_I\|\,\mathrm{d}t$$

is finite, where the supremum is taken over all intervals $I\subseteq\mathbb{T}$ and $f_I := \frac{1}{|I|}\int_I f(t)\,\mathrm{d}t$ is the average of f over I.

The vector-valued extension of the H^1_{at}-BMO duality reads as follows:

Theorem 1.4.9. *For a Banach space X, the following assertions are equivalent:*

(1) $(H^1_{\mathrm{at}}(\mathbb{T};X))^* = \mathrm{BMO}(\mathbb{T};X^*)$ *with equivalent norms;*
(2) X^* *has the Radon–Nikodým property.*

For more information on vector-valued H^1 and BMO the reader may consult Pisier [2016].

2

Operators on Bochner spaces

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

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

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

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

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

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 63, DOI 10.1007/978-3-319-48520-1_2

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

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

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

2.1 The L^p-extension problem

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

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

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

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

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

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

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

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

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

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

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

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

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

if the isomorphism is isometric.

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

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

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

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

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

$$
\begin{aligned}
\|(T \otimes I_X)f\|^p_{L^p(S;X)} &= \int_{S'} \Big\| \sum_{n=1}^{N} Tg_n \otimes x_n(s')\Big\|^p_{L^p(S)} \,\mathrm{d}\mu'(s') \\
&\leqslant \|T\|^p \int_{S'} \Big\| \sum_{n=1}^{N} g_n \otimes x_n(s')\Big\|^p_{L^p(S)} \,\mathrm{d}\mu'(s') = \|T\|^p \|f\|^p_{L^p(S;X)}.
\end{aligned}
$$

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

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

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

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

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

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

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

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

with obvious modifications for $p_2 = \infty$. Since the μ_1-simple functions are dense in $L^{p_1}(S_1; X)$, this proves that $T \otimes I$ has a unique extension to a bounded operator from $L^{p_1}(S_1;X)$ to $L^{p_2}(S_2;X)$ of norm $\|T\otimes I_X\|\leqslant\|T\|$. Equality $\|T\otimes I_X\|=\|T\|$ is obtained by considering functions f of the form $g\otimes x$ with $g\in L^{p_1}(S_1)$ and $x\in X$ of norm one. $\square$

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

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

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

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

$$A_N=\big\{\|f-x_N\|<\varepsilon\big\}\cap\Big(\bigcap_{n=1}^{N-1}\big\{\|f-x_n\|\geqslant\varepsilon\big\}\in\mathscr{A}$$

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

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

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

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

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

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

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

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

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

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

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

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

by the dominated convergence theorem. $\square$

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

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

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

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

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

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

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

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

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

Gaussian random variables

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\Big\| \sum_{n\geqslant 1} a_n\gamma_n \Big\|_{L^p(\Omega)} = \|\gamma\|_p \|a\|_{\ell^2}. \tag{2.3}$$

As a consequence, the mapping

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

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

Extension theorem for Hilbert spaces

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

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

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

with

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

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

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

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

Now we consider two cases.

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

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

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

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

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

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

Extension results for complexifications of real Banach spaces

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

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

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

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

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

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

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

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

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

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

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

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

This gives the asserted norm identity. □

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

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

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

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

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

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

$$\begin{aligned}
\|\gamma\|_p\|T_{\mathbb{C}}(f+ig)\|_{L^p(S_2;Y_{\mathbb{C}}^{\gamma,p})} &= \|\gamma_1Tf+\gamma_2Tg\|_{L^p(S_2;L^p(\Omega;Y))}\\
&= \|T(\gamma_1f+\gamma_2g)\|_{L^p(\Omega;L^p(S_2;Y))}\\
&\leqslant \|T\|\,\|\gamma_1f+\gamma_2g\|_{L^p(\Omega;L^p(S_1;X))}\\
&= \|T\|\,\|\gamma_1f+\gamma_2g\|_{L^p(S_1;L^p(\Omega;X))}\\
&= \|\gamma\|_p\|T\|\,\|f+ig\|_{L^p(S_1;X_{\mathbb{C}}^{\gamma,p})}.
\end{aligned}$$

□

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

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

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

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

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

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

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

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

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

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

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

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

2.1.c Counterexamples

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

Example 2.1.15 (Fourier transform). The *Fourier transform*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\mathscr{F}(Hf)(\xi) = -i\,\mathrm{sgn}(\xi)\mathscr{F}f(\xi),$$

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

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

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

On the other hand, from the straightforward identity

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

we obtain

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

so that

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

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

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

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

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

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

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

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

defines the Riesz projection

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

2.2 Interpolation of Bochner spaces

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

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

2.2.a The Riesz–Thorin interpolation theorem

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

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

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

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

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

where

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

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

Consider the open strip

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

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

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

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

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

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

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

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

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

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

$$\frac{1}{p_z} = \frac{1-z}{p_0} + \frac{z}{p_1}, \quad \frac{1}{q_z'} = \frac{1-z}{q_0'} + \frac{z}{q_1'},$$

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

$z \in \overline{\mathbb{S}}$, the μ_0- and μ_1-simple functions $f_z \in L^{p_0}(S_0; X) \cap L^{p_1}(S_0; X)$ and $g_z \in L^{q_0'}(S_1; Y^*) \cap L^{q_1'}(S_1; Y^*)$ by

$$f_z(s_0) = \mathbf{1}_{\{a(s_0)\neq 0\}} \|a(s_0)\|^{p_\theta/p_z} \frac{a(s_0)}{\|a(s_0)\|},$$
$$g_z(s_1) = \mathbf{1}_{\{b(s_1)\neq 0\}} \|b(s_1)\|^{q_\theta'/q_z'} \frac{b(s_1)}{\|b(s_1)\|}.$$

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

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

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

$$|F(iv)| \leqslant A_0 \|f_{iv}\|_{p_0} \|g_{iv}\|_{q_0'} \leqslant A_0 \|a\|_{p_\theta}^{p_\theta/p_0} \|b\|_{q_\theta'}^{q_\theta'/q_0'}$$

and similarly

$$|F(1+iv)| \leqslant A_1 \|a\|_{p_\theta}^{p_\theta/p_1} \|b\|_{q_\theta'}^{q_\theta'/q_1'}.$$

Hence by the three lines lemma,

$$\begin{aligned}
\Big| \int_{S_1} \langle Ta, b \rangle \, \mathrm{d}\mu_1 \Big| &= |F(\theta)| \\
&\leqslant A_0^{1-\theta} A_1^\theta \|a\|_{p_\theta}^{(1-\theta)p_\theta/p_0} \|b\|_{q_\theta'}^{(1-\theta)q_\theta'/q_0'} \|a\|_{p_\theta}^{\theta p_\theta/p_1} \|b\|_{q_\theta'}^{\theta q_\theta'/q_1'} \\
&\leqslant A_0^{1-\theta} A_1^\theta \|a\|_{p_\theta} \|b\|_{q_\theta'}.
\end{aligned}$$

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

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

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

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

2.2.b The Marcinkiewicz interpolation theorem

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

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

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

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

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

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

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

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

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

$$\|T(f+g)\| \leqslant \|Tf)\| + \|T(g)\| \ \text{ almost everywhere.}$$

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

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

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

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

Then we have the estimates

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

where

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

and

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

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

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

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

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

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

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

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

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

and

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

Hence

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

as claimed. ∎

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

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

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

where

$$\widetilde{f}^{\lambda} := \Big(f - \lambda\frac{f}{\|f\|}\Big)\cdot \mathbf{1}_{\{\|f\|>\lambda\}},$$
$$\widetilde{f}_{\lambda} := f\cdot \mathbf{1}_{\{\|f\|\leqslant\lambda\}} + \lambda\frac{f}{\|f\|}\cdot \mathbf{1}_{\{\|f\|>\lambda\}}.$$

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

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

as well as

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

moreover, we have $\|\widetilde{f}_{\lambda}\|_\infty \leqslant \lambda$.

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

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

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

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

would work just as well.

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

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

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

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

Concerning the weak inequalities (2.6), we have

$$\begin{aligned}\|\widetilde{f}^\lambda\|_{p_0}^{p_0} &= \int_S(\|f\|-\lambda)_+^{p_0}\,\mathrm{d}\mu = \int_0^\infty p_0t^{p_0-1}\mu(\|f\|>\lambda+t)\,\mathrm{d}t\\ &\leqslant \int_0^\infty\frac{p_0t^{p_0-1}}{(\lambda+t)^p}\|f\|_{L^{p,\infty}}^p\,\mathrm{d}t = \int_0^\infty\frac{p_0u^{p_0-1}}{(1+u)^p}\,\mathrm{d}u\cdot\lambda^{p_0-p}\|f\|_{L^{p,\infty}}^p\\ &= p_0\frac{\Gamma(p_0)\Gamma(p-p_0)}{\Gamma(p)}\lambda^{p_0-p}\|f\|_{L^{p,\infty}}^p\\ &= \frac{\Gamma(p_0+1)\Gamma(p-p_0)}{\Gamma(p)}\lambda^{p_0-p}\|f\|_{L^{p,\infty}}^p,\end{aligned}$$

and

$$\begin{aligned}\|\widetilde{f}_\lambda\|_{p_1}^{p_1} &= \int_S(\|f\|\wedge\lambda)^{p_1}\,\mathrm{d}\mu = \int_0^\infty p_1t^{p_1-1}\mu(\|f\|\wedge\lambda>t)\,\mathrm{d}\mu\\ &= \int_0^\lambda p_1t^{p_1-1}\mu(\|f\|>t)\,\mathrm{d}t \leqslant \int_0^\lambda p_1t^{p_1-p-1}\|f\|_{L^{p,\infty}}^p\,\mathrm{d}t\\ &= \frac{p_1}{p_1-p}\lambda^{p_1-p}\|f\|_{L^{p,\infty}}^p.\end{aligned}$$

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

Now we are ready for:

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

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

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

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

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

Case $p_1 < \infty$

For the strong inequality, we have

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

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

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

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

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

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

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

Case $p_1 = \infty$

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

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

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

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

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

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

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

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

The proof will be based on the following approximation result.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

For k large enough,

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

Also,

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

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

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

from which the density follows. □

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

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

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

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

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

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

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

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

writing

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

for brevity. This function is holomorphic on $\mathbb{S}$ and satisfies $F(\theta) = f(\theta)$. Moreover, $v \mapsto F(iv)$ and $v \mapsto F(1+iv)$ are continuous as functions with values $L^{p_0}(S; X_0)$ and $L^{p_1}(S; X_1)$, respectively. Noting that $\theta p p_0(\frac{1}{p_0} - \frac{1}{p_1}) = -pp_0\big(\frac{1}{p} - \frac{1}{p_0}\big) = p - p_0$ we obtain

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

Similarly,

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

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

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

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

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

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

Taking the infimum over all admissible F it follows that

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

This concludes the proof of the theorem. □

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

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

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

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

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

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

and

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

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

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

and

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

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

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

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

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

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

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

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

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

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

This is the second inequality in (1). □

Complex interpolation of real Banach spaces

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\|f(s)\|_{(X_0,X_1)_{\theta,p_0,p_1}} \leqslant \|t \mapsto t^{-\theta}u(t,s)\|^{1-\theta}_{L^{p_0}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)} \cdot \|t \mapsto t^{1-\theta}u(t,s)\|^{\theta}_{L^{p_1}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)}.$$

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

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

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

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

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

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

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

where the exponent ν will be chosen shortly. Then

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

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

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

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

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

□

2.3 The Hardy–Littlewood maximal operator

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

$$Mf(x) := \sup_{B \ni x} \fint_B \|f(y)\| \, \mathrm{d}y := \sup_{B \ni x} \frac{1}{|B|} \int_B \|f(y)\| \, \mathrm{d}y, \tag{2.8}$$

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

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

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

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

Since

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

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

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

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

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

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

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

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

$$w(B) := \int_B w(x)\,\mathrm{d}x.$$

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

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

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

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

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

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

$$\begin{aligned} w(K) \leqslant w\Big(\bigcup_{B\in\mathscr{B}} B\Big) &\leqslant w\Big(\bigcup_{B\in\mathscr{B}_0} 3B\Big) \\ &\leqslant \sum_{B\in\mathscr{B}_0} \frac{w(3B)}{|3B|} 3^d |B| \\ &\overset{(*)}{\leqslant} \sum_{B\in\mathscr{B}_0} \inf_{z\in B} Mw(z) \frac{3^d}{t} \int_B \|f(y)\| \,\mathrm{d}y \\ &\leqslant \frac{3^d}{t} \sum_{B\in\mathscr{B}_0} \int_B \|f(y)\| Mw(y) \,\mathrm{d}y \leqslant \frac{3^d}{t} \|f\|_{L^1(\mathbb{R}^d,Mw;X)}, \end{aligned} \tag{2.9}$$

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

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

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

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

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

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

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

2.3.a Lebesgue points and differentiation

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

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

$$\lim_{\substack{B\ni x\\|B|\to 0}} \fint_B \|f(y)-f(x)\|\,\mathrm{d}y = 0,$$

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

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

$$\lim_{\substack{B\ni x\\|B|\to 0}} \fint_B f(y)\,\mathrm{d}y = f(x).$$

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

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

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

$$\varLambda f(x) := \limsup_{\substack{B\ni x\\|B|\to 0}} \fint_B \|f(y)-f(x)\|\,\mathrm{d}y.$$

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

$$\varLambda f = \varLambda(f-g+g)\leqslant \varLambda(f-g)+\varLambda g\leqslant M(f-g)+\|f-g\|+0,$$

and hence

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

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

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

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

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

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

$$\frac{1}{h}\int_x^{x+h} \|f(y)-f(x)\|\,\mathrm{d}y \leqslant \frac{2}{|I|}\int_I \|f(y)-f(x)\|\,\mathrm{d}y.$$

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

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

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

$$\Big\| \limsup_{n\to\infty} \|T_n x\|_Y \Big\|_{L^{p,\infty}(S)} \leqslant C\|x\|.$$

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

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

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

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

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

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

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

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

From this and the assumption we find

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

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

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

2.3.b Convolutions and approximation

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

$$\phi_\varepsilon * f(x) := \int_{\mathbb{R}^d} \phi_\varepsilon(y) f(x-y)\,\mathrm{d}y, \qquad \phi_\varepsilon(y) := \frac{1}{\varepsilon^d}\phi\Big(\frac{y}{\varepsilon}\Big).$$

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

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

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

takes place under either of the following additional assumptions:

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

$$z \mapsto \varPhi(z) := \operatorname*{ess\,sup}_{|y|\geqslant|z|} |\phi(y)| \quad \textit{belongs to } L^1(\mathbb{R}^d), \tag{2.11}$$

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

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

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

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

will do.

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

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

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

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

$$F_\varepsilon \leqslant C \fint_{B_0} \|f(x-\varepsilon y) - f(x)\|\,\mathrm{d}y = C\fint_{x+\varepsilon B_0} \|f(z) - f(x)\|\,\mathrm{d}z \to 0,$$

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

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

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

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

$$\begin{aligned} F_\varepsilon &\leqslant \int_{\mathbb{R}^d} \Phi(y) \|f(x-\varepsilon y) - f(x)\| \,\mathrm{d}y \\ &= \int_{\mathbb{R}_+} \Big(\int_{|y|\leqslant t} \|f(x-\varepsilon y) - f(x)\| \,\mathrm{d}y \Big) \,\mathrm{d}\mu(t) \\ &= \int_{\mathbb{R}_+} |B(0,t)| \Big(\fint_{B(x,\varepsilon t)} \|f(z) - f(x)\| \,\mathrm{d}y \Big) \,\mathrm{d}\mu(t). \end{aligned} \tag{2.12}$$

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

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

by assumption.

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

$$\fint_B \|f(y) - f(x)\| \,\mathrm{d}y \leqslant 1$$

whenever $B \ni x$ satisfies $|B| < \varepsilon$, and hence $\fint_B \|f\| \,\mathrm{d}x \leqslant \|f(x)\| + 1$ for all such balls. On the other hand, if $|B| \geqslant \varepsilon$, then

$$\fint_B \|f(y)\| \,\mathrm{d}y \leqslant |B|^{-1/p} \|f\|_p \leqslant \varepsilon^{-1/p} \|f\|_p.$$

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

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

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

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

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

2.4 The Fourier transform

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

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

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

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

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

$$\mathscr{F}^{-1} f(x) := \widecheck{f}(x) := \int_{\mathbb{R}^d} f(\xi) e^{2\pi i x \cdot \xi} \, \mathrm{d}\xi, \qquad x \in \mathbb{R}^d.$$

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

$$\Big(\sum_{n-1}^{N} g_n \otimes x_n \Big)^{\widehat{}} = \sum_{n=1}^{N} \widehat{g}_n \otimes x_n,$$

and similarly for the inverse Fourier transform.

Writing

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

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

$$\widecheck{f} = (\widehat{f}\,)^{\widetilde{}} = (\widetilde{f}\,)^{\widehat{}}, \quad \widehat{f} = (\widecheck{f}\,)^{\widetilde{}} = (\widetilde{f}\,)^{\widecheck{}}\,, \tag{2.14}$$

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

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

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

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

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

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

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

after an integration by parts, and thus

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

□

2.4.a The inversion formula and Plancherel's theorem

To proceed any further we need the following classical identity.

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

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

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

$$= \exp(-\pi\xi^2) \int_{\mathbb{R}} \exp(-\pi z^2)\,\mathrm{d}z$$
$$= \exp(-\pi\xi^2).$$

The general case follows from this via separation of variables. □

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

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

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

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

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

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

$$\begin{aligned}
f(x) &= \lim_{\varepsilon\to 0} \int_{\mathbb{R}^d} g_\varepsilon(y) f(x-y)\,\mathrm{d}y \\
&= \lim_{\varepsilon\to 0} \int_{\mathbb{R}^d} \Big[\int_{\mathbb{R}^d} g(\varepsilon\xi) e^{2\pi i y\cdot\xi}\,\mathrm{d}\xi\Big] f(x-y)\,\mathrm{d}y \\
&= \lim_{\varepsilon\to 0} \int_{\mathbb{R}^d} g(\varepsilon\xi) \Big[\int_{\mathbb{R}^d} e^{-2\pi i (x-y)\cdot\xi} f(x-y)\,\mathrm{d}y\Big] e^{2\pi i x\cdot\xi}\,\mathrm{d}\xi \\
&= \lim_{\varepsilon\to 0} \int_{\mathbb{R}^d} g(\varepsilon\xi) \widehat{f}(\xi) e^{2\pi i x\cdot\xi}\,\mathrm{d}\xi \\
&= \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{2\pi i x\cdot\xi}\,\mathrm{d}\xi = (\widehat{f})^{\smallsmile}(x),
\end{aligned}$$

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

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

The Plancherel theorem

It will be useful to introduce the class

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

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

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

$$\mathscr{F}f := \widehat{f} := g.$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

We are now ready to prove:

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

$$\|\widehat{f}\|_2 = \|f\|_2. \tag{2.15}$$

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

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

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

$$\int_{\mathbb{R}^d} (\widehat{f}(x)|g(x)) \,\mathrm{d}x = \int_{\mathbb{R}^d} (f(x)|\breve{g}(x)) \,\mathrm{d}x.$$

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

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

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

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

2.4.b Fourier type

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

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

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

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

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

$$\Big(\frac{p^{1/p}}{(p')^{1/p'}}\Big)^{(d-1)/2}\varphi_{p,X}(\mathbb{R})\leqslant\varphi_{p,X}(\mathbb{R}^d)\leqslant(\varphi_{p,X}(\mathbb{R}))^d. \tag{2.16}$$

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

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

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

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

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

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

$$
\begin{aligned}
\|\mathscr{F} f\|_{L^{p'}(\mathbb{R}^{d+1};X)} &= \Big(\int_{\mathbb{R}^d}\int_{\mathbb{R}} \|\mathscr{F}_1(\mathscr{F}_d(f)(\cdot,\eta))(\xi)\|^{p'}\,\mathrm{d}\xi\,\mathrm{d}\eta\Big)^{1/p'} \\
&\leqslant \varphi_{p,X}(\mathbb{R})\Big(\int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}} \|\mathscr{F}_d(f)(x,\eta)\|^{p}\,\mathrm{d}x\Big)^{p'/p}\mathrm{d}\eta\Big)^{1/p'} \\
&\leqslant \varphi_{p,X}(\mathbb{R})\Big(\int_{\mathbb{R}}\Big(\int_{\mathbb{R}^d} \|\mathscr{F}_d(f)(x,\eta)\|^{p'}\,\mathrm{d}\eta\Big)^{p/p'}\mathrm{d}x\Big)^{1/p} \\
&\leqslant \varphi_{p,X}(\mathbb{R})\varphi_{p,X}(\mathbb{R}^d)\Big(\int_{\mathbb{R}}\int_{\mathbb{R}^d} \|f(x,y)\|^{p}\,\mathrm{d}y\,\mathrm{d}x\Big)^{1/p} \\
&= \varphi_{p,X}(\mathbb{R})\varphi_{p,X}(\mathbb{R}^d)\|f\|_{L^p(\mathbb{R}^{d+1};X)}.
\end{aligned}
$$

This also proves the right-hand side estimate in (2.16). □

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

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

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

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

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

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

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

and the result follows.

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

Specialising to the scalar-valued setting we obtain:

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

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

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

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

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

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

Hence,

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

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

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

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

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

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

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

$$\varphi_{p,X_\theta}(\mathbb{R}^d) \leqslant (\varphi_{p_0,X_0}(\mathbb{R}^d))^{1-\theta}(\varphi_{p_1,X_1}(\mathbb{R}^d))^\theta,$$
$$\varphi_{p,(X_0,X_1)_{\theta,p_0,p_1}} \leqslant (\varphi_{p_0,X_0}(\mathbb{R}^d))^{1-\theta}(\varphi_{p_1,X_1}(\mathbb{R}^d))^\theta.$$

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

Proof. For X_θ the result follows from Theorem 2.2.6:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

and, by 1-periodicity,

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

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

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

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

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

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

Then

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

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

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

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

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

where, for every $r \geqslant 2$,

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

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

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

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

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

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

and

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Let X be a Banach space.

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

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

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

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

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

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

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

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

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

Thus

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

and

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

where finally

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

so that altogether

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

from which the assertion follows. $\square$

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

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

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

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

We record a basic density result:

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

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

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

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

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

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

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

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

Definition 2.4.24 (Tempered distributions). *The space*

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

is called the space of X-valued tempered distributions.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Clearly, the right-hand side grows polynomially in the variable x. ☐

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

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

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

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

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

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

2.5 Sobolev spaces and differentiability

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

2.5.a Weak derivatives

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

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

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

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

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

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

$$\int_D f(x)\partial^\alpha\phi(x)\,\mathrm{d}x=(-1)^{|\alpha|}\int_D g(x)\phi(x)\,\mathrm{d}x\quad \text{for all } \phi\in C_{\rm c}^\infty(D).$$

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

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

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

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

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

$$\int_D f(x)\phi(x)\,\mathrm{d}x=0\quad\forall\phi\in C_{\rm c}^\infty(D),$$

then $f=0$ almost everywhere.

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

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

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

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

For later use we also record the following fact.

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

Proof. *Step 1* – For all $\phi, \psi \in C_{\mathrm{c}}^{\infty}(D)$ we have

$$\int_D (\phi(x)\nabla\psi(x) + \nabla\phi(x)\psi(x)) f(x)\,\mathrm{d}x = -\int_D \phi(x)\psi(x)\nabla f(x)\,\mathrm{d}x,$$

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

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

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

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

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

Let X be a Banach space.

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

Endowed with the norm

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

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

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

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

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

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

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

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

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

2.5.c Almost everywhere differentiability

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

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

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

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

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

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

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

where $\partial_j f$ denotes the jth partial derivative of f.

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

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

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

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

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

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

$$f(x+h) - f(x) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s} f(x+sh)\,\mathrm{d}t = \int_0^1 \nabla f(x+sh) \cdot h\,\mathrm{d}s. \tag{2.19}$$

Taking $h = te_j$ we find

$$\|\delta_j^t f\|_p \leqslant \int_0^1 \|\partial_j f(\cdot + se_j)\|_p\,\mathrm{d}s = \|\partial_j f\|_p.$$

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

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

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

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

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

$$\Big\| \int_{\mathbb{R}^d} f(x)\partial_j \phi(x)\,\mathrm{d}x \Big\| \leqslant C\|\phi\|_{p'}. \tag{2.20}$$

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

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

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

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

By linearity, for simple functions ϕ this gives

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

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

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

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

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

$$\Big| \int_{\mathbb{R}^d} \langle f(x), \partial_j\phi(x)\rangle\,\mathrm{d}x \Big| \leqslant C\|\phi\|_{p'}.$$

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

$$\Big| \int_{\mathbb{R}^d} \langle \partial_j f(x), \phi(x)\rangle\,\mathrm{d}x \Big| \leqslant C\|\phi\|_{p'}.$$

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

As an application of the Lebesgue differentiation theorem, we shall prove next that every Lipschitz continuous function $f : \mathbb{R} \to X$ is almost everywhere differentiable if and only if X has the Radon–Nikodým property. In Proposition 2.5.7 we have already seen that the RNP can be used to obtain the existence of weak derivatives for functions $f : \mathbb{R}^d \to X$ satisfying a certain Lipschitz estimate. We will now focus on the one-dimensional case and obtain equivalences between weak differentiability and almost everywhere differentiability. Some of the results do extend to higher dimensions, but this requires a more thorough development of Sobolev spaces.

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

It will be convenient to put

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

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

$$f(t) := \int_a^t g(s)\,\mathrm{d}s.$$

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

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

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

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

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

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

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

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

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

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

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

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

A function $f:\mathbb{R}\to X$ is said to belong to $W^{1,p}_{\mathrm{loc}}(\mathbb{R};X)$ if f belongs to $L^p_{\mathrm{loc}}(\mathbb{R},X)$ and has a weak derivative belonging to $L^p_{\mathrm{loc}}(\mathbb{R};X)$.

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

(1) $f\in W^{1,p}_{\mathrm{loc}}(\mathbb{R};X)$*;*
(2) *there exists a function $g\in L^p_{\mathrm{loc}}(\mathbb{R};X)$ such that*

$$f(t)-f(s)=\int_s^t g(r)\,\mathrm{d}r \quad \textit{for all } -\infty<s\leqslant t<\infty;$$

(3) *f is almost everywhere differentiable and $f'\in L^p_{\mathrm{loc}}(\mathbb{R};X)$.*

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

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

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

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

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

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

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

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

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

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

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

so eventually $|E_n| \leqslant \delta$. $\square$

We use this lemma to prove:

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

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

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

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

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

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

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

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

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

(2)⇒(1) and (2)⇒(3) both follow from Lemma 2.5.8.

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

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

(1) *X has the Radon–Nikodým property;*
(2) *every locally absolutely continuous function $f : \mathbb{R} \to X$ is differentiable almost everywhere;*
(3) *every locally Lipschitz continuous function $f : \mathbb{R} \to X$ is differentiable almost everywhere;*
(4) *every locally absolutely continuous function $f : \mathbb{R} \to X$ belongs to $W^{1,1}_{\mathrm{loc}}(\mathbb{R};X)$;*
(5) *every locally Lipschitz continuous function $f : \mathbb{R} \to X$ belongs to $W^{1,\infty}_{\mathrm{loc}}(\mathbb{R};X)$.*

Proof. (1)⇒(3): It is possible to deduce this from Proposition 2.5.7, but we prefer to give a different argument here. It suffices to prove the almost everywhere differentiability on a finite interval $I = (a,b)$. For $a \leqslant s < t \leqslant b$ define

$$T\mathbf{1}_{(s,t)} := f(t) - f(s)$$

and extend this definition to step functions by linearity. For such functions $g = \sum_{n=1}^{N} c_n \mathbf{1}_{(t_{n-1},t_n)}$,

$$\|Tg\| \leqslant L \sum_{n=1}^{N} |c_n|(t_n - t_{n-1}) = L\|g\|_1,$$

where L is the Lipschitz constant of f. Therefore T extends to a bounded operator from $L^1(I)$ to X of norm $\|T\| \leqslant L$. By Theorem 1.3.10, T is representable by a function $\phi \in L^\infty(I;X)$. For all $t \in I$ this gives

$$f(t) - f(a) = T\mathbf{1}_{(a,t)} = \int_a^t \phi(r)\,\mathrm{d}r,$$

and Lemma 2.5.8 gives the result.

(3)⇔(5): This follows from Proposition 2.5.9, noting that the almost everywhere derivative of a Lipschitz functions is bounded.

(5)⇒(1): Let $T : L^1(0,1) \to X$ be a bounded operator. Define $f : [0,1] \to X$ by

$$f(t) := T\mathbf{1}_{(0,t)}.$$

Then for all $0 \leqslant s < t \leqslant 1$ we have

$$\|f(t) - f(s)\| \leqslant \|T\|\|\mathbf{1}_{(s,t)}\|_1 = (t-s)\|T\|,$$

so f is Lipschitz continuous. Now extend f to a Lipschitz function on $\mathbb{R}$ by setting $f(t) = f(0)$ for $t < 0$ and $f(t) = f(1)$ for $t > 1$. Let f' be the almost everywhere defined derivative of f. Clearly, $f' \in L^\infty(\mathbb{R};X)$ with $\|f'\|_\infty \leqslant L$, the Lipschitz constant of f. Using Proposition 2.5.9, for all $0 \leqslant a < b \leqslant 1$ we obtain

$$T\mathbf{1}_{(a,b)} = f(b) - f(a) = \int_a^b f'(t)\,\mathrm{d}t = \int_0^1 \mathbf{1}_{(a,b)}(t) f'(t)\,\mathrm{d}t.$$

It follows that $T\phi = \int_0^1 \phi(t) f'(t)\,\mathrm{d}t$ holds for all step function ϕ. Since these are dense in $L^1(0,1)$ and T is bounded, this identity extends to arbitrary $\phi \in L^1(0,1)$. This proves that T is representable, and therefore X has the RNP with respect to $[0,1]$ by Theorem 1.3.10.

(5)⇒(4): Let $f : \mathbb{R} \to X$ be absolutely continuous. It suffices to prove that f is almost everywhere differentiable with integrable derivative on finite intervals. By a scaling and translation argument it suffices to consider $I = (0,1)$. Define the function $V_f : [0,1] \to \mathbb{R}$ by

$$V_f(t) := \sup \sum_{n=1}^{N} \|f(t_n) - f(t_{n-1})\|,$$

where the supremum is taken over all finite partitions $0 = t_0 < \cdots < t_N = 1$. Note that $V_f(t)$ is finite, because f is absolutely continuous, and as a function of t, V_f is non-decreasing. The function

$$h(t) := \frac{V_f(t) + t}{V_f(1) + 1}$$

is absolutely continuous, strictly increasing, and satisfies $h(0) = 0$, $h(1) = 1$. Moreover, for all $0 \leqslant s \leqslant t \leqslant 1$,

$$\|f(t) - f(s)\| \leqslant V_f(t) - V_f(s) \leqslant (V_f(1) + 1)(h(t) - h(s)).$$

Hence $f \circ h^{-1}$ is Lipschitz continuous on $[0,1]$, with Lipschitz constant $\leqslant V_f(1) + 1$. By assumption, this implies that $f \circ h^{-1}$ is differentiable almost everywhere, say outside some null set N_0, and clearly its derivative is bounded by $V_f(1) + 1$ almost everywhere, say outside some null set N_1. Also, h is differentiable almost everywhere as well, say outside some null set N_2, and the derivative h' is integrable; the proof of this fact will be given in Lemma 2.5.13 below. From $|h(t) - h(s)| \geqslant |t - s|/(V_f(1) + 1)$ it follows that h^{-1} maps null sets to null sets. Therefore, by the chain rule, $f = (f \circ h^{-1}) \circ h$ is differentiable on the complement of the null set $N := h^{-1}(N_0 \cup N_1) \cup N_2$ and there it satisfies $f'(t) = (f \circ h^{-1})'(h(t)) \cdot h'(t)$ and

$$\|f'(t)\| \leqslant (V_f(1) + 1)|h'(t)|.$$

This shows that f' is Bochner integrable.

(4)⇒(2): This is a consequence of Proposition 2.5.9.

(2)⇒(3): Every Lipschitz function is absolutely continuous. □

It remains to prove the following fact, which was used in the proof of the implication (5)⇒(2):

Lemma 2.5.13. *If $f : \mathbb{R} \to [0, \infty)$ is non-decreasing and absolutely continuous, then f is differentiable almost everywhere and f' is locally integrable.*

Proof. For $a < b$ set $\mu([a, b)) := f(b) - f(a)$ and extend this definition to finite unions of such half-open intervals. By the Carathéodory extension theorem, μ extends to a Borel measure μ on $\mathbb{R}$ which is finite on bounded intervals. We will show that μ is absolutely continuous with respect to the Lebesgue measure. To this end fix $\varepsilon > 0$ and let δ be as in the definition of absolute continuity.

If B is a Borel set of Lebesgue measure $|B| < \frac{1}{2}\delta$, invoking Lemma A.1.2 we find sets $I_n = [a_n, b_n)$, $n = 1, \dots, N$, such that both $|B \Delta \bigcup_{n=1}^N I_n| < \frac{1}{2}\delta$ and $\mu(B \Delta \bigcup_{n=1}^N I_n) < \varepsilon$. The former implies that

$$\sum_{n=1}^{n} (b_n - a_n) = \bigcup_{n=1}^{N} |I_n| < \delta.$$

Hence by the absolute continuity of f,

$$\mu(\bigcup_{n=1}^{N} I_n) = \sum_{n=1}^{N}(f(b_n) - f(a_n)) < \varepsilon.$$

Therefore $\mu(B) < 2\varepsilon$. This proves the absolute continuity of μ. By the Radon–Nikodým theorem, applied to the intervals $(-r, r)$, μ has a density $g \in L^1_{\rm loc}(\mathbb{R})$. In particular,

$$f(b) - f(a) = \int_a^b g(s)\,\mathrm{d}s$$

for all $a < b$. Now the result, with $f' = g$, follows from Lemma 2.5.8. □

Remark 2.5.14. The implication (1)⇒(2) of the theorem admits the following direct and more transparent proof, although its rigorous implementation is somewhat laborious. The idea is that the X-valued Stieltjes integral

$$G(A) := \int_0^1 \mathbf{1}_A\,\mathrm{d}f$$

defines an X-valued measure of bounded variation which is absolutely continuous with respect to the Lebesgue measure. If $\phi \in L^1(0, 1; X)$ denotes its density, we have

$$f(t) = \int_0^t \phi(t)\,\mathrm{d}t$$

and the result follows from Lemma 2.5.8.

Corollary 2.5.15. *Let X be reflexive.*

(1) *Every locally absolutely continuous function $f : \mathbb{R} \to X$ is differentiable almost everywhere, and $f \in W^{1,1}_{\rm loc}(\mathbb{R}; X)$.*
(2) *Every locally Lipschitz continuous function $f : \mathbb{R} \to X$ is differentiable almost everywhere, and $f \in W^{1,\infty}_{\rm loc}(\mathbb{R}; X)$.*

In both situations, the almost everywhere derivative and the weak derivative agree.

2.5.d The fractional Sobolev spaces $W^{s,p}(\mathbb{R}^d; X)$

In this subsection we introduce the scale of Sobolev spaces $W^{s,p}(\mathbb{R}^d; X)$ and identify them, for $0 < s < 1$, as real interpolation spaces between $L^p(\mathbb{R}^d; X)$ and $W^{1,p}(\mathbb{R}^d; X)$. The results of this subsection will not be needed elsewhere in this volume. We will revisit Sobolev spaces in Chapter 5.

Definition 2.5.16 (Sobolev–Slobodetskii spaces). *Let $D \subseteq \mathbb{R}^d$ be an open set. For $p \in [1, \infty)$ and $0 < s < 1$ we define $W^{s,p}(D; X)$ as the space of all $f \in L^p(D; X)$ for which*

$$[f]_{W^{s,p}(D;X)} := \Big(\int_D \int_D \frac{\|f(x) - f(y)\|^p}{|x - y|^{sp+d}}\,\mathrm{d}x\,\mathrm{d}y\Big)^{1/p} < \infty.$$

Endowed with the norm

$$\|f\|_{W^{s,p}(D;X)} := \|f\|_p + [f]_{W^{s,p}(D;X)},$$

$W^{s,p}(\mathbb{R}^d; X)$ is a Banach space.

In this subsection we will take $D = \mathbb{R}^d$ and note that

$$W^{1,p}(\mathbb{R}^d; X) \hookrightarrow W^{s,p}(\mathbb{R}^d; X) \hookrightarrow L^p(\mathbb{R}^d; X)$$

with continuous embeddings. For the first embedding note that

$$[f]_{W^{s,p}(\mathbb{R}^d;X)} = \Big(\int_{\mathbb{R}^d} \frac{\|f(x+h) - f(x)\|^p}{|h|^{sp+d}} \,\mathrm{d}x \,\mathrm{d}h\Big)^{1/p}.$$

Now for $|h| > 1$ one may use $\|f(\cdot + h) - f(\cdot)\|_p \leqslant 2\|f\|_p$, and for $|h| \leqslant 1$ one may use that $\|f(\cdot+h) - f(\cdot)\|_p \leqslant \|f\|_{W^{1,p}(\mathbb{R}^d;X)}|h|$, which follows from (2.19).

We are now ready to state and prove the main result of this section:

Theorem 2.5.17 (Real interpolation). *Let X be a Banach space. For all $p \in [1, \infty)$ and $0 < s < 1$ we have*

$$(L^p(\mathbb{R}^d; X), W^{1,p}(\mathbb{R}^d; X))_{s,p} = W^{s,p}(\mathbb{R}^d; X)$$

with equivalent norms.

Proof. Below we write $K(t, f) := K(t, f, L^p, W^{1,p})$ for the K-functional as introduced in Section C.3.

'$\subseteq$': Fix $f \in (L^p(\mathbb{R}^d; X), W^{1,p}(\mathbb{R}^d; X))_{s,p}$. For fixed $h \in \mathbb{R}^d$ choose $a = a_h \in L^p(\mathbb{R}^d; X)$ and $b = b_h \in W^{1,p}(\mathbb{R}^d; X)$ such that $f = a + b$ and

$$\|a\|_p + |h|\|b\|_{W^{1,p}} \leqslant 2K(|h|, f).$$

Using the triangle inequality and Proposition 2.5.7 we find

$$\begin{aligned}\|f(\cdot + h) - f\|_p &\leqslant \|a(\cdot + h) - a\|_p + \|b(\cdot + h) - b\|_p \\ &\leqslant 2\|a\|_p + |h|\|\nabla b\|_p \leqslant 2\|a\|_p + |h|\|b\|_{W^{1,p}} \leqslant 4K(|h|, f).\end{aligned}$$

Integrating over h and turning to polar coordinates,

$$[f]^p_{W^{s,p}(\mathbb{R}^d;X)} \leqslant 4^p \int_{\mathbb{R}^d} \frac{(K(|h|, f))^p}{|h|^{sp+d}} \,\mathrm{d}h \tag{2.21}$$

$$= C_{d,p} \int_0^\infty t^{-sp}(K(t, f))^p \,\frac{\mathrm{d}t}{t} = C_{d,p}\|f\|^p_{(L^p(\mathbb{R}^d;X),W^{1,p}(\mathbb{R}^d;X))_{s,p}} \tag{2.22}$$

with $C_{d,p} = 4^p \times$ the area of the unit sphere of $\mathbb{R}^d$. Secondly, since $W^{1,p}(\mathbb{R}^d; X)$ embeds into $L^p(\mathbb{R}^d; X)$ continuously, we also have $\min\{1, t\}\|f\|_p \lesssim_{d,p} K(t, f)$

(see (C.7)) for all $t > 0$. Multiplying both sides with t^{-s} and integrating, we obtain

$$\|f\|_p \leqslant C_{p,s}\|f\|_{(L^p(\mathbb{R}^d;X),W^{1,p}(\mathbb{R}^d;X))_{s,p}}. \tag{2.23}$$

Combining the estimates (2.21) and (2.23), the desired embedding is obtained.

'$\supseteq$': For $t \geqslant 1$ the trivial estimate $K(t,f) \leqslant \|f\|_p$ already gives, for $f \in W^{s,p}(\mathbb{R}^d;X)$,

$$\int_1^\infty t^{-sp}(K(t,f))^p \frac{\mathrm{d}t}{t} \leqslant \int_1^\infty t^{-sp}\|f\|_p^p \frac{\mathrm{d}t}{t} = C_{s,p}\|f\|_p^p \leqslant C_{s,p}\|f\|_{s,p}^p.$$

To estimate the integral over $0 < t < 1$, let $\phi \in C_\mathrm{c}^1(\mathbb{R}^d)$ be a non-negative test function supported in the unit ball of $\mathbb{R}^d$ satisfying $\int_{\mathbb{R}^d}\phi(x)\,\mathrm{d}x = 1$. Set $\phi_t(x) := t^{-d}\phi(t^{-1}x)$ and define

$$b_t := \int_{\mathbb{R}^d}\phi_t(x-y)f(y)\,\mathrm{d}y, \quad a_t := f - b_t.$$

In order to prove the required embedding it remains to be shown that the right-hand side terms of

$$\begin{aligned}
&\Big(\int_0^1 t^{-sp}(K(t,f))^p \frac{\mathrm{d}t}{t}\Big)^{1/p} \\
&\qquad \leqslant \Big(\int_0^1 t^{-sp}(\|a_t\|_p + t\|b_t\|_{W^{1,p}})^p \frac{\mathrm{d}t}{t}\Big)^{1/p} \\
&\qquad \leqslant \Big(\int_0^1 t^{-sp}\|a_t\|_p^p \frac{\mathrm{d}t}{t}\Big)^{1/p} + \Big(\int_0^1 t^{(1-s)}\Big(\sum_{|\alpha|\leqslant 1}\|\partial^\alpha b_t\|_p\Big)^p \frac{\mathrm{d}t}{t}\Big)^{1/p}
\end{aligned}$$

can be estimated by $C_{s,d,p}\|f\|_{s,p}$.

From $f(x) = \int_{\mathbb{R}^d}\phi_t(x-y)f(x)\,\mathrm{d}y$ we find

$$\begin{aligned}
\|a_t\|_p^p &= \int_{\mathbb{R}^d}\Big\|\int_{\mathbb{R}^d}\phi_t(x-y)[f(x)-f(y)]\,\mathrm{d}y\Big\|^p \mathrm{d}x \\
&\leqslant \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\phi_t(x-y)\|f(y)-f(x)\|^p\,\mathrm{d}y\,\mathrm{d}x,
\end{aligned}$$

applying Jensen's inequality to the probability measures $\phi_t(x-y)\,\mathrm{d}y$. Hence,

$$\begin{aligned}
\int_0^\infty t^{-sp}\|a_t\|_p^p \frac{\mathrm{d}t}{t} &\leqslant \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\Big(\int_0^\infty t^{-sp}\phi_t(x-y)\frac{\mathrm{d}t}{t}\Big)\|f(y)-f(x)\|^p\,\mathrm{d}y\,\mathrm{d}x \\
&\overset{(*)}{=} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\Big(\int_{|x-y|}^\infty t^{-sp-d}\phi\big(\frac{x-y}{t}\big)\frac{\mathrm{d}t}{t}\Big)\|f(y)-f(x)\|^p\,\mathrm{d}y\,\mathrm{d}x \\
&\leqslant \frac{\|\phi\|_\infty}{sp+d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{\|f(y)-f(y)\|^p}{|x-y|^{sp+d}}\,\mathrm{d}y\,\mathrm{d}x
\end{aligned}$$

$$= \frac{\|\phi\|_\infty}{sp+d}[f]^p_{W^{s,p}(\mathbb{R}^d;X)},$$

where $(*)$ uses the definition of ϕ_t and the assumption that ϕ be supported in the unit ball.

Clearly $b_t \in W^{1,p}(\mathbb{R}^d)$. Since $\|b_t\|_p \leqslant \|f\|_p$ by Young's inequality,

$$\int_0^1 t^{(1-s)p}\|b_t\|_p^p \frac{\mathrm{d}t}{t} \leqslant \int_0^1 t^{(1-s)p}\|f\|_p^p \frac{\mathrm{d}t}{t} = C_{p,s}\|f\|_p^p.$$

Next,

$$\begin{aligned}
\partial_j b_t(x) &= \frac{1}{t^{d+1}}\int_{\mathbb{R}^d} \partial_j\phi\Big(\frac{x-y}{t}\Big) f(y)\,\mathrm{d}y \\
&= \frac{1}{t^{d+1}}\int_{\mathbb{R}^d} \partial_j\phi\Big(\frac{x-y}{t}\Big)[f(y)-f(x)]\,\mathrm{d}y \\
&= \frac{1}{t}\int_{\mathbb{R}^d} (\partial_j\phi)_t(x-y)[f(y)-f(x)]\,\mathrm{d}y,
\end{aligned}$$

since $\int_{\mathbb{R}^d} \partial_j\phi(\frac{x-y}{t})\,\mathrm{d}y = 0$ by an integration by parts. Arguing along similar lines as above, applying Jensen's inequality after normalising the measures $|(\partial_j\phi)_t(x-y)|\,\mathrm{d}y$, we obtain

$$\begin{aligned}
&\int_0^\infty t^{(1-s)p}\|\partial_j b_t\|_p^p \frac{\mathrm{d}t}{t} \\
&\quad= \int_0^\infty t^{(1-s)p}\int_{\mathbb{R}^d}\Big\|\frac{1}{t}\int_{\mathbb{R}^d}(\partial_j\phi)_t(x-y)[f(y)-f(x)]\,\mathrm{d}y\Big\|^p\,\mathrm{d}x\,\frac{\mathrm{d}t}{t} \\
&\quad\lesssim_{\phi,j} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\Big(\int_0^\infty t^{-sp}|(\partial_j\phi)_t(x-y)|\,\frac{\mathrm{d}t}{t}\Big)\|f(y)-f(x)\|^p\,\mathrm{d}y\,\mathrm{d}x \\
&\quad= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\Big(\int_{|x-y|}^\infty t^{-sp-d}\Big|\partial_j\phi\Big(\frac{x-y}{t}\Big)\Big|\,\frac{\mathrm{d}t}{t}\Big)\|f(y)-f(x)\|^p\,\mathrm{d}y\,\mathrm{d}x \\
&\quad\leqslant \frac{\|\partial_j\phi\|_\infty}{sp+d}[f]^p_{W^{s,p}(\mathbb{R}^d;X)}.
\end{aligned}$$

□

Remark 2.5.18. We deliberately refrain from discussing the more general problem of characterising the interpolation spaces

$$(L^p(\mathbb{R}^d;X), W^{1,p}(\mathbb{R}^d;X))_{s,q}$$

for $p \neq q$ as this would take us into the realm of the Besov spaces. For similar reasons, we restrict ourselves to the Sobolev spaces over $\mathbb{R}^d$. A comprehensive treatment of Sobolev spaces over open domains $D \subseteq \mathbb{R}^d$ can be developed as well, but again their treatment requires the use of extension techniques that would take us too far afield. All these topics will be covered in Volume III.

Remark 2.5.19. The identification of the complex interpolation spaces

$$[L^p(\mathbb{R}^d;X), W^{1,p}(\mathbb{R}^d;X)]_s$$

for UMD spaces X will be presented in Chapter 5 as an application of the Mihlin multiplier theorem.

2.6 Conditional expectations

We now introduce and study an important class of operators, the conditional expectations, which will then accompany us for the rest of this volume. In particular, they are used to define the fundamental concept of a martingale in the following chapter.

In harmonic analysis, conditional expectations are used through the ubiquitous averaging procedures. In this context we are usually working over $\mathbb{R}^d$ endowed with the Lebesgue measure, and it is for this reason that we have chosen to give a unified presentation of the topic in the setting of σ-finite measure spaces.

In probability theory, it is customary to define conditional expectations only for integrable functions. This automatically implies that the conditional expectation of an L^p-function is well defined. In the σ-finite context this is no longer the case and for this reason we choose to work with strongly measurable functions which enjoy a minimal 'local' integrability condition needed to make the definition of a conditional expectation a meaningful one. Even in the case of probability spaces this leads to a conveniently general set-up.

Throughout this section we fix a measure space $(S,\mathscr{A},\mu)$ and a Banach space X. We recall that $L^0(S;X)$ denotes the vector space of all strongly μ-measurable functions $f: S\to X$, identifying functions that are equal almost everywhere. When we wish to emphasise the underlying σ-algebra we shall write $L^0(S,\mathscr{A};X)$. In particular, when $\mathscr{F}\subseteq\mathscr{A}$ is a sub-σ-algebra, the space $L^0(S,\mathscr{F};X)$ is defined with reference to the measure space $(S,\mathscr{F},\mu|_{\mathscr{F}})$. It has already been observed that this space coincides with the subspace of $L^0(S;X)$ of functions having a strongly $\mathscr{F}$-measurable representative. The elements of $L^0(S,\mathscr{F};X)$ will be referred to as the *$\mathscr{F}$-measurable* elements of $L^0(S;X)$.

Given a sub-σ-algebra $\mathscr{F}\subseteq\mathscr{A}$, for a function $f\in L^0(S;X)$ we denote

$$\mathscr{F}_f := \{F\in\mathscr{F}:\ \mathbf{1}_F f\in L^1(S;X)\}.$$

Definition 2.6.1. *A function $f\in L^0(S;X)$ is called* σ-integrable over $\mathscr{F}$ *if S can be covered by at most countably many sets in $\mathscr{F}_f$. Any covering sequence in $\mathscr{F}_f$ will be called an* exhausting sequence for f in $\mathscr{F}$.

The sets in an exhausting sequence for f can be taken disjoint: if $(F_n)_{n\geqslant 1}$ is exhausting, then so is the sequence $(\widetilde{F}_n)_{n\geqslant 1}$, where $\widetilde{F}_1 = F_1$ and $\widetilde{F}_{n+1} = F_{n+1}\setminus\bigcup_{j=1}^n F_j$, $n\geqslant 1$.

The σ-integrability of f with respect to $\mathscr{F}$ allows us to split the function f in integrable parts which respect the σ-algebra $\mathscr{F}$. This will enable us to build a conditional expectation without assuming global integrability on S.

The next two propositions provide easy examples of σ-integrability.

Proposition 2.6.2. *Every function $f \in L^0(S;X)$ is σ-integrable over $\mathscr{A}$.*

Proof. By Proposition 1.1.15 we have a disjoint decomposition $S = F_0 \cup F_1$ with $F_0, F_1 \in \mathscr{A}$ such that $f \equiv 0$ almost everywhere on F_0 and μ is σ-finite on F_1. Let $(F^{(n)})_{n\geqslant 1}$ be an exhausting sequence for F_1. For each $j \geqslant 1$ let $F_j = \{\|f\| \leqslant j\}$. Then f is integrable on F_0 and on each of the sets $F_{jn} := F^{(n)} \cap F_j$, $j, n \geqslant 1$; together, these sets cover S and therefore provide an exhausting sequence for f in $\mathscr{A}$. □

Although every function $f \in L^0(S;X)$ is σ-integrable over $\mathscr{A}$, such functions need not to be σ-integrable over a sub-σ-algebra of $\mathscr{A}$. For example, this happens when $S = \mathbb{R}$ and $\mathscr{A} = \mathscr{B}(\mathbb{R})$ is the Borel σ-algebra; the constant function $\mathbf{1}$ is σ-integrable over $\mathscr{A}$ but not over the sub-σ-algebra $\mathscr{F} = \{\varnothing, \mathbb{R}\}$. This defect cannot be remedied by assuming that μ be σ-finite on $\mathscr{F}$: if f is a function which fails to be integrable on every interval $[n, n+1)$, then f fails to be σ-integrable over the σ-finite sub-σ-algebra generated by these intervals. If we add a global integrability assumption such problems do not arise:

Proposition 2.6.3. *If μ is σ-finite on the sub-σ-algebra $\mathscr{F}$ of $\mathscr{A}$ and if $1 \leqslant p \leqslant \infty$, then every function $f \in L^p(S;X)$ is σ-integrable over $\mathscr{F}$.*

Proof. Let $(S_n)_{n\geqslant 1}$ is an exhaustion of S by sets in $\mathscr{F}$ of finite μ-measure. The functions $\mathbf{1}_{S_n} f$ belong to $L^1(S;X)$ by Hölder's inequality, and therefore $(S_n)_{n\geqslant 1}$ is an exhausting sequence for f in $\mathscr{F}$. □

Definition 2.6.4 (Conditional expectation). *Let $\mathscr{F} \subseteq \mathscr{A}$ be a sub-σ-algebra. A function $g \in L^0(S, \mathscr{F}; X)$ is said to be a* conditional expectation with respect to $\mathscr{F}$ *of the function $f \in L^0(S;X)$ if*

$$\int_F g \, \mathrm{d}\mu = \int_F f \, \mathrm{d}\mu, \quad \textit{for all } F \in \mathscr{F}_f \cap \mathscr{F}_g. \tag{2.24}$$

Without a σ-integrability assumption on f it may happen that $\mathscr{F}_f = \{\varnothing\}$, in which case the above definition is vacuous. Accordingly, in what follows we will usually assume f to be σ-integrable over $\mathscr{F}$. Among other things this assumption implies that a conditional expectation, if one exists, is unique (Theorem 2.6.18). In this situation we moreover have $\mathscr{F}_f \subseteq \mathscr{F}_g$ and therefore (2.24) holds for all $F \in \mathscr{F}_f$. If we furthermore assume that μ is σ-finite on $\mathscr{F}$, then f admits a unique conditional expectation g with respect to $\mathscr{F}$ (Theorem 2.6.20).

In will be convenient to have a criterion which enables us to verify the defining condition (2.24) by testing it on an appropriate sub-collection $\mathscr{C}$ of $\mathscr{F}_f \cap \mathscr{F}_g$. For this purpose we introduce the following definition.

Definition 2.6.5. *Let $\mathscr{F} \subseteq \mathscr{A}$ be a sub-σ-algebra. A family $\mathscr{C} \subseteq \mathscr{F}$ will be called an* exhausting ideal for $\mathscr{F}$ *if it has the following two properties:*

(i) *$C \cap F \in \mathscr{C}$ for all $C \in \mathscr{C}$ and $F \in \mathscr{F}$;*
(ii) *S can be covered by at most countably many sets from $\mathscr{C}$.*

Example 2.6.6. In a σ-finite measure space $(S, \mathscr{A}, \mu)$, the sets of finite μ-measure form an exhausting ideal for $\mathscr{A}$.

Any exhausting ideal contains a disjoint sequence covering S. Indeed, by (ii) there is a sequence $(C_n)_{n\geqslant 1}$ in $\mathscr{C}$ covering S. Set $C_1' = C_1$ and $C_{n+1}' = C_{n+1} \setminus \bigcup_{j=1}^n C_j$. In view of $C_{n+1}' = C_{n+1} \cap \complement(\bigcup_{j=1}^n C_j)$ and $\complement(\bigcup_{j=1}^n C_j) \in \mathscr{F}$, (i) implies that $C_{n+1}' \in \mathscr{C}$. The sets C_n' are disjoint and cover S.

Exhausting ideals are closed under taking intersections:

Lemma 2.6.7. *If $\mathscr{C}_1$ and $\mathscr{C}_2$ are exhausting ideals for $\mathscr{F}$, then $\mathscr{C}_1 \cap \mathscr{C}_2$ is a exhausting ideal for $\mathscr{F}$ and*

$$\mathscr{C}_1 \cap \mathscr{C}_2 = \{C_1 \cap C_2 : \ C_1 \in \mathscr{C}_1, \ C_2 \in \mathscr{C}_2\}.$$

We leave the simple proof to the reader.

Example 2.6.8. The constant function $\mathbf{1}$ is σ-integrable over $\mathscr{F}$ if and only if μ is σ-finite on $\mathscr{F}$, if and only if $\mathscr{F}_{\mathbf{1}}$ (i.e., $\mathscr{F}_h$ for the constant function $h = \mathbf{1}$) is an exhausting ideal for $\mathscr{F}$. More generally, a function $f \in L^0(S;X)$ is σ-integrable over $\mathscr{F}$ if and only if the measure $d\nu := \|f\| \, \mathrm{d}\mu$ is σ-finite on $\mathscr{F}$, if and only if $\mathscr{F}_f$ is an exhausting ideal for $\mathscr{F}$. In particular, if $f \in L^0(S, \mathscr{F}; X)$, then $\mathscr{F}_f$ is an exhausting ideal for $\mathscr{F}$ by Proposition 2.6.2.

The next result simplifies the task of verifying condition (2.24) considerably.

Proposition 2.6.9. *Let $\mathscr{F} \subseteq \mathscr{A}$ be a sub-σ-algebra, let $f \in L^0(S;X)$ be σ-integrable over $\mathscr{F}$, and let $g \in L^0(S, \mathscr{F}; X)$ be given. The following assertions are equivalent:*

(1) *g is a conditional expectation of f with respect to $\mathscr{F}$;*
(2) *there exists an exhausting ideal $\mathscr{C} \subseteq \mathscr{F}_f \cap \mathscr{F}_g$ for $\mathscr{F}$ such that*

$$\int_C g \, \mathrm{d}\mu = \int_C f \, \mathrm{d}\mu \quad \textit{for all } C \in \mathscr{C}. \tag{2.25}$$

Proof. (1)⇒(2): It follows from Example 2.6.8 that both $\mathscr{F}_f$ and $\mathscr{F}_g$ are exhausting ideals for $\mathscr{F}$. By Lemma 2.6.7, $\mathscr{C} = \mathscr{F}_f \cap \mathscr{F}_g$ is an exhausting ideal for $\mathscr{F}$, and (2.25) holds by the definition of a conditional expectation.

(2)⇒(1): Choose a cover $(C_n)_{n\geqslant 1}$ of S by pairwise disjoint sets and fix a set $F \in \mathscr{F}_f \cap \mathscr{F}_g$. It follows from (2.25) and the fact that $F \cap C_n \in \mathscr{C}$ that

$$\int_{C_n \cap F} g \, \mathrm{d}\mu = \int_{C_n \cap F} f \, \mathrm{d}\mu.$$

Summing over n, the identity (2.24) follows by dominated convergence. □

Example 2.6.10. If f_1 and f_2 are σ-integrable and have conditional expectations g_1 and g_2, then for all scalars c_1 and c_2 the function $c_1g_1+c_2g_2$ is a conditional expectation for $c_1f_1+c_2f_2$. This easily follows from Proposition 2.6.9, noting that $\mathscr{C} := \mathscr{F}_{f_1}\cap\mathscr{F}_{f_2}\cap\mathscr{F}_{g_1}\cap\mathscr{F}_{g_2}$ is an exhausting ideal for $\mathscr{F}$ by Lemma 2.6.7. This example clearly shows the flexibility of the notion of an exhausting ideal. Working with the 'maximal' exhausting ideal $\mathscr{F}_{c_1f_1+c_2f_2}\cap\mathscr{F}_{c_1g_1+c_2g_2}$ would make the exercise much harder!

Example 2.6.11 (Restriction to sets in $\mathscr{F}$). If $f \in L^0(S;X)$ admits a conditional expectation g with respect to $\mathscr{F}$, then for every $A \in \mathscr{F}$, $\mathbf{1}_A g$ is a conditional expectation of $\mathbf{1}_A f$ with respect to $\mathscr{F}$. Indeed, for every $F \in \mathscr{F}_f \cap \mathscr{F}_g$ also $F \cap A \in \mathscr{F}_f \cap \mathscr{F}_g$ and

$$\int_F \mathbf{1}_A f \,\mathrm{d}\mu = \int_{F\cap A} f \,\mathrm{d}\mu = \int_{F\cap A} g \,\mathrm{d}\mu = \int_F \mathbf{1}_A g \,\mathrm{d}\mu.$$

Example 2.6.12 (Applying bounded operators). If $f \in L^0(S;X)$ admits a conditional expectation g with respect to $\mathscr{F}$, and if $T : X \to Y$ is a bounded operator, then Tg is a conditional expectation of Tf with respect to $\mathscr{F}$. The proof is similar to the previous one; this time one uses that $\mathscr{F}_f \cap \mathscr{F}_g$ is contained in $\mathscr{F}_{Tf} \cap \mathscr{F}_{Tg}$.

In particular it follows that for every $x^* \in X^*$ the function $\langle g, x^*\rangle$ is a conditional expectation of $\langle f, x^*\rangle$.

Example 2.6.13 (Averaging). For each $n \in \mathbb{Z}$, let $\mathscr{F}_n$ be the σ-algebra in $\mathbb{R}^d$ generated by the family of dyadic cubes $Q_k^n = 2^{-n}(k + [0,1)^d)$, $k \in \mathbb{Z}^d$. A function $f \in L^0(\mathbb{R}^d;X)$ is σ-integrable with respect to $\mathscr{F}_n$ if and only if $f \in L^1_{\rm loc}(\mathbb{R}^d;X)$. In this case, a conditional expectation of f with respect to $\mathscr{F}_n$ exists and is given by

$$\sum_{k\in\mathbb{Z}^d} \Big(\frac{1}{|Q_k^n|}\int_{Q_k^n} f(x)\,\mathrm{d}x\Big)\mathbf{1}_{Q_k^n}.$$

Example 2.6.14 (Even part of a function). Let $f \in L^0(\mathbb{R};X)$, where on $\mathbb{R}$ we take the Borel σ-algebra $\mathscr{B}(\mathbb{R})$ and the Lebesgue measure. Let $\mathscr{F} = \{B \in \mathscr{B}(\mathbb{R}) : -B = B\}$. Then $g : \mathbb{R} \to X$ defined by $g(t) = \frac{1}{2}(f(t) + f(-t))$ is a conditional expectation of f given $\mathscr{F}$. Indeed, $g \in L^0(S,\mathscr{F};X)$, $\mathscr{F}_f \subseteq \mathscr{F}_g$, and for any set in $B \in \mathscr{F}_f = \mathscr{F}_f \cap \mathscr{F}_g$ we have

$$\int_B f(t)\,\mathrm{d}t = \int_B f(-t)\,\mathrm{d}t = \int_B g(t)\,\mathrm{d}t.$$

2.6.a Uniqueness

Our task in this section is to prove the uniqueness of conditional expectations. We begin with some lemmas concerning scalar-valued functions.

We fix a measure space $(S,\mathscr{A},\mu)$ and let $\mathscr{F} \subseteq \mathscr{A}$ be sub-σ-algebra.

Lemma 2.6.15. *Let $f \in L^0(S, \mathscr{F})$ be given and suppose that $\mathscr{C} \subseteq \mathscr{F}_f$ is an exhausting ideal for $\mathscr{F}$. If $\int_C f \,\mathrm{d}\mu \geqslant 0$ (respectively, $= 0$, $\leqslant 0$) for all $C \in \mathscr{C}$, then $f \geqslant 0$ (respectively, $= 0$, $\leqslant 0$) almost everywhere.*

Proof. For any $\mathscr{F}$-measurable representative of f we have $F_k := \{f < -1/k\} \in \mathscr{F}$ for all $k \geqslant 1$. Fix $C \in \mathscr{C}$. Then $C \cap F_k \in \mathscr{C}$, the integrability of f on this set implies $\mu(C \cap F_k) < \infty$, and

$$0 \leqslant \int_{C\cap F_k} f \,\mathrm{d}\mu \leqslant -\frac{1}{k}\mu(C \cap F_k) \leqslant 0.$$

Hence $\mu(C \cap F_k) = 0$. Since $\{f < 0\} = \bigcup_{k\geqslant 1} F_k$, we see that $\mu(C \cap \{f < 0\}) = 0$. Applying this to the sets C in a countable cover of S, we see that $\mu(\{f < 0\}) = 0$, which is the same as saying that $f \geqslant 0$ almost everywhere.

The assertion concerning '$\leqslant 0$' follows by applying the previous case to $-f$, and the assertion concerning '$= 0$' follows by combining the two previous cases. □

Lemma 2.6.16. *Let $f \in L^0(S)$ be σ-integrable over $\mathscr{F}$. If $g \in L^0(S)$ is a conditional expectation of f with respect to $\mathscr{F}$, then $\mathscr{F}_f \subseteq \mathscr{F}_g$ and*

$$\int_F f \,\mathrm{d}\mu = \int_F g \,\mathrm{d}\mu \quad \forall F \in \mathscr{F}_f. \tag{2.26}$$

Proof. In the complex case we have $\mathscr{F}_f = \mathscr{F}_{\Re f} \cap \mathscr{F}_{\Im f}$ and $\mathscr{F}_g = \mathscr{F}_{\Re g} \cap \mathscr{F}_{\Im g}$, from which it is easy to see that $\Re g$ and $\Im g$ are conditional expectations of $\Re f$ and $\Im f$, respectively. It suffices therefore to prove the lemma over the real scalars.

By Proposition 2.6.9, $\mathscr{C} = \mathscr{F}_f \cap \mathscr{F}_g$ is an exhausting ideal for $\mathscr{F}$ and therefore we can choose a sequence $(C_n)_{n\geqslant 1}$ in $\mathscr{C}$ covering S. By the observation made after Definition 2.6.5, we may assume that the sets C_n are pairwise disjoint.

Fix a set $F \in \mathscr{F}_f$. Choosing an $\mathscr{F}$-measurable representative of g, we have $\{g \geqslant 0\} \in \mathscr{F}$, so that $F \cap C_n \cap \{g \geqslant 0\} \in \mathscr{C}$ and

$$\int_{F\cap C_n\cap\{g\geqslant 0\}} g \,\mathrm{d}\mu = \int_{F\cap C_n\cap\{g\geqslant 0\}} f \,\mathrm{d}\mu.$$

Summing over n, we obtain (by monotone convergence on the left-hand side and dominated convergence on the right-hand side)

$$\int_{F\cap\{g\geqslant 0\}} g \,\mathrm{d}\mu = \int_{F\cap\{g\geqslant 0\}} f \,\mathrm{d}\mu.$$

In the same way we prove that $\int_{F\cap\{g<0\}} g \,\mathrm{d}\mu = \int_{F\cap\{g<0\}} f \,\mathrm{d}\mu$ Therefore,

$$\int_F |g| \,\mathrm{d}\mu = \int_{F\cap\{g\geqslant 0\}} g \,\mathrm{d}\mu - \int_{F\cap\{g<0\}} g \,\mathrm{d}\mu$$

$$= \int_{F\cap\{g\geqslant 0\}} f \,\mathrm{d}\mu - \int_{F\cap\{g<0\}} f \,\mathrm{d}\mu \leqslant \int_F |f| d\mu < \infty$$

and

$$\int_F g \,\mathrm{d}\mu = \int_{F\cap\{g\geqslant 0\}} g \,\mathrm{d}\mu + \int_{F\cap\{g<0\}} g \,\mathrm{d}\mu = \int_F f \,\mathrm{d}\mu.$$

This first relation shows that $F \in \mathscr{F}_g$ and the second proves (2.26). □

Before turning to the proof of the uniqueness of conditional expectations we record a useful consequence of the previous two lemmas.

Lemma 2.6.17. *Let $f_1, f_2 \in L^0(S)$ be σ-integrable over $\mathscr{F}$ and $g_1, g_2 \in L^0(S)$ be conditional expectations of these functions with respect to $\mathscr{F}$. If $f_1 \leqslant f_2$ almost everywhere, then $g_1 \leqslant g_2$ almost everywhere.*

Proof. Under the stated assumptions, $g := g_2 - g_1$ is a conditional expectation of the non-negative function $f := f_2 - f_1 \geqslant 0$, and it suffices to prove that g is non-negative.

By Lemma 2.6.16, for all $F \in \mathscr{F}_f$ we have $F \in \mathscr{F}_g$ and

$$\int_F g \,\mathrm{d}\mu = \int_F f \,\mathrm{d}\mu \geqslant 0.$$

We now apply Lemma 2.6.15 (to g and $\mathscr{C} = \mathscr{F}_f$) to conclude that $g \geqslant 0$ almost everywhere. □

Theorem 2.6.18 (Uniqueness of conditional expectations). *Suppose that $f \in L^0(S;X)$ is σ-integrable with respect to $\mathscr{F}$. If $g \in L^0(S;X)$ and $\widetilde{g} \in L^0(S;X)$ are conditional expectations of f with respect to $\mathscr{F}$, then $g = \widetilde{g}$ almost everywhere.*

Proof. For any $x^* \in X^*$, both $\langle g, x^*\rangle$ and $\langle \widetilde{g}, x^*\rangle$ are conditional expectations of $\langle f, x^*\rangle$. Moreover, for all $F \in \mathscr{F}_{\langle f,x^*\rangle}$ we have $F \in \mathscr{F}_{\langle g,x^*\rangle} \cap \mathscr{F}_{\langle \widetilde{g},x^*\rangle}$ by Lemma 2.6.16, and therefore

$$\int_F \langle g, x^*\rangle \,\mathrm{d}\mu = \int_F \langle f, x^*\rangle \,\mathrm{d}\mu = \int_F \langle \widetilde{g}, x^*\rangle \,\mathrm{d}\mu.$$

Hence $\langle g, x^*\rangle = \langle \widetilde{g}, x^*\rangle$ almost everywhere by Lemma 2.6.15 (applied to $\langle g - \widetilde{g}, x^*\rangle$ and $\mathscr{C} = \mathscr{F}_{\langle f,x^*\rangle}$). An appeal to Corollary 1.1.25 concludes the proof. □

Thus a conditional expectation of a σ-integrable function $f \in L^0(S;X)$, if it exists, is unique as an element of $L^0(S,\mathscr{F};X)$. This allows us to speak of *the* conditional expectation of f with respect to $\mathscr{F}$.

Notation. The conditional expectation of a σ-integrable function $f \in L^0(S;X)$ with respect to $\mathscr{F}$, if it exists, will be denoted by $\mathbb{E}(f|\mathscr{F})$.

For later use we state the following simple lemma.

Lemma 2.6.19. *Let $f \in L^0(S;X)$ be σ-integrable over $\mathscr{F} \subseteq \mathscr{A}$. If the conditional expectations of f and $\|f\|$ with respect to $\mathscr{F}$ exist, then almost everywhere we have*

$$\|\mathbb{E}(f|\mathscr{F})\| \leqslant \mathbb{E}(\|f\||\mathscr{F}).$$

Proof. By forgetting the complex scalar multiplication we may assume that $\mathbb{K} = \mathbb{R}$. Since a strongly measurable function takes its values in a separable subspace almost everywhere, we may assume that X is separable. Pick a norming sequence $(x_n^*)_{n\geqslant 1}$ of unit vectors in X^*.

By Example 2.6.12, the conditional expectation of $\langle f, x_n^*\rangle$ exists for every $n \geqslant 1$, and

$$\|\mathbb{E}(f|\mathscr{F})\| = \sup_{n\geqslant 1}\langle \mathbb{E}(f|\mathscr{F}), x_n^*\rangle = \sup_{n\geqslant 1}\mathbb{E}(\langle f, x_n^*\rangle|\mathscr{F}) \leqslant \mathbb{E}(\|f\||\mathscr{F})$$

almost everywhere, where the last step follows from Lemma 2.6.17 since $\langle f, x_n^*\rangle \leqslant \|f\|$ almost everywhere. $\square$

2.6.b Existence

We now turn to the existence problem for conditional expectations. As always we fix a measure space $(S, \mathscr{A}, \mu)$ and let $\mathscr{F} \subseteq \mathscr{A}$ be sub-σ-algebra.

Theorem 2.6.20 (Existence of conditional expectations). *If the function $f \in L^0(S;X)$ is σ-integrable over $\mathscr{F}$ and μ is σ-finite on the sub-σ-algebra $\mathscr{F}$, then f admits a conditional expectation with respect to $\mathscr{F}$. In this case, the conditional expectation is uniquely defined as an element of $L^0(S, \mathscr{F}; X)$. Denoting it by $\mathbb{E}(f|\mathscr{F})$, we have $\mathscr{F}_f \subseteq \mathscr{F}_{\mathbb{E}(f|\mathscr{F})}$ and*

$$\int_F f \,\mathrm{d}\mu = \int_F \mathbb{E}(f|\mathscr{F}) \,\mathrm{d}\mu \quad \forall F \in \mathscr{F}_f. \tag{2.27}$$

Furthermore, if $(F_n)_{n\geqslant 1}$ is an exhausting sequence for f in $\mathscr{F}$ consisting of disjoint sets, then

$$\sum_{n\geqslant 1} \mathbf{1}_{F_n}\mathbb{E}(\mathbf{1}_{F_n} f|\mathscr{F}). \tag{2.28}$$

As we have already noted in the previous subsection, the σ-integrability over $\mathscr{F}$ of the constant function $\mathbf{1}$ is equivalent to the σ-finiteness of μ on $\mathscr{F}$. It is for this reason that in the sequel we only consider conditional expectations with respect to σ-finite σ-algebras. We will also provide an example of the anomalies that arise in the absence of this assumption.

The proof of Theorem 2.6.20 will be accomplished in several steps. The starting point is provided by the next lemma.

Lemma 2.6.21 (Existence in $L^2(S)$). *If μ is σ-finite on the sub-σ-algebra $\mathscr{F}$, then every $f \in L^2(S)$ has a unique conditional expectation with respect to $\mathscr{F}$. It belongs to $L^2(S)$ and is given by Pf, where P be the orthogonal projection of $L^2(S)$ onto $L^2(S, \mathscr{F})$.*

Proof. Since μ is assumed to be σ-finite on $\mathscr{F}$, the family $\mathscr{F}_{\mathbf{1}}$ (i.e., $\mathscr{F}_h$ for the constant function $h = \mathbf{1}$) is an exhausting ideal for $\mathscr{F}$. Moreover, $\mathscr{F}_{\mathbf{1}} \subseteq \mathscr{F}_f \cap \mathscr{F}_{Pf}$ by the fact that $f, Pf \in L^2(S)$ and the Cauchy–Schwarz inequality. Finally, for $F \in \mathscr{F}_{\mathbf{1}}$, using that orthogonal projections are self-adjoint, we have

$$\int_F Pf \,\mathrm{d}\mu = \int_S \mathbf{1}_F(Pf) \,\mathrm{d}\mu = \int_S (P\mathbf{1}_F)f \,\mathrm{d}\mu = \int_S \mathbf{1}_F f \,\mathrm{d}\mu = \int_F f \,\mathrm{d}\mu.$$

This completes the proof. □

Lemma 2.6.22 (Existence in $L^1(S)$). *If μ is σ-finite on the sub-σ-algebra $\mathscr{F}$, then every $f \in L^1(S)$ has a unique conditional expectation with respect to $\mathscr{F}$. It belongs to $L^1(S)$ and satisfies*

$$\|\mathbb{E}(f|\mathscr{F})\|_1 \leqslant \|f\|_1$$

and

$$\int_F \mathbb{E}(f|\mathscr{F}) \,\mathrm{d}\mu = \int_F f \,\mathrm{d}\mu \quad \forall F \in \mathscr{F}.$$

Proof. First let $f \in L^1(S) \cap L^\infty(S)$. Then $f, |f| \in L^2(S)$, so their conditional expectations exist by Lemma 2.6.21, and Lemma 2.6.19 shows that

$$\int_S |\mathbb{E}(f|\mathscr{F})| \,\mathrm{d}\mu \leqslant \int_S \mathbb{E}(|f||\mathscr{F}) \,\mathrm{d}\mu = \int_S |f| \,\mathrm{d}\mu,$$

where the last step follows from (2.26) and the fact that $S \in \mathscr{F}_{|f|}$ for $f \in L^1(S)$.

This proves that the conditional expectation is contractive with respect to the L^1-norm on the dense subspace $L^1(S) \cap L^\infty(S)$ of $L^1(S)$. As a consequence, this operator admits a unique extension to a contraction, denoted by P, on $L^1(S)$ which takes its values in the closed subspace $L^1(S, \mathscr{F})$ of $L^1(S)$. To see that for any $f \in L^1(S)$, Pf is the conditional expectation of f, we set $f_n = \mathbf{1}_{\{|f| \leqslant n\}} f$. Then $f_n \in L^1(S) \cap L^\infty(S)$, $f_n \to f$ in $L^1(S)$, and for all $F \in \mathscr{F}$ we have

$$\begin{aligned}\int_F Pf \,\mathrm{d}\mu = \lim_{n\to\infty} \int_F Pf_n \,\mathrm{d}\mu &= \lim_{n\to\infty} \int_F \mathbb{E}(f_n|\mathscr{F}) \,\mathrm{d}\mu \\ &= \lim_{n\to\infty} \int_F f_n \,\mathrm{d}\mu = \int_F f \,\mathrm{d}\mu,\end{aligned}$$

where the second-to-last step used (2.26) and the fact that $\mathscr{F}_{f_n} = \mathscr{F}$. □

After these scalar-valued auxiliary results, we return to the realm of vector-valued functions:

Theorem 2.6.23 (Existence in $L^1(S;X)$). *If μ is σ-finite on the sub-σ-algebra $\mathscr{F}$, then every $f \in L^1(S;X)$ admits a unique conditional expectation with respect to $\mathscr{F}$. It belongs to in $L^1(S;X)$ and satisfies*

$$\|\mathbb{E}(f|\mathscr{F})\|_1 \leqslant \|f\|_1$$

and

$$\int_F \mathbb{E}(f|\mathscr{F})\,\mathrm{d}\mu = \int_F f\,\mathrm{d}\mu \quad \forall F \in \mathscr{F}.$$

Proof. By Lemma 2.6.22, the operator $\mathbb{E}(\cdot|\mathscr{F})$ is a well defined contraction on $L^1(S)$, which is positive by Lemma 2.6.17. By Theorem 2.1.3, it extends to a contraction on $L^1(S;X)$. To see that this extension is indeed the conditional expectation on $L^1(S;X)$ we note that for μ-simple functions this follows from the scalar case (it is immediate from the definition of conditional expectations that $\mathbb{E}(g\otimes x|\mathscr{F}) = \mathbb{E}(g|\mathscr{F})\otimes x$), and for the general case we approximate f by a sequence of μ-simple functions f_n and argue as in the proof of Lemma 2.6.22.

The final identity follows from the scalar case by a similar approximation, noting that it is true for μ-simple functions since it is true for scalar functions by Lemma 2.6.22. □

Proof of Theorem 2.6.20. First we prove that if μ is σ-finite on the sub-σ-algebra $\mathscr{F}$, then every $f \in L^0(S;X)$ that is σ-integrable on $\mathscr{F}$ admits a conditional expectation with respect to $\mathscr{F}$.

Fix an arbitrary exhausting sequence $(F_n)_{n\geqslant 1}$ for f in $\mathscr{F}$ consisting of disjoint sets. Since integrable functions have a conditional expectation by Theorem 2.6.23, the function

$$\sum_{n\geqslant 1} \mathbf{1}_{F_n}\mathbb{E}(\mathbf{1}_{F_n} f|\mathscr{F})$$

is well defined. By Proposition 2.6.9 $\mathscr{C}_n = \mathscr{F}_{\mathbf{1}_{F_n} f} \cap \mathscr{F}_{\mathbb{E}(\mathbf{1}_{F_n} f|\mathscr{F})}$ is an exhausting ideal in $\mathscr{F}$. It is easily checked that $\mathscr{C} := \bigcup_{n\geqslant 1} \mathscr{C}_n|_{F_n}$ is an exhausting ideal in $\mathscr{F}$, where $\mathscr{C}_n|_{F_n} := \{C \cap F_n : C \in \mathscr{C}_n\}$. Proposition 2.6.9 then implies that that the function just defined is a conditional expectation of f.

The proof of Theorem 2.6.20 is completed by checking that the assertion (2.27) holds true. Let $F \in \mathscr{F}_f$, so that $\mathbf{1}_F f \in L^1(S;X)$ by definition. Theorem 2.6.23 guarantees that this function has a conditional expectation $h \in L^1(S;X)$ and

$$\int_S h\,\mathrm{d}\mu = \int_S \mathbf{1}_F f\,\mathrm{d}\mu = \int_F f\,\mathrm{d}\mu.$$

On the other hand, by Example 2.6.11, the conditional expectation of $\mathbf{1}_F f$, which we denoted by h, is given by $\mathbf{1}_F\mathbb{E}(f|\mathscr{F})$. So in fact we checked that $\mathbf{1}_F\mathbb{E}(f|\mathscr{F}) \in L^1(S;X)$ and

$$\int_F \mathbb{E}(f|\mathscr{F})\,\mathrm{d}\mu = \int_S h\,\mathrm{d}\mu = \int_F f\,\mathrm{d}\mu,$$

as we wanted to prove. □

Examples of non-existence

We now give an example to demonstrate the failure of existence of conditional expectations, even for functions $f \in L^1(S)$ (which is σ-integrable over any sub-σ-algebra), when μ is not assumed to be σ-finite on $\mathscr{F}$.

Example 2.6.24. Let S be a set with a distinct point s_1 and equipped with a measure μ such that $\mu\{s_1\} = 1$ and $\mu(S \setminus \{s_1\}) = \infty$. The function $f(s_1) = a$, $f(s) = 0$ for $s \neq s_1$ is σ-integrable over $\mathscr{F} = \{\varnothing, S\}$ but has a conditional expectation with respect to $\mathscr{F}$ if and only if $a = 0$ (in which case the conditional expectation is the zero function). Indeed, the conditional expectation has to be a constant function which, by the requirement of strong μ-measurability, can only be the zero function. For $a \neq 0$ the defining condition (2.24) for conditional expectations is clearly violated.

2.6.c Conditional limit theorems

We proceed with "conditional" versions of the fundamental limit theorems of integration theory. Their proofs follow the same logical pattern as for their unconditional counterparts: we start with monotone convergence and arrive, via Fatou's lemma, at dominated convergence.

Theorem 2.6.25 (Conditional monotone convergence). *Let $(f_n)_{n\geqslant 1}$ be a non-decreasing sequence of non-negative functions in $L^0(S)$ and let $f \in L^0(S)$ be σ-integrable over a sub-σ-algebra $\mathscr{F}$ on which μ is σ-finite. If $f_n \uparrow f$ almost everywhere, then*

$$\mathbb{E}(f_n|\mathscr{F}) \uparrow \mathbb{E}(f|\mathscr{F}),$$

the existence of the conditional expectations being part of the assertion.

Proof. The σ-integrability of f over $\mathscr{F}$ implies the same property for each f_n, and hence the existence of the conditional expectations is guaranteed by Theorem 2.6.20.

By monotonicity and the positivity of the conditional expectation operator it follows that we have the almost everywhere inequalities

$$0 \leqslant \mathbb{E}(f_n|\mathscr{F}) \leqslant \mathbb{E}(f_{n+1}|\mathscr{F}) \leqslant \mathbb{E}(f|\mathscr{F}).$$

Hence $(\mathbb{E}(f_n|\mathscr{F}))_{n\geqslant 1}$ is an almost everywhere bounded non-decreasing sequence, so it has an almost everywhere pointwise limit $g \in L^0(S, \mathscr{F})$ which satisfies $0 \leqslant g \leqslant \mathbb{E}(f|\mathscr{F})$. It remains to be shown that $g = \mathbb{E}(f|\mathscr{F})$.

For all $F \in \mathscr{F}_f$ we have $F \in \mathscr{F}_{f_n}$ and

$$\int_F g \,\mathrm{d}\mu = \lim_{n\to\infty} \int_F \mathbb{E}(f_n|\mathscr{F}) \,\mathrm{d}\mu = \lim_{n\to\infty} \int_F f_n \,\mathrm{d}\mu = \int_F f \,\mathrm{d}\mu,$$

where the first and last steps were based on the usual monotone convergence theorem. This gives the desired result. $\square$

The σ-integrability assumption on f cannot be omitted:

Example 2.6.26. Let $S = (0,1)$ with the Borel σ-algebra and Lebesgue measure, let $\mathscr{F} = \{\varnothing, (0,1)\}$, and consider the functions $f(t) = 1/t$ and $f_n(t) = \mathbf{1}_{(1/n,1)}f$. Each f_n has a conditional expectation with respect to $\mathscr{F}$, but f does not.

Theorem 2.6.27 (Conditional Fatou lemma). *Let μ be σ-finite on the sub-σ-algebra $\mathscr{F}$, let $(f_n)_{n\geqslant 1}$ be a sequence of non-negative functions in $L^0(S)$, let $f \in L^0(S)$, and assume that all these functions are σ-integrable over $\mathscr{F}$. Suppose furthermore that $f = \liminf_{n\to\infty} f_n$ almost everywhere. Then, almost everywhere,*

$$\mathbb{E}(f|\mathscr{F}) \leqslant \liminf_{n\to\infty} \mathbb{E}(f_n|\mathscr{F}).$$

Proof. Putting $\phi_n := \inf_{m\geqslant n} f_m$ we have $0 \leqslant \phi_n \uparrow f = \sup_{n\geqslant 1} \inf_{m\geqslant n} f_m$. By the conditional monotone convergence theorem,

$$\begin{aligned}\mathbb{E}(f|\mathscr{F}) &= \mathbb{E}(\lim_{n\to\infty} \phi_n|\mathscr{F}) \\ &= \lim_{n\to\infty} \mathbb{E}(\phi_n|\mathscr{F}) \leqslant \lim_{n\to\infty} \inf_{m\geqslant n} \mathbb{E}(f_m|\mathscr{F}) = \liminf_{n\to\infty} \mathbb{E}(f_n|\mathscr{F}).\end{aligned}$$

The estimate $\leqslant$ above was based on the fact that $\phi_n \leqslant f_m$ for all $m \geqslant n$ and hence $\mathbb{E}(\phi_n|\mathscr{F}) \leqslant \mathbb{E}(f_m|\mathscr{F})$ for all $m \geqslant n$. □

Theorem 2.6.28 (Conditional dominated convergence). *Let μ be σ-finite on the sub-σ-algebra $\mathscr{F}$. Let $(f_n)_{n\geqslant 1}$ be a sequence in $L^0(S;X)$ and suppose that $\lim_{n\to\infty} f_n = f$ and $\|f_n\| \leqslant g$ almost everywhere, where $g \in L^0(S)$ is σ-integrable over $\mathscr{F}$. Then, almost everywhere*

$$\lim_{n\to\infty} \mathbb{E}(\|f_n - f\||\mathscr{F}) = 0,$$

and

$$\lim_{n\to\infty} \mathbb{E}(f_n|\mathscr{F}) = \mathbb{E}(f|\mathscr{F}),$$

the existence of these conditional expectations being part of the assertion.

Proof. The existence of the conditional expectations of f and f_n follows from Theorem 2.6.20 and the fact that they, too, must be σ-integrable over $\mathscr{F}$.

Let $g_n := \|f_n - f\|$. Then almost everywhere we have $g_n \to 0$ and $0 \leqslant g_n \leqslant 2g$. Let $\psi_n = \sup_{m\geqslant n} g_m$. Then $0 \leqslant \psi_n \leqslant 2g$ and $\psi_n \downarrow 0$ almost everywhere. Hence $\mathbb{E}(2g - \psi_n|\mathscr{F}) \uparrow \mathbb{E}(2g|\mathscr{F})$ almost everywhere by the conditional monotone convergence theorem and therefore $\mathbb{E}(\psi_n|\mathscr{F}) \downarrow 0$ almost everywhere as $n \to \infty$. In view of $0 \leqslant g_n \leqslant \psi_n$, this proves that $\mathbb{E}(g_n|\mathscr{F}) \downarrow 0$ almost everywhere. Now the final assertion can be derived if one uses Lemma 2.6.19. □

2.6.d Inequalities and identities

The following key inequality is the conditional analogue of Proposition 1.2.11 and extends Lemma 2.6.19.

Proposition 2.6.29 (Conditional Jensen's inequality). *Let μ be σ-finite on the sub-σ-algebra $\mathscr{F}$. Let $\phi : X \to \mathbb{R}$ be convex and lower semi continuous, let $f \in L^0(S;X)$, and suppose that f and $\phi \circ f$ are σ-integrable over $\mathscr{F}$. Then, almost everywhere,*

$$\phi \circ \mathbb{E}(f|\mathscr{F}) \leqslant \mathbb{E}(\phi \circ f|\mathscr{F}).$$

Proof. Since f takes its values in a separable subspace almost everywhere, we may assume that X is separable. Then Lemma 1.2.10 provides a sequence of affine functions $\phi_i(x) = \Re\langle x, y_i^*\rangle + a_i$ on X such that $\phi(x) = \sup_i \phi_i(x)$. Thus, almost everywhere,

$$\phi \circ \mathbb{E}(f|\mathscr{F}) = \sup_i \phi_i \circ \mathbb{E}(f|\mathscr{F}) = \sup_i \mathbb{E}(\phi_i \circ f|\mathscr{F}) \leqslant \mathbb{E}(\phi \circ f|\mathscr{F}).$$

□

Corollary 2.6.30 (L^p-contractivity). *Let μ be σ-finite on the sub-σ-algebra $\mathscr{F}$. Let $p \in [1,\infty)$ and suppose that $f \in L^p(S;X)$. Then the conditional expectations below exist, and satisfy the almost everywhere inequality*

$$\|\mathbb{E}(f|\mathscr{F})\|^p \leqslant \mathbb{E}(\|f\|^p|\mathscr{F}). \tag{2.29}$$

In particular, we have $\mathbb{E}(f|\mathscr{F}) \in L^p(S;X)$ and

$$\|\mathbb{E}(f|\mathscr{F})\|_p \leqslant \|f\|_p.$$

Proof. Both f and $\|f\|^p$ are σ-integrable over $\mathscr{F}$ (see Proposition 2.6.3). Therefore, the conditional expectations exist. Now (2.29) follows from Jensen's inequality applied to $\phi(x) = \|x\|^p$, and the norm inequality follows by integrating (2.29) over S and applying Lemma 2.6.22 on the right. □

We continue with some useful identities for conditional expectations. In Propositions 2.6.31 and 2.6.32 we let X_1, X_2, Y be Banach spaces and

$$\beta : X_1 \times X_2 \to Y$$

be a bounded bilinear mapping. The main examples we have in mind are

$$\begin{aligned} &X_1 = \mathbb{K}, \;\; X_2 = X, \;\; Y = X, \;\; \beta(c,x) = cx && \text{(scalar multiplication);} \\ &X_1 = X, \;\; X_2 = X^*, \;\; Y = \mathbb{K}, \;\; \beta(x,x^*) = \langle x, x^*\rangle && \text{(duality).} \end{aligned}$$

Proposition 2.6.31 (Taking out $\mathscr{F}$-measurable terms). *Let μ be σ-finite on the sub-σ-algebra $\mathscr{F}$. Suppose that $g \in L^0(S;\mathscr{F};X_1)$ and that $f \in L^0(S;X_2)$ is σ-integrable over $\mathscr{F}$. Then $\beta(g,f) \in L^0(S;Y)$ is σ-integrable over $\mathscr{F}$, and almost everywhere we have*

$$\mathbb{E}(\beta(g,f)|\mathscr{F}) = \beta(g, \mathbb{E}(f|\mathscr{F})).$$

Proof. Choosing an exhausting sequence of sets $F_i \in \mathscr{F}_f$, which exist by σ-integrability, the function $\beta(g, f)$ is integrable over any of the sets $F_i \cap \{\|g\| \leqslant j\} \in \mathscr{F}$, which form a countable cover of S. Thus $\beta(g, f)$ is σ-integrable over $\mathscr{F}$, which implies the existence of the conditional expectation by Theorem 2.6.20.

To prove the identity, let us first consider the case that $g = \sum_{i=1}^n x_i \mathbf{1}_{F_i}$ is simple. Then, for $F \in \mathscr{F}_f \subseteq \mathscr{F}_{\mathbb{E}(f|\mathscr{F})}$ we simply compute

$$\begin{aligned}
\int_F \beta(g, \mathbb{E}(f|\mathscr{F}))\,\mathrm{d}\mu &= \sum_{i=1}^n \beta\Big(x_i, \int_{F\cap F_i} \mathbb{E}(f|\mathscr{F})\,\mathrm{d}\mu\Big) \\
&= \sum_{i=1}^n \beta\Big(x_i, \int_{F\cap F_i} f\,\mathrm{d}\mu\Big) = \int_F \beta(g, f)\,\mathrm{d}\mu,
\end{aligned}$$

which shows by definition that $\beta(g, \mathbb{E}(f|\mathscr{F}))$ is a conditional expectation of $\beta(f, g)$ as $\mathscr{F}_f$ is an exhausting ideal for $\mathscr{F}$..

In the general case, we pick a sequence of simple functions

$$g_n \in L^0(S; \mathscr{F}; X_1)$$

such that $g_n \to g$ pointwise and $\|g_n\| \leqslant \|g\|$. Then $\beta(g_n, h) \to \beta(g, h)$ pointwise for $h \in \{f, \mathbb{E}(f|\mathscr{F})\}$, and

$$\|\beta(g_n, f)\| \leqslant \|\beta\| \cdot \|g_n\| \cdot \|f\| \leqslant \|\beta\| \cdot \|g\| \cdot \|f\|,$$

which is a σ-integrable over $\mathscr{F}$, as shown by the sets $F_i \cap \{\|g\| \leqslant j\}$ from the beginning of the proof. Thus the conditional dominated convergence theorem proves that

$$\begin{aligned}
\mathbb{E}(\beta(g, f)|\mathscr{F}) &= \lim_{n\to\infty} \mathbb{E}(\beta(g_n, f)|\mathscr{F}) \\
&= \lim_{n\to\infty} \beta(g_n, \mathbb{E}(f|\mathscr{F})) = \beta(g, \mathbb{E}(f|\mathscr{F})).
\end{aligned}$$

□

Another version of this result, where the function g is allowed to be operator-valued, will be proved in the next chapter (Lemma 3.5.2).

As an application we prove that the conditional expectation has a useful self-adjointness property.

Proposition 2.6.32 (Self-adjointness property). *Suppose that μ is σ-finite on the sub-σ-algebra $\mathscr{F}$. Let $f_1 \in L^0(S; X_1)$ and $f_2 \in L^0(S; X_2)$ be σ-integrable over $\mathscr{F}$. If both $\beta(f_1, \mathbb{E}(f_2|\mathscr{F}))$ and $\beta(\mathbb{E}(f_1|\mathscr{F}), f_2)$ are integrable on some $F \in \mathscr{F}$, then so is $\beta(\mathbb{E}(f_1|\mathscr{F}), \mathbb{E}(f_2|\mathscr{F}))$, and we have*

$$\int_F \beta(f_1, \mathbb{E}(f_2|\mathscr{F}))\,\mathrm{d}\mu = \int_F \beta(\mathbb{E}(f_1|\mathscr{F}), f_2)\,\mathrm{d}\mu = \int_F \beta(\mathbb{E}(f_1|\mathscr{F}), \mathbb{E}(f_2|\mathscr{F}))\,\mathrm{d}\mu.$$

Proof. Denote $g_i := \mathbb{E}(f_i|\mathscr{F})$ for $I = 1, 2$. By Proposition 2.6.31, $\beta(f_1, g_2)$ has a conditional expectation with respect to $\mathscr{F}$ given by

$$\mathbb{E}(\beta(f_1, g_2)|\mathscr{F}) = \beta(g_1, g_2).$$

The left side, and thus the right, is integrable over $F \in \mathscr{F}_{\beta(f_1,g_2)} \cap \mathscr{F}_{\beta(g_1,f_2)}$, and we have

$$\int_F \beta(f_1, g_2)\,\mathrm{d}\mu = \int_F \mathbb{E}(\beta(f_1, g_2)|\mathscr{F})\,\mathrm{d}\mu = \int_F \beta(g_1, g_2)\,\mathrm{d}\mu.$$

The other identity follows by exchanging the roles of f_1 and f_2, observing that the right side is symmetric with respect to both f_1 and f_2. □

Proposition 2.6.33 (Tower property). *Let μ be σ-finite on $\mathscr{G}$, where $\mathscr{G} \subseteq \mathscr{F}$ are sub-σ-algebras. If $h \in L^0(S; X)$ is σ-integrable over $\mathscr{G}$, then the conditional expectations below exist and satisfy*

$$\mathbb{E}(\mathbb{E}(h|\mathscr{F})|\mathscr{G}) = \mathbb{E}(h|\mathscr{G}).$$

Remark 2.6.34. If we drop the condition of σ-finiteness on $\mathscr{G}$ it may happen that exactly one of $\mathbb{E}(h|\mathscr{F})$ and $\mathbb{E}(h|\mathscr{G})$ exists. Let $(S, \mathscr{A}, \mu) = (\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathrm{d}x)$, $\mathscr{G} = \{\varnothing, \mathbb{R}\}$, $\mathscr{F} = \sigma([0, \infty))$ and $h(t) = \mathrm{sgn}(t)/(1 + t^2)$, we have $\mathbb{E}(h|\mathscr{G}) = 0$ but $\mathbb{E}(h|\mathscr{F})$ does not exist. For the same $\mathscr{G}$, $\mathscr{F} = \sigma(\{[k, k + 1), k \in \mathbb{Z}\})$ and $h = \mathbf{1}$, we have $\mathbb{E}(h|\mathscr{F}) = h$, whereas $\mathbb{E}(h|\mathscr{G})$ does not exist. Note that in the second example the measure is σ-finite on $\mathscr{F}$.

Proof. The existence of $g := \mathbb{E}(h|\mathscr{G})$ and $f := \mathbb{E}(h|\mathscr{F})$ follows by Theorem 2.6.20 since h is σ-integrable over $\mathscr{G}$, and hence over $\mathscr{F}$, and μ is σ-finite on $\mathscr{G}$, and hence on $\mathscr{F}$. Moreover,

$$\mathscr{G}_f = \mathscr{F}_f \cap \mathscr{G} \supseteq \mathscr{F}_h \cap \mathscr{G} = \mathscr{G}_h$$

so that f is also σ-integrable over $\mathscr{G}$ and hence has a conditional expectation.

For $G \in \mathscr{G}_h \subseteq \mathscr{F}_h$, we have

$$\int_G g\,\mathrm{d}\mu = \int_G \mathbb{E}(h|\mathscr{G})\,\mathrm{d}\mu = \int_G h\,\mathrm{d}\mu = \int_G \mathbb{E}(h|\mathscr{F})\,\mathrm{d}\mu = \int_G f\,\mathrm{d}\mu,$$

and thus g is the conditional expectation of f with respect to $\mathscr{G}$, as we wanted to show. □

Independence and conditional expectations

The next two propositions are concerned with independence. Readers not familiar with this notion from probability theory may safely skip them.

Proposition 2.6.35 (Independence I). *Let $(S, \mathscr{A}, \mu)$ be a probability space. If $f \in L^1(S; X)$ is independent of the sub-σ-algebra $\mathscr{F}$, then its conditional expectation with respect to $\mathscr{F}$ is given by the constant function $\mathbb{E}(f|\mathscr{F}) = \mathbb{E}f$.*

Proof. The result is a consequence of the following identities, valid for all $F \in \mathscr{F}$:

$$\int_F f \, \mathrm{d}\mu = \mathbb{E}(\mathbf{1}_F f) = \mathbb{E}(\mathbf{1}_F)\mathbb{E}f = \int_F \mathbb{E}f \, \mathrm{d}\mu.$$

□

A more refined version of this proposition reads as follows.

Proposition 2.6.36 (Independence II). *Let $(S, \mathscr{A}, \mu)$ be a probability space and let $\mathscr{F}$ and $\mathscr{G}$ be sub-σ-algebras. Let $f \in L^0(S; X)$ be σ-integrable over $\mathscr{F}$. If $\mathscr{G}$ and $\sigma(f, \mathscr{F})$ are independent, then f has a conditional expectation with respect to $\sigma(\mathscr{G}, \mathscr{F})$ and*

$$\mathbb{E}(f|\sigma(\mathscr{G}, \mathscr{F})) = \mathbb{E}(f|\mathscr{F}).$$

Proof. For all $F \in \mathscr{F}_f$ and $G \in \mathscr{G}$ we have $F \in \mathscr{F}_{\mathbb{E}(f|\mathscr{F})}$ and

$$\begin{aligned}\int_{F\cap G} \mathbb{E}(f|\mathscr{F}) \, \mathrm{d}\mu &= \int_S \mathbf{1}_G \cdot (\mathbf{1}_F \mathbb{E}(f|\mathscr{F})) \, \mathrm{d}\mu \\ &\overset{\text{(i)}}{=} \mu(G) \int_F \mathbb{E}(f|\mathscr{F}) \, \mathrm{d}\mu = \mu(G) \int_F f \, \mathrm{d}\mu \\ &\overset{\text{(ii)}}{=} \int_S \mathbf{1}_G \cdot (\mathbf{1}_F f) \, \mathrm{d}\mu = \int_{F\cap G} f \, \mathrm{d}\mu.\end{aligned}$$

In (i) we used the independence of $\mathscr{G}$ and $\mathscr{F}$, in (ii) the independence of $\mathscr{G}$ and $\mathbf{1}_F f$. Let

$$\mathscr{C} = \{H \in \sigma(\mathscr{F}, \mathscr{G}) : \ H \subseteq F \cap G \text{ for some } F \in \mathscr{F}_f \text{ and } G \in \mathscr{G}\}.$$

Since S can be covered by countably many sets in $\mathscr{F}_f$, S can be covered by countably many sets in $\mathscr{C}$. It follows that $\mathscr{C}$ is an exhausting ideal for $\mathscr{H} := \sigma(\mathscr{F}, \mathscr{G})$.

By Proposition 2.6.9 It remains to be shown that for all $C \in \mathscr{C}$ we have

$$\int_C \mathbb{E}(f|\mathscr{F}) \, \mathrm{d}\mu = \int_C f \, \mathrm{d}\mu.$$

Fix a set $C \in \mathscr{C}$. By assumption we have $C \subseteq F_0 \cap G_0$ for some $F_0 \in \mathscr{F}_f$ and $G_0 \in \mathscr{G}$. Fix this set G_0 and consider the restricted σ-algebra $\mathscr{H}|_{G_0}$ on G_0. The collection of all $F \cap G$ with $F \in \mathscr{F}_f$ and $G \subseteq G_0$ is closed under taking finite intersections and generates $\mathscr{H}|_{G_0}$. Hence by Dynkin's Lemma A.1.3 we conclude that

$$\int_H \mathbb{E}(f|\mathscr{F}) \, \mathrm{d}\mu = \int_H f \, \mathrm{d}\mu \quad \forall H \in \mathscr{H}|_{G_0}.$$

Since $C \in \mathscr{H}|_{G_0}$, this concludes the proof. □

As an illustration we include the following lemma which will be used in Theorems 3.3.10 and 4.4.1.

Lemma 2.6.37. *Let $(S,\mathscr{A},\mu)$ be a probability space. Let $\xi_1,\dots,\xi_n$ be independent and identically distributed random variables in $L^0(S;X)$. Let*

$$\overline{\xi}_n := \frac{1}{k}\sum_{k=1}^n \xi_k.$$

Let $\mathscr{G}\subseteq\mathscr{A}$ be a σ-algebra which is independent of $\sigma(\xi_1,\overline{\xi}_n)$. Then

$$\mathbb{E}(\xi_j|\sigma(\overline{\xi}_n,\mathscr{G})) = \mathbb{E}(\xi_j|\sigma(\overline{\xi}_n)) = \overline{\xi}_n$$

for all $j=1,\dots,n$.

Proof. The first identity follows from Proposition 2.6.36. Turing to the second identity, we claim that for any Borel set $B\subseteq X$ the following identity holds

$$\int_{\{\overline{\xi}_n\in B\}} \xi_j \,\mathrm{d}\mathbb{P} = \int_{\{\overline{\xi}_n\in B\}} \xi_k \,\mathrm{d}\mathbb{P}, \quad 1\leqslant j,k\leqslant n. \tag{2.30}$$

Taking the average on the right-hand side of (2.30) with respect to k, keeping j fixed, the identity $\mathbb{E}(\xi_j|\sigma(\overline{\xi}_n)) = \overline{\xi}_n$ follows.

To prove (2.30) it suffices to take $j=1$. For $k=1,\dots,n$ let $h_k : X^n\to X$ be given by

$$h_k(x_1,\dots,x_n) := \mathbf{1}_{g(x_1,\dots,x_n)\in B}\otimes x_k \quad\text{with}\quad g(x_1,\dots,x_n) := \frac{1}{n}\sum_{j=1}^n x_j.$$

Then the following symmetry property holds: $h_j\circ\phi_{j,k} = h_k$, where $\phi_{j,k} : X^n\to X^n$ flips the jth and kth variable.

Let $\nu = \mathbb{P}\circ(\xi_1,\dots,\xi_n)^{-1}$ be the image measure on X^n. By the properties of the ξ_k's we have $\nu\circ\phi_{j,k}^{-1} = \nu$. Therefore, by substitution (see Proposition 1.2.6), we find

$$\begin{aligned}\int_{\{\overline{\xi}_n\in B\}} \xi_1 \,\mathrm{d}\mathbb{P} &= \int_{X^n} h_1\,\mathrm{d}\nu = \int_{X^n} h_1\,\mathrm{d}\nu\circ\phi_{1,k}^{-1}\\ &= \int_{X^n} h_1\circ\phi_{1,k}\,\mathrm{d}\nu = \int_{X^n} h_k\,\mathrm{d}\nu = \int_{\{\overline{\xi}_n\in B\}} \xi_k\,\mathrm{d}\mathbb{P}.\end{aligned}$$

□

Uniform integrability of conditional expectations

Definition 2.6.38 (Uniform integrability). *Let $1\leqslant p<\infty$. A subset $T\subseteq L^p(S;X)$ is said to be* uniformly p-integrable *(or just* uniformly integrable *when $p=1$) if*

(i) *For all $\varepsilon>0$, there exists an $r>0$ such that*

$$\sup_{f\in T}\int_S \mathbf{1}_{\{\|f\|>r\}}\|f\|^p\,\mathrm{d}\mu \leqslant \varepsilon.$$

(ii) *For every $\varepsilon > 0$, there is a set $B \in \mathscr{A}$ of finite measure such that*

$$\sup_{f \in T} \int_{\complement B} \|f\|^p \, \mathrm{d}\mu \leqslant \varepsilon.$$

The reader may check that these conditions are satisfied if and only if the family $\{\|f\|^p : f \in T\}$ is uniformly integrable in the sense of Definition A.3.1.

Proposition 2.6.39. *Let $(\mathscr{F}_i)_{i\in I}$ be a family of sub-σ-algebras of $\mathscr{A}$ such that μ is σ-finite on their intersection $\mathscr{F} := \bigcap_{i\in I} \mathscr{F}_i$, and let $1 \leqslant p < \infty$. Then for all $f \in L^p(S;X)$ the family $\{\mathbb{E}(f|\mathscr{F}_i) : i \in I\}$ is uniformly p-integrable.*

It will follow from the proof that the sets B in condition (ii) may be chosen from $\mathscr{F}$.

Remark 2.6.40. The assumption of σ-finiteness of $\mathscr{F}$ cannot be weakened to σ-finiteness on each of the $\mathscr{F}_i$. Indeed, let $(S, \mathscr{A}, \mu) = (\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathrm{d}x)$. Let

$$\mathscr{D}_n := \{2^{-n}([0,1) + k) : k \in \mathbb{Z}\}, \quad n \in \mathbb{Z},$$

be the standard dyadic cubes of side-length 2^{-n}, and let $\mathscr{F}_n := \sigma(\mathscr{D}_n)$ the σ-algebra generated by them. Then $(\mathscr{F}_n)_{n\in\mathbb{Z}}$ is σ-finite filtration. For $f := \mathbf{1}_{[0,1)}$ and $n \geqslant 0$, we have $f_{-n} := \mathbb{E}(f|\mathscr{F}_{-n}) = 2^{-n}\mathbf{1}_{[0,2^n)}$. Given a set $B \in \mathscr{B}(\mathbb{R})$ of finite measure, most of the mass of f_{-n} eventually gets concentrated outside the set B as $n \to \infty$, and therefore condition (ii) cannot be satisfied. Notice that $\mathscr{F} = \{\varnothing, (-\infty, 0), [0, \infty), \mathbb{R}\}$ is purely infinite in this case.

Proof of Proposition 2.6.39. First we observe that $\|\mathbb{E}(f|\mathscr{F}_i)\|^p \leqslant \mathbb{E}(\|f\|^p|\mathscr{F}_i)$ by the conditional Jensen inequality. Thus by Proposition A.3.4 it suffices to consider the case $X = \mathbb{R}$ and $p = 1$, and we need to show the uniform integrability of $\{\mathbb{E}(|f||\mathscr{F}_i) : i \in I\}$ for every $f \in L^1(S)$.

Fix $\varepsilon > 0$ arbitrary and choose $\delta > 0$ such that $\mu(A) \leqslant \delta$ implies $\int_A |f| \, \mathrm{d}\mu \leqslant \varepsilon$. Fix $i \in I$ and write $g_i := \mathbb{E}(|f||\mathscr{F}_i)$ for brevity. Set $r := \|f\|_1/\delta$ and note that $\mu(|g_i| > r) \leqslant \|g_i\|_1/r = \delta$. Noting the $\mathscr{F}_i$-measurability of $\{|g_i| > r\}$, we conclude that

$$\int_S \mathbf{1}_{\{|g|>r\}} |g_i| \, \mathrm{d}\mu = \int_S \mathbf{1}_{\{|g_i|>r\}} |f| \, \mathrm{d}\mu \leqslant \varepsilon.$$

Choose an increasing sequence of sets $B_k \in \mathscr{F}$ of finite measure that exhaust S; they exist by assumption. Then $\int_{\complement B_k} |f| \, \mathrm{d}\mu \to 0$ by dominated convergence. Given $\varepsilon > 0$, choose $B := B_k$ so large that the mentioned integral is at most ε. Then, for any $i \in I$, we have, using the conditional Jensen inequality and the defining property of the conditional expectation on $\complement B \in \mathscr{F} \subseteq \mathscr{F}_i$,

$$\int_{\complement B} |\mathbb{E}(f|\mathscr{F}_i)| \, \mathrm{d}\mu \leqslant \int_{\complement B} \mathbb{E}(|f||\mathscr{F}_n) \, \mathrm{d}\mu = \int_{\complement B} |f| \, \mathrm{d}\mu \leqslant \varepsilon.$$

$\square$

2.7 Notes

References on the general L^p-extension problem include Bukhvalov [1975], Hernandez [1984], Herz [1971], Pisier [1986a, 2010], Virot [1981].

Section 2.1

The following characterisation of "L^p-extension spaces" is due to Kwapień [1972b] and establishes a converse to Proposition 2.1.2.

Theorem 2.7.1 (Kwapień). *Let $1 \leqslant p \leqslant \infty$ be fixed. For a Banach space X the following assertions are equivalent:*

(1) *for any bounded operator $T \in \mathscr{L}(L^p(0,1))$, $T \otimes I_X$ extends to a bounded operator on $L^p(0,1;X)$;*
(2) *there exists a measure space $(S, \mathscr{A}, \mu)$ such that X is isomorphic to a closed subspace of a quotient of a space $L^p(S)$.*

A good starting point for the structure theory of L^p-spaces is the Handbook article by Alspach and Odell [2001]; classical papers on the subject include Rosenthal [1970, 1973].

Theorem 2.1.3 admits an extension to regular operators, and in this formulation a converse holds. Recall that a bounded operator $T : L^{p_1}(S_1) \to L^{p_2}(S_2)$ is called *regular* if there exists a positive operator $S : L^{p_1}(S_1) \to L^{p_2}(S_2)$ such that

$$|Tx| \leqslant S|x| \quad \forall x \in L^{p_1}(S_1). \tag{2.31}$$

For such an operator we define

$$\|T\|_{\mathrm{reg}} := \inf\big\{\|S\| : \ S \in \mathscr{L}(L^{p_1}(S_1), L^{p_2}(S_2)),\ S \geqslant 0,\ \text{and (2.31) holds}\big\}.$$

Theorem 2.7.2 (Bukhvalov). *Let $1 \leqslant p_1, p_2 < \infty$. For a bounded operator T from $L^{p_1}(S_1)$ to $L^{p_2}(S_2)$ the following assertions are equivalent:*

(1) *$T \otimes I_X$ extends to a bounded operator from $L^{p_1}(S_1;X)$ to $L^{p_2}(S_2;X)$ for every Banach space X;*
(2) *$T \otimes I_{\ell^1}$ extends to a bounded operator from $L^{p_1}(S_1;\ell^1)$ to $L^{p_2}(S_2;\ell^1)$;*
(3) *T is regular.*

In this situation, for every Banach space X we have

$$\|T \otimes I_X\| \leqslant \|T\|_{\mathrm{reg}} \leqslant \|T \otimes I_{\ell^1}\|.$$

An analogue for adjoint operators $T : L^\infty(S_1) \to L^p(S_2)$ in the style of Theorem 2.1.7 can be formulated as well.

The proof of Theorem 2.7.2 uses the fact, to be proved shortly, that T is regular if and only if there is a constant C such that for all $x_1, \dots, x_N$ in $L^{p_1}(S_1)$,

$$\Big\| \sum_{n=1}^{N} |Tx_n| \Big\|_{L^{p_2}(S_2)} \leqslant C \Big\| \sum_{n=1}^{N} |x_n| \Big\|_{L^{p_1}(S_1)}. \tag{2.32}$$

Moreover, $\|T\|_{\rm reg}$ equals the least admissible constant C.

Proof. (3)⇒(1): This follows by repeating the proof of Theorem 2.1.3 line by line; this argument produces the estimate $\|T \otimes I_X\| \leqslant \|T\|_{\rm reg}$.

(1)⇒(2): This implication is trivial.

(2)⇒(3): Let $x_1, \dots, x_N \in L^{p_1}(S_1)$ be given. Then $x := (x_n)_{n=1}^N$ defines an element of $L^{p_1}(S;\ell_1)$ and $Tx := (Tx_n)_{n=1}^N$ satisfies $Tx = (T \otimes I_{\ell^1})x$ and

$$\begin{aligned}\Big\| \sum_{n=1}^{N} |Tx_n| \Big\|_{L^{p_2}(S_2)} &= \|(T \otimes I_{\ell^1})x\|_{L^{p_2}(S_2;\ell^1)} \\ &\leqslant \|T \otimes I_{\ell^1}\| \, \|x\|_{L^{p_1}(S_1;\ell^1)} = \|T \otimes I_{\ell^1}\| \Big\| \sum_{n=1}^{N} |x_n| \Big\|_{L^{p_1}(S_1)}.\end{aligned}$$

This shows that (2.32) holds, with constant $\|T \otimes I_{\ell^1}\|$. Therefore T is regular and $\|T\|_{\rm reg} \leqslant \|T \otimes I_{\ell^1}\|$. □

Let us now prove the characterisation of regular operators alluded to in (2.32). Recall that an operator T from $L^{p_1}(S_1)$ to $L^{p_2}(S_2)$ is called *regular* if there exists a positive operator $V : L^{p_1}(S_1) \to L^{p_2}(S_2)$ such that $|Tx| \leqslant S|V|$ for all $x \in L^{p_1}(S_1)$. For such an operator we define $\|T\|_{\rm reg}$ as the infimum of the norms $\|V\|$, where V ranges over all positive operators from $L^{p_1}(S_1)$ to $L^{p_2}(S_2)$ satisfying $|Tx| \leqslant V|x|$ for all $x \in L^{p_1}(S_1)$.

Proposition 2.7.3. *Let $1 \leqslant p_1, p_2 \leqslant \infty$ and let $T : L^{p_1}(S_1) \to L^{p_2}(S_2)$ be a bounded linear operator. The following assertions are equivalent:*

(1) *T is regular;*
(2) *there exists a constant C such that for all $x_1, \dots, x_N$ in $L^{p_1}(S_1)$,*

$$\Big\| \sum_{n=1}^{N} |Tx_n| \Big\|_{L^{p_2}(S_2)} \leqslant C \Big\| \sum_{n=1}^{N} |x_n| \Big\|_{L^{p_1}(S_1)};$$

(3) *T is the difference of two positive operators.*

Moreover, $\|T\|_{\rm reg}$ is the infimum of admissible constants C in (2).

Sketch. The implications (3)⇒(1)⇒(2) are easy. The latter also shows that (2) holds with constant $C \leqslant \|T\|_{\rm reg}$; the converse inequality follows from (2) by taking $N = 1$.

The proof of (2)⇒(3) is taken from Bukhvalov [1975, Lemma 1.7] and uses the fact that in the spaces $L^p(S)$ every norm bounded set A that is directed upward (in the sense that for any two $y, y' \in A$ there exists $z \in A$

such that $z \geqslant y$ and $z \geqslant y'$) has a supremum. This property, in turn, implies that every non-empty set that is bounded above has a supremum (see, e.g., Meyer-Nieberg [1991, Section 2.4]).

Let us put $X = L^{p_1}(S_1)$ and $Y = L^{p_2}(S_1)$ for brevity. For $0 \leqslant x \in X$ consider the set

$$A_x = \Big\{ \sum_{n=1}^N |Tx_n| : \ x = \sum_{n=1}^N x_n, \ x_n \geqslant 0 \Big\}.$$

Using the Riesz decomposition property of Banach lattices (see Aliprantis and Burkinshaw [2006, Theorem 1.15]) this set is easily seen to be directed upward and, by our assumption, it is norm bounded. It follows that $\sup A_x$ exists. If $|z| \leqslant x$, then

$$|Tz| \leqslant |Tz^+| + |Tz^-| + T(x - |z|) \leqslant \sup A_x.$$

Thus we may define a mapping $T^+ : X^+ \to Y^+$ by $T^+x := \sup\{Tz : 0 \leqslant z \leqslant x\}$. Standard arguments from the theory of Banach lattices (see Aliprantis and Burkinshaw [2006, Chapter 1]) show that T^+ extends to a bounded positive linear operator from X to Y by setting, for any $x \in X$, $T^+x := T^+x^+ - T^+x^-$. Similarly, $T^-x := \sup\{-Tz : 0 \leqslant z \leqslant x\}$ defines a bounded positive operator from X to Y, and we have $T = T^+ - T^-$. □

It is known that every bounded operator $T : L^1 \to L^p$ is regular (this follows, e.g., from Meyer-Nieberg [1991, Proposition 3.3.14]). Proposition 2.1.1 then appears as a special case of Theorem 2.7.2.

The idea to use Lemma 2.1.6 to prove Theorem 2.1.7 was shown to us by N. Lindemulder. Further results along this line may be found in Lindemulder [2016].

Our proof of Theorem 2.1.9 is adapted from García-Cuerva and Rubio de Francia [1985]. In its present form it is due to Marcinkiewicz and Zygmund [1939]; the case $p_1 = p_2$ (without sharp constant) was shown earlier by Paley [1932]. Some further developments are presented in Garling [2007], which also contains a wealth of material relating to other topics treated in this chapter.

In Example 2.1.15 we have dealt with the Fourier transform on the real line. A similar example can be given for the Fourier transform on the circle. In this formulation, the first to notice the non-extendability of the Hausdorff–Young inequality to certain Banach spaces X was Bochner [1933a]. His example is interesting enough to be mentioned:

Example 2.7.4. The Fourier transform on the circle $\mathbb{T} = [0, 1)$ is defined by $\mathscr{F} f(k) := \widehat{f}(k)$, $k \in \mathbb{Z}$, with

$$\mathscr{F} f(k) := \int_0^1 f(k)e^{-2\pi ikt}\,\mathrm{d}t.$$

It is elementary to see that $\mathscr{F}$ an isometry from $L^2(\mathbb{T})$ onto $\ell^2(\mathbb{Z})$. We wish to check that $\mathscr{F} \otimes I_{L^1(\mathbb{T})}$ does not extend to a bounded operator from $L^2(\mathbb{T}; L^1(\mathbb{T}))$ to $\ell^2(\mathbb{Z}; L^1(\mathbb{T}))$.

Suppose the contrary. Fix an arbitrary $f \in L^1(\mathbb{T})$ and define $g_f \in L^2(\mathbb{T}; L^1(\mathbb{T}))$ by $(g_f(t))(s) := f(t + s \bmod 1)$. By approximation one sees that its image under $\mathscr{F} \otimes I_{L^1(\mathbb{T})}$ must be the the sequence $(\widehat{g_f}(k))_{k\in\mathbb{Z}}$, where $\widehat{g_f}(k) \in L^1(\mathbb{T})$ is the function $s \mapsto e^{2\pi iks}\widehat{f}(k)$, so that

$$\|(\mathscr{F} \otimes I_{L^1(\mathbb{T})})g_f\|^2_{\ell^2(\mathbb{Z};L^1(\mathbb{T}))} = \sum_{k\in\mathbb{Z}} \|e^{2\pi ik(\cdot)}\widehat{f}(k)\|^2_{L^1(\mathbb{T})} = \sum_{k\in\mathbb{Z}} |\widehat{f}(k)|^2.$$

By assumed boundedness of $\mathscr{F} \otimes I_{L^1(\mathbb{T})}$, it follows that

$$\|f\|_2^2 = \sum_{k\in\mathbb{Z}} |\widehat{f}(k)|^2 \leqslant \|\mathscr{F} \otimes I_{L^1(\mathbb{T})}\|^2 \|g_f\|^2_{L^2(\mathbb{T};L^1(\mathbb{T}))} = \|\mathscr{F} \otimes I_{L^1(\mathbb{T})}\|^2 \|f\|_1^2.$$

an obvious contradiction.

Incidentally, this example also shows that the partial Fourier sums $S_N g$ of a function in $g \in L^2(\mathbb{T}; X)$ need not converge to g in the norm of $L^2(\mathbb{T}; X)$, an observation from Bochner and Taylor [1938]. Indeed, with the above notation we have $S_N g_f(k) = \sum_{|n|\leqslant N} e^{2\pi ikt} e^{ik(\cdot)}\widehat{f}(k)$. Now convergence $S_N g_f \to g_f$ in $L^2(\mathbb{T}; L^1(\mathbb{T}))$ would lead to the contradiction

$$\begin{aligned}
\|f\|^2_{L^2(\mathbb{T})} &= \lim_{N\to\infty} \int_0^1 \Big| \sum_{|k|\leqslant N} e^{2\pi ikt}\widehat{f}(k)\Big|^2 \,\mathrm{d}t \\
&\leqslant \lim_{N\to\infty} \int_0^1 \Big\|S_N g_f(t)\Big\|^2_{L^1(\mathbb{T})} \,\mathrm{d}t = \|g_f\|^2_{L^2(\mathbb{T};L^1(\mathbb{T}))} \leqslant \|f\|^2_{L^1(\mathbb{T})}.
\end{aligned}$$

Example 2.1.15 is from Peetre [1969]. Example 2.1.16 is a variation of an example in Defant [1989]. Example 2.1.17 is due to Rosiński and Suchanecki [1980]. Previous examples of bounded functions which fail to admit a stochastic integral were given in Yor [1974]. Theorem 2.1.20 is due to Hoffmann-Jørgensen and Pisier [1976].

Weighted norm inequalities

In harmonic analysis, weighted norm inequalities in the scalar case often imply the boundedness of vector-valued extensions. This principle was first observed by Rubio de Francia [1984]. This rich and deep subject is covered in the monographs García-Cuerva and Rubio de Francia [1985] and Cruz-Uribe, Martell, and Pérez [2011]. A typical example is as follows. Let $1 < p_0 < \infty$ and suppose T is a bounded operator on $L^{p_0}(\mathbb{R}^d, w)$ for all Muckenhoupt A_{p_0}-weights w, with a bound on the norms which depends on the A_{p_0}-constant of w in a suitable 'monotone' way. Then T extends to a bounded operator on $L^p(\mathbb{R}^d, w; \ell^q)$

for all $1 < p, q < \infty$ and all A_p-weights w. This result remains true for sub-linear operators, and in this formulation it covers the Hardy–Littlewood maximal operator. Instead of ℓ^q it is possible to use more general classes of Banach function spaces; see Rubio de Francia [1986]. Recall that a *Banach function space* over a measure space $(S, \mathscr{A}, \mu)$ is a subspace $X \subseteq L^0(S)$ equipped with a complete norm that respects the pointwise ordering, in the sense that if $|f| \leqslant |g|$ for some $f \in L^0(S)$ and $g \in X$, then $f \in X$ with $\|f\|_X \leqslant \|g\|_X$.

Section 2.2

The chronologically first result on the interpolation of operators in the modern sense is the Riesz–Thorin Theorem 2.2.1. It is originally due to Riesz [1927], whose proof is quite different from the present one; it is completely "elementary" (although perhaps not easy), as it has nothing to do with complex variables but relies at the bottom on a careful analysis of the conditions for equality in Hölder's inequality. In the introduction to his article, Riesz [1927] indicates the work of Jensen [1906] on convex function inequalities as his source of inspiration. It was Thorin [1938] who introduced the now-standard complex-variable framework for this result, foreshadowing the abstract notion of complex interpolation. Our proof of Theorem 2.2.1 is standard and may be found in most textbooks on interpolation theory. We mention Bergh and Löfström [1976], Kreĭn, Petunīn, and Semënov [1982], Lunardi [2009], Triebel [1978].

The Marcinkiewicz Interpolation Theorem 2.2.3 was announced without proof in a two-page note by Marcinkiewicz [1939a]. The full theorem, already formulated in Marcinkiewicz [1939a], also deals with operators T mapping L^p into another L^q. such a mapping is called *quasi-linear* if, for all $f \in L^p$ and $c \in \mathbb{K}$,

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

and there is a constant $C \geqslant 0$ such that for all $f, g \in L^p$,

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

Theorem 2.7.5 (Marcinkiewicz [1939a]). *Let $T : L^{p_0} + L^{p_1} \to L^0$ be a quasi-linear operator that maps $T : L^{p_i} \to L^{q_i,\infty}$ for $i = 0, 1$, where $p_i \leqslant q_i$ and $p_0 < p_0$. Then*

$$T : L^p \to L^q \quad \textit{for} \quad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \ \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q}, \quad \theta \in (0,1).$$

According to Zygmund [1956, p. 228], the proof of this theorem was indicated in Marcinkiewicz's letter to Zygmund in June 1939, and finally published by Zygmund [1956]. The argument found there is already the one presented in most modern textbooks; we have followed the same lines, except for some fine-tuning of the constants which may be new. Our proof also gives Theorem

2.7.5, but a variation of Lemma 2.2.5 is needed which involves a Minkowski type argument.

The theorem of Marcinkiewicz [1939a] apparently attracted little attention until it was revisited by Zygmund [1956]. It seems that Zygmund himself had forgotten about the result in the first place, as he later adds that his fundamental work on singular integrals, Calderón and Zygmund [1952], "implicitly uses the same ideas", whereas "an explicit application of [Marcinkiewicz's interpolation theorem] would have shortened the proof" (Zygmund [1956, p. 235]). The key novelty of Marcinkiewicz theorem, allowing applications beyond the scope of the Riesz–Thorin theorem, was the use of weak-type norms. In the preface to Marcinkiewicz [1964], Zygmund writes that this was "suggested in all probability by an old result of Kolmogorov [1925] about properties of conjugate functions." An extension of Marcinkiewicz's result to Lorentz spaces is due to Hunt [1964].

Comprehensive treatments of vector-valued function spaces and their interpolation theory are contained in Amann [1995], König [1986], Kreĭn, Petunīn, and Semënov [1982], Schmeisser [1987]. Lemma 2.2.7 is due to Calderón [1964] with a different proof. Our presentations of Theorems 2.2.6 and 2.2.10 follow Triebel [1978, Section 1.18.4] and Kreĭn, Petunīn, and Semënov [1982, Theorem IV.2.11], respectively. The Clarkson inequalities (Corollary 2.2.9) are from Clarkson [1936]. The proof presented here was kindly shown to us by J. Voigt. The idea of applying complex interpolation to the mapping $(f, g) \mapsto (f+g, f-g)$ appears in Kato and Takahashi [1997], where the validity of the Clarkson inequalities in $L^p(S;X)$ is related to the (Rademacher) type and cotype constants of X. Further results along this line are given in Milman [1984], where Clarkson inequalities in $L^p(S;X)$ are related to the Fourier type of X with respect to the group $\{0,1\}$ with addition modulo 2.

Section 2.3

Hardy and Littlewood introduced the maximal function that now bears their name in their famous 1930 paper Hardy and Littlewood [1930]. They proved its L^p-boundedness as a consequence of an elaborate argument using decreasing rearrangements, which were introduced for that purpose in the same paper. The latter's usefulness is explained as follows:

> "The problem is most easily grasped when stated in the language of cricket, or any other game in which a player compiles a series of scores of which an average is recorded."

They then consider the batman's satisfaction as an increasing function of his average over the number of inning played to date. This leads them to the notion of decreasing rearrangements, for it is easy to see that

> "(The batmans's) total satisfaction for the season (...) is a maximum, for a given stock of innings, when the innings are played in decreasing order."

Instead of the batsman's average for the complete season to date, they proceed to consider his maximum average for any consecutive series of innings to date, for which bounds are established that ultimately lead to the L^p-boundedness of the maximal function.

The limitation of this argument is that it works in dimension $d = 1$ only. The extension to higher dimensions through the use of a covering lemma is due to Wiener [1939]. The version with weights presented in Theorem 2.3.2, which works for functions with values in an arbitrary Banach space X, should be carefully distinguished from the Hardy–Littlewood theorem for A_p-weights hinted at in the Notes to Section 2.1. The latter states that if $1 < p < \infty$ and w is a Muckenhoupt A_p-weight on $\mathbb{R}^d$, then the Hardy–Littlewood function M is bounded on $L^p(\mathbb{R}^d; w)$ and, more generally, on $L^p(\mathbb{R}^d, w; \ell^q)$ for all $1 < q < \infty$.

The literature contains many versions of Theorem 2.3.8, making it difficult to trace its exact origins. The reader is referred to Stein and Weiss [1971] for a fuller discussion. It was shown in Burkholder [1964] that, in many situations, maximal inequalities are not only sufficient, but also necessary to obtain the almost everywhere convergence of a given sequence of functions.

Section 2.4

With the exception of the material on Fourier type, the results of this section are standard. Vector-valued spaces of test functions and the corresponding classes of (tempered) distributions are discussed in Amann [1995].

The notion of Fourier type goes back to Peetre [1969], where most of the results presented here can be found. Peetre was concerned with improvement of the classical inclusion relations

$$(X_0, X_1)_{\theta,1} \hookrightarrow [X_0, X_1]_\theta \hookrightarrow (X_0, X_1)_{\theta,\infty}$$

between the real and complex interpolation spaces in the situation where X_0 and X_0 have Fourier type p_0 and p_1, respectively. He was able to prove that under this assumption one has the embeddings

$$(X_0, X_1)_{\theta,p_\theta} \hookrightarrow [X_0, X_1]_\theta \hookrightarrow (X_0, X_1)_{\theta,p'_\theta},$$

where $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$. Peetre's proof is reproduced in Appendix C.

The equivalence of Fourier type for the line and the circle of Proposition 2.4.20 is due to Kwapień [1972a] whose approach we follow. A detailed proof that the function considered in (2.17) takes its minimum in the points $\frac{1}{2} + \mathbb{Z}$ is written out in Gijswijt and Van Neerven [2016]. The proof is long and tedious, and ultimately relies on a variant of the classical Hardy–Littlewood-Pólya inequality.

The notion of Fourier type is connected with the probabilistic notions of type and cotype (see Definition 4.3.12; a detailed study will follow in Volume II). It is straightforward to see that every Banach space with Fourier type p has type p and cotype q. It is a deep theorem of Bourgain [1988] that the following converse holds:

Theorem 2.7.6 (Bourgain). *If X has non-trivial type, then X has non-trivial Fourier type.*

Here, 'non-trivial' means 'belonging to $(1, 2]$'. It is not true, however, that non-trivial type p implies non-trivial Fourier type p; Bourgain's proof gives a non-linear dependence of the Fourier type on the type and its constant. Another proof of Bourgain's theorem can be found in Hinrichs [1996].

In König [1991], various summability properties of the Fourier coefficients of X-valued functions were established under Fourier type assumptions on X. For instance, it was shown that X has non-trivial Fourier type if and only if the Fourier coefficients of every $f \in C^1(\mathbb{T}; X)$ are absolutely summable.

Expository references for this material are García-Cuerva, Kazaryan, Kolyada, and Torrea [1998] and Pietsch and Wenzel [1998]. Bourgain's result has various application to vector-valued harmonic analysis; in the presence of non-trivial Fourier type it is often possible to improve certain parameters, much in the way as we saw above in connection with interpolation.

Section 2.5

A classical reference on Sobolev spaces is the book by Adams, which appeared in its second edition as Adams and Fournier [2003]. Gentle introductions may be found in Brezis [2011] and Evans [2010]. A thorough treatment of fractional Sobolev spaces which includes various other related classes of function spaces is presented in Runst and Sickel [1996] and Triebel [1978]. A self-contained introduction is presented in Di Nezza, Palatucci, and Valdinoci [2012]. Sobolev spaces of vector-valued functions are treated mostly in connection with the theory of evolution equations; see for instance Amann [1995], Prüss and Simonett [2016]. A detailed discussion of the theory of vector-valued Sobolev spaces can be found in Schmeisser [1987], Triebel [1997], and references therein.

The results connecting weak derivatives and almost everywhere derivatives are mostly classical, although it is hard to find a systematic reference in the literature. For more on the connection with the Radon–Nikodým property we refer to Diestel and Uhl [1977]. The proofs of Lemma 2.5.10 and 2.5.11 are taken from Stein and Shakarchi [2005]. We learnt the trick in the proof of (3)⇒(2) of Theorem 2.5.12 from Arendt, Batty, Hieber, and Neubrander [2011] whose presentation we follow. A direct real variable proof of Lemma 2.5.13 that avoids the use of the Carathéodory theorem and the Radon–Nikodým theorem can be found in many textbooks. A detailed presentation is given in Stein and Shakarchi [2005].

The proof of Theorem 2.5.17 is taken from Lunardi [2009]. The result can be generalised to open domains D in $\mathbb{R}^d$ that admit suitable extension operators, such as domains with a uniformly C^1 boundary. For such domains one then obtains

$$(L^p(D; X), W^{1,p}(D; X))_{\theta,p} = W^{\theta,p}(D; X)$$

with equivalent norms.

Section 2.6

Conditional expectations in a σ-finite setting have been considered by many authors. Some aspects of our approach seem to be new, however, even in the scalar-valued situation. Our approach is related to that of Stromberg [1994] and Tanaka and Terasawa [2013], the principal difference being that we take the collection $\mathscr{F}_f$ of sets over which f is integrable as the point of departure rather than the collection of sets of finite μ-measure, our philosophy being that what really matters is the integrability properties of f on a gives set, and not so much the size of the set.

Neveu [1975] takes a different approach which we will outline and compare with ours. Let first $(S, \mathscr{A}, \mathbb{P})$ be a probability space and let $\mathscr{F} \subseteq \mathscr{A}$ be a sub-σ-algebra. For arbitrary $\mathscr{A}$-measurable $f : S \to [0, \infty]$, Neveu checks that $\mathbb{E}(f|\mathscr{F})$ may be defined as the almost everywhere limit of $\mathbb{E}(f \wedge n|\mathscr{F})$ as $n \to \infty$, in the sense that this limit is the unique (up to a $\mathbb{P}$-null set) $\mathscr{F}$-measurable function with values in $[0, \infty]$ such that

$$\int_F \mathbb{E}(f|\mathscr{F})\,\mathrm{d}\mathbb{P} = \int_F f\,\mathrm{d}\mathbb{P} \quad \forall F \in \mathscr{F}. \tag{2.33}$$

As a second step, let $(S, \mathscr{A}, \mu)$ be a σ-finite measure space and let $\mathscr{F} \subseteq \mathscr{A}$ be a sub-σ-algebra. We do not assume μ to be σ-finite on $\mathscr{F}$. Fix a strictly positive integrable function $w : S \to (0, \infty)$. Then the measure $\mathbb{P} := w\mu$ is a probability measure and we have $\mu = w^{-1}\mathbb{P}$ (where of course $w^{-1} = 1/w$). On $\mathscr{F}$, by (2.33) we have $\mu = \mathbb{E}(w^{-1}|\mathscr{F})\mathbb{P}$, the conditional expectation being taken with respect to $\mathbb{P}$. The set

$$F_\mu := \{\mathbb{E}(w^{-1}|\mathscr{F}) < \infty\}$$

belongs to $\mathscr{F}$ and has the following properties:

(i) $\mu|_{\mathscr{F}}$ is σ-finite on F_μ;
(ii) $\mu|_{\mathscr{F}}$ is purely infinite on $\complement F_\mu$.

Recall that the second property means that $\mu|_{\mathscr{F}}$ only takes the values 0 and ∞ on the $\mathscr{F}$-measurable subsets of $\complement F_\mu$. Then Neveu observes that $\mathbb{E}(w^{-1}f|\mathscr{F})$ is strictly positive μ-almost everywhere on F_μ and that, for an $\mathscr{A}$-measurable function $f : S \to [0, \infty]$, the prescription

$$\mathbb{E}_\mu(g|\mathscr{F}) := \begin{cases} \dfrac{\mathbb{E}(w^{-1}f|\mathscr{F})}{\mathbb{E}(w^{-1}|\mathscr{F})} & \text{on } F_\mu \\ \\ 0 & \text{on } \complement F_\mu \end{cases} \tag{2.34}$$

defines the conditional expectation of g with respect to $\mathscr{F}$, taken with respect to μ, in the sense that it agrees with the orthogonal projection in $L^2(S, \mathscr{F}, \mu)$

onto $L^2(S,\mu)$ and is the unique (up to a $\mathbb{P}$-null set) $\mathscr{F}$-measurable function with values in $[0,\infty]$ such that

$$\int_F \mathbb{E}_\mu(g|\mathscr{F})\,\mathrm{d}\mu = \int_F g\,\mathrm{d}\mu \quad \forall F\in\mathscr{F},\ F\subseteq F_\mu. \tag{2.35}$$

In particular, the definition of $\mathbb{E}_\mu(g|\mathscr{F})$ is independent of the choice of w. This conditional expectation defines a positive linear contraction from $L^p(S,\mathscr{F},\mu)$ onto $L^p(S,\mu)$ for each $p\in[1,\infty]$.

To compare this approach with ours, we first observe that if $(S,\mathscr{A},\mu)$ is a σ-finite measure space and $\mathscr{F}\subseteq\mathscr{A}$ is a sub-σ-algebra, then by Proposition A.1.4 there exist disjoint sets $S_0,S_1\in\mathscr{F}$ such that $S_0\cup S_1=S$ such that the restriction of μ to $\mathscr{F}|_{S_0}$ is σ-finite and the restriction of μ to $\mathscr{F}|_{S_1}$ is purely infinite. To see that Neveu's approach coincides with ours, we simply note that $S_0=F_\mu$ and $S_1=\complement F_\mu$ up to μ-null sets, for there can be only one such partition of S up to μ-null sets. Now it follows from (2.35) and uniqueness that on F_μ the conditional expectation given by (2.34) agrees with ours.

The extension of this approach to the vector-valued case would create some difficulties, since the step which extends the conditional expectation on $L^2(S,\mathbb{P})$ to $L^0(S,\mathbb{P})$ is based on a monotonicity argument. This can easily be circumvented by using the fact (see Proposition 2.6.2) that any function in $L^0(S;X)$ is σ-integrable and then define the conditional expectation along the lines of (2.28), but this would bring us straight to the construction we have taken. Let us also emphasise that our approach develops conditional expectations from first principles.

Yet another approach could be based on extending the regular conditional probability (see Kallenberg [2002, Chapter 6]) and to use it to define the vector-valued conditional expectation. Using this method, "conditional versions" of several classical theorems can be derived immediately.

Subsections 2.6.c and 2.6.d follow a more or less standard pattern. An alternative proof of Proposition 2.6.29 (for functions $f\in L^p(S;X)$) can be found in Haase [2007].

Conditional expectations in $L^p(S;L^q(T;X))$

Let $(S,\mathscr{A},\mu)$ and $(T,\mathscr{B},\nu)$ be measure spaces. For a sub-σ-algebra $\mathscr{F}$ of the product σ-algebra $\mathscr{A}\times\mathscr{B}$, we define $L^p_{\mathscr{F}}(S;L^q(T;X))$ to be the closed subspace in $L^p(S;L^q(T;X))$ of all functions having a strongly $\mathscr{F}$-measurable representative. The following result is proved in Lü and Van Neerven [2016].

Theorem 2.7.7. *Let $(A,\mathscr{A},\mu)$ and $(B,\mathscr{B},\nu)$ be probability spaces, let X be a Banach space, and let $1<p,q<\infty$. Then the conditional expectation operator $\mathbb{E}(\cdot|\mathscr{F})$ on $L^1(S\times T)$ restricts to a bounded projection on the space $L^p(S;L^q(T;X))$ with range $L^p_{\mathscr{F}}(S;L^q(T;X))$. The norm of this projection is bounded by a constant independent of X.*

The next example, due to Qiu [2012], shows that the conditional expectation may fail to be contractive. It will be refined in Chapter 4 (see Proposition 4.3.14).

Example 2.7.8. Let $A = B = \{0,1\}$ with $\mathscr{A} = \mathscr{B} = \{\varnothing, \{0\}, \{1\}, \{0,1\}\}$ and $\mu = \nu$ the measure on $\{0,1\}$ that gives each point mass $\frac{1}{2}$, and let

$$\mathscr{F} = \{\{(0,0),(1,0)\}, \{(0,1)\}, \{(1,1)\}\}.$$

Let $f : A \times B \to \mathbb{R}$ be defined by $f(0,0) = 0$, $f(1,0) = 1$, $f(0,1) = 1$, $f(1,1) = 0$. Then

$$\mathbb{E}(f|\mathscr{F})(0,0) = \frac{1}{2}, \ \ \mathbb{E}(f|\mathscr{F})(1,0) = \frac{1}{2}, \ \ \mathbb{E}(f|\mathscr{F})(0,1) = 1, \ \ \mathbb{E}(f|\mathscr{F})(1,1) = 0.$$

Hence in this example we have

$$\begin{aligned}\|f\|_{L^p(S;L^2(T))} &= \Big(\big(\frac{1}{2}\big)^{p/2} + \big(\frac{1}{2}\big)^{p/2}\Big)^{1/p},\\ \|\mathbb{E}(f|\mathscr{F})\|_{L^\infty(S;L^2(T))} &= \Big(\big(\frac{1}{8}\big)^{p/2} + \big(\frac{5}{8}\big)^{p/2}\Big)^{1/p}.\end{aligned}$$

Consequently, for large enough p the conditional expectation fails to be contractive in $L^p(S;L^2(T))$.

As an immediate consequence of Proposition 2.7.7 we obtain the following duality result, due to Lü, Yong, and Zhang [2012]. Their proof, however, requires an additional assumption on $\mathscr{F}$.

Corollary 2.7.9. *Let $(A, \mathscr{A}, \mu)$ and $(B, \mathscr{B}, \nu)$ be a probability space and a σ-finite measure space respectively, and let X be a Banach space whose dual has the Radon-Nikodým property, and let $1 < p, q < \infty$. Then we have a natural isometric isomorphism of Banach spaces*

$$(L^p_{\mathscr{F}}(S;L^q(T;X)))^* = L^{p'}_{\mathscr{F}}(S;L^{q'}(T;X)).$$

This corollary cannot be extended to $p = 1$, even when $q = 2$. In fact it is shown in Lü and Van Neerven [2016] that if $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space supporting a Brownian motion $(B_t)_{t\in[0,1]}$, and if $\mathscr{P}$ is the so-called progressive σ-algebra associated with it (see, e.g., Chung and Williams [1990] for its definition), then

$$L^\infty_{\mathscr{P}}(\Omega;L^2(0,1)) \subsetneq (L^1_{\mathscr{P}}(\Omega;L^2(0,1)))^*. \tag{2.36}$$

3

Martingales

This chapter on martingale theory provides the techniques which will serve as the main tool in the next two chapters. With a view towards averaging and approximation schemes in harmonic analysis, we continue in the framework adopted in the section on conditional expectations in the previous chapter, where we have set up the theory for σ-finite measure spaces. Our treatment covers both forward and backward martingales.

After introducing the main definitions in Section 3.1, maximal inequalities form the main topic of Section 3.2. We present a general form of Doob's maximal inequality for non-negative martingales $(f_n)_{n\in\mathbb{Z}}$ with a non-negative weight w. This leads quickly to the Hardy–Littlewood and Fefferman–Stein maximal inequalities.

Section 3.3 contains the martingale convergence theorems for forward and backward martingales. The treatment of L^1-bounded martingales requires the Radon–Nikodým property introduced in Chapter 1.

The decomposition theorems of Davis and Gundy in Section 3.4 pave the way for good-λ inequalities for martingale transforms in Section 3.5. These lead to the highlight of this chapter, the extrapolation theorem for operator-valued martingale transforms. The endpoint case also requires a martingale version of Cotlar's inequality for singular integrals.

As an illustration we use operator-valued martingale transforms to introduce equivalent norms in the martingale type and cotype inequalities. More importantly, we will show in future chapters that estimates for the Hilbert transform and more general singular integral operators and Fourier multipliers can be reduced to the rather well-structured and well-behaved martingale transforms.

In Section 3.6 we prove some approximation results for martingales by Paley–Walsh martingales. Such martingales have a much simpler structure, and in later sections we will apply this approximation procedure to reduce the proofs of many statements to this simpler class of martingales.

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 63, DOI 10.1007/978-3-319-48520-1_3

Throughout the chapter we comment on the quantitative behaviour of the constants in the various martingale inequalities. In some cases we are even able to determine the exact best constants.

3.1 Definitions and basic properties

Let $(S, \mathscr{A}, \mu)$ be a measure space, $(I, \leqslant)$ an ordered set, and X a Banach space. We recall from Theorem 2.6.20 that if μ is σ-finite on the sub-σ-algebra $\mathscr{F}$ of $\mathscr{A}$ and $f \in L^0(S; X)$ is σ-integrable over $\mathscr{F}$, then the conditional expectation $\mathbb{E}(f|\mathscr{F})$ exists and is uniquely determined as an element of $L^0(S; X)$.

Definition 3.1.1 (Filtrations, adaptedness, martingales).

(i) *A family of sub-σ-algebras $(\mathscr{F}_n)_{n\in I}$ of $\mathscr{A}$ is called a* filtration *in $(S, \mathscr{A}, \mu)$ if $\mathscr{F}_m \subseteq \mathscr{F}_n$ whenever $m, n \in I$ and $m \leqslant n$. The filtration is called σ-*finite *if μ is σ-finite on each $\mathscr{F}_n$.*

(ii) *A family of functions $(f_n)_{n\in I}$ in $L^0(S; X)$ is* adapted *to the filtration $(\mathscr{F}_n)_{n\in I}$ if $f_n \in L^0(S, \mathscr{F}_n; X)$ for all $n \in I$.*

(iii) *A family of functions $(f_n)_{n\in I}$ in $L^0(S; X)$ is called a* martingale *with respect to a σ-finite filtration $(\mathscr{F}_n)_{n\in I}$ if it is adapted to $(\mathscr{F}_n)_{n\in I}$, and for all indices $m \leqslant n$ the function f_n is σ-integrable over $\mathscr{F}_m$ and satisfies*

$$\mathbb{E}(f_n|\mathscr{F}_m) = f_m.$$

If $(f_n)_{n\in I}$ is a martingale and each f_n belongs to $L^p(S; X)$ (with $1 \leqslant p \leqslant \infty$), then $(f_n)_{n\in I}$ is called an *L^p-martingale.* Such a martingales is called if $\sup_{n\in I} \|f_n\|_p < \infty$. If $(f_n)_{n\in I}$ is a martingale and each f_n is μ-simple, then $(f_n)_{n\in I}$ is called a μ-simple martingale.

An adapted family of real-valued functions $(f_n)_{n\in I}$ in $L^0(S)$ is called a *submartingale* with respect to a σ-finite filtration $(\mathscr{F}_n)_{n\in I}$ if for all indices $m \leqslant n$, f_n is σ-integrable over $\mathscr{F}_m$ and satisfies

$$f_m \leqslant \mathbb{E}(f_n|\mathscr{F}_m),$$

and a *supermartingale* with respect to this filtration if for all indices $m \leqslant n$, f_n is σ-integrable over $\mathscr{F}_m$ and satisfies

$$f_m \geqslant \mathbb{E}(f_n|\mathscr{F}_m).$$

The notions of *L^p-submartingale* and *L^p-supermartingale* are defined similarly. Clearly, a sequence of real-valued functions is a martingale if and only if it is both a submartingale and a supermartingale.

Example 3.1.2 (Martingales generated by a function). Let $(\mathscr{F}_n)_{n\in I}$ be a σ-finite filtration in $(S, \mathscr{A}, \mu)$. If a function $f \in L^0(S; X)$ is σ-integrable over each $\mathscr{F}_n$, then

$$f_n := \mathbb{E}(f|\mathscr{F}_n), \quad n \in I,$$

defines a martingale $(f_n)_{n\in I}$ with respect to $(\mathscr{F}_n)_{n\in I}$. Indeed, Proposition 2.6.33 shows that for all $m \leqslant n$, we have

$$\mathbb{E}(f_n|\mathscr{F}_m) = \mathbb{E}(\mathbb{E}(f|\mathscr{F}_n)|\mathscr{F}_m) = \mathbb{E}(f|\mathscr{F}_m) = f_m.$$

Moreover, if $f \in L^p(S; X)$, then by the contractivity of the conditional expectation (Corollary 2.6.30) $(f_n)_{n\in I}$ is an L^p-bounded martingale.

Example 3.1.3 (Dyadic harmonic analysis). Let $(S, \mathscr{A}, \mu) = (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d), \mathrm{d}x)$. Let

$$\mathscr{D}_n := \{2^{-n}([0,1)^d + k) : k \in \mathbb{Z}^d\}, \quad n \in \mathbb{Z},$$

be the standard dyadic cubes of side-length 2^{-n}, and $\mathscr{F}_n := \sigma(\mathscr{D}_n)$ its generated σ-algebra. Then $(\mathscr{F}_n)_{n\in\mathbb{Z}}$ is a σ-finite filtration. Every function $f \in L^1_{\rm loc}(\mathbb{R}^d; X)$ is σ-integrable over every $\mathscr{F}_n$ and therefore generates a martingale $(\mathbb{E}(f|\mathscr{F}_n))_{n\in\mathbb{Z}}$. These *dyadic martingales* are fundamental to harmonic analysis and will be explored in more detail in Chapter 5.

Example 3.1.4 (Sums of independent mean zero random variables). The partial sum sequence of any sequence $(\xi_n)_{n\geqslant 0}$ of independent mean zero random variables in $L^1(\Omega; X)$ is a martingale with respect to the filtration generated by $(\xi_n)_{n\geqslant 0}$ given by $\mathscr{F}_n = \sigma(\xi_0, \dots, \xi_n)$. Indeed, by Proposition 2.6.35, for all $m < n$ the conditional expectations satisfy $\mathbb{E}(\xi_n|\mathscr{F}_m) = \mathbb{E}(\xi_n) = 0$.

Example 3.1.5 (Submartingales arising from a martingale). If $(f_n)_{n\in I}$ is a martingale with respect to a σ-finite filtration $(\mathscr{F}_n)_{n\in I}$, then $(\|f_n\|)_{n\in I}$ is a submartingale with respect to $(\mathscr{F}_n)_{n\in I}$. If $(f_n)_{n\in I}$ is an L^p-martingale with respect to $(\mathscr{F}_n)_{n\in I}$ for some $1 \leqslant p < \infty$, and if $1 \leqslant q \leqslant p$, then $(\|f_n\|^q)_{n\in I}$ is a submartingale with respect to $(\mathscr{F}_n)_{n\in I}$.

To check this, first note that $\|f_n\|^q$ is integrable over any set of finite measure if $f_n \in L^p(S; X)$ and $q \leqslant p$. Since $\mathscr{F}_m$ is σ-finite, the σ-integrability of $\|f_n\|^q$ over $\mathscr{F}_m$ follows. In each case, the submartingale property follows from the conditional Jensen inequality applied to $\phi(x) = \|x\|$ or $\phi(x) = \|x\|^q$.

3.1.a Difference sequences

In Example 3.1.4 we saw that the partial sums of a sequence of independent random variables form a martingale. From this martingale, the random variables can be recovered as the corresponding martingale differences. This simple observation suggests the following basic definition.

Definition 3.1.6 (Difference sequence).

(i) *The* difference sequence *of a sequence* $(f_n)_{n\in\mathbb{Z}}$ *is the sequence* $(df_n)_{n\in\mathbb{Z}}$ *given by*

$$df_n = f_n - f_{n-1}, \quad n \in \mathbb{Z}.$$

(ii) *A sequence of functions* $(d_n)_{n\in\mathbb{Z}}$ *is a* martingale difference sequence *with respect to a* σ*-finite filtration* $(\mathscr{F}_n)_{n\in\mathbb{Z}}$ *if it is adapted, each* d_n *is* σ*-integrable with respect to* $\mathscr{F}_{n-1}$*, and*

$$\mathbb{E}(d_n|\mathscr{F}_{n-1}) = 0 \qquad \forall n \in \mathbb{Z}.$$

Clearly, a sequence $(f_n)_{n\in\mathbb{Z}}$ in $L^0(S;X)$ is a martingale if and only if its difference sequence $(df_n)_{n\in\mathbb{Z}}$ is a martingale difference sequence.

Martingales $(f_n)_{n\in I}$ over the index sets $I = \mathbb{N}$ and $I = \{0,\dots,N\}$ can be embedded in martingales over $\mathbb{Z}$ by extending them trivially with $f_n = f_0$ for $n \leqslant -1$ and, in the second case, with $f_n = f_N$ for $n \geqslant N+1$. Their difference sequences are then defined as the difference sequences of the extended martingales. Of course, $df_n = 0$ for $n \leqslant 0$ and, in the second case $df_n = 0$ for $n \geqslant N+1$; usually these differences are ignored.

Example 3.1.7 (Martingale transforms). Let $I = \mathbb{N}$ or $I = \{0,\dots,N\}$. If $f = (f_n)_{n\in I}$ is a martingale and $v = (v_n)_{n\in I}$ is a *predictable sequence*, i.e., v_0 is $\mathscr{F}_0$-measurable and v_n is $\mathscr{F}_{n-1}$-measurable for all $n \in I$ with $n \geqslant 1$, the sequence

$$g_n := v_0 f_0 + \sum_{k=1}^{n} v_k df_k, \quad n \in I,$$

defines a martingale, the so-called *martingale transform* of f by v. The martingale property for g is an immediate consequence of Proposition 2.6.31.

A similar definition can be given for martingales over $\mathbb{Z}$, but some delicate convergence issues have to be handled in this case; this is taken up in Definition 3.5.1. We shall see in later chapters that a great many questions in Analysis can be reduced to problems about martingale transforms. We shall take up their study in Section 3.5.

3.1.b Paley–Walsh martingales

A special class of martingales, the class of so-called Paley–Walsh martingales, plays a distinguished role in the theory: many assertions concerning martingales hold for arbitrary martingales once they hold for Paley–Walsh martingales.

Let $(S,\mathscr{A},\mu)$ be a measure space. Recall that a set $A \in \mathscr{A}$ is an *atom* if $\mu(A) > 0$ and $A = A_0 \cup A_1$ with disjoint $A_0, A_1 \in \mathscr{A}$ implies that $\mu(A_0) = 0$ or $\mu(A_1) = 0$.

Definition 3.1.8 (Paley–Walsh martingales). *A filtration* $(\mathscr{F}_n)_{n\geqslant 0}$ *on a probability space* $(\Omega,\mathscr{F},\mathbb{P})$ *is called a* Paley–Walsh filtration *if each set belonging to* $\mathscr{F}_n$ *is a union of atoms of size* 2^{-n}*.*

A Paley–Walsh martingale *is a martingale with respect to a Paley–Walsh filtration.*

It follows immediately from the definition that $\mathscr{F}_n$ has precisely 2^n atoms and that each atom $F \in \mathscr{F}_n$ is the union of two atoms in $\mathscr{F}_{n+1}$, which we may label in an arbitrary fashion as F^+ and F^-. Thus, $\mathscr{F}_0$ has Ω as its unique atom of size 1, $\mathscr{F}_1$ has two atoms Ω^+ and Ω^- of size $\frac{1}{2}$, $\mathscr{F}_2$ has four atoms Ω^{++}, Ω^{+-}, Ω^{-+}, Ω^{--} of size $\frac{1}{4}$, and so on.

There is a useful representation formula for Paley–Walsh martingales in terms of so-called Rademacher variables. A *real Rademacher variable* (*complex* Rademacher variables will be defined later on) is a random variable $r : \Omega \to \{-1, 1\}$ satisfying $\mathbb{P}(r = -1) = \mathbb{P}(r = 1) = \frac{1}{2}$. A *real Rademacher sequence* is a sequence of independent Rademacher variables.

Example 3.1.9. Two standard constructions of real Rademacher variables are as follows.

(a) On the product space $\prod_{n\geqslant 1}(\{-1, 1\}, \mu)$ with $\mu(\{-1\}) = \mu(\{1\}) = \frac{1}{2}$, the coordinate functions

$$r_n(\omega) := \omega_n, \quad n \geqslant 1,$$

define a Rademacher sequence.

(b) On the unit interval $[0, 1)$, the functions

$$r_n(t) := \operatorname{sgn}(\sin(2^n \pi t)), \quad n \geqslant 1,$$

define a Rademacher sequence.

Now we are ready for:

Proposition 3.1.10 (Representation of Paley–Walsh martingales). *Let $(f_n)_{n\geqslant 0}$ be a Paley–Walsh martingale, i.e., a martingale with respect to a Paley–Walsh filtration $(\mathscr{F}_n)_{n\geqslant 0}$. Then there exist a Rademacher sequence $(r_n)_{n\geqslant 1}$ and functions $\phi_n : \{-1, 1\}^{n-1} \to X$ so that for all $n \geqslant 1$ we have $\mathscr{F}_n = \sigma(r_1, \dots, r_n)$, the σ-algebra generated by $r_1, \dots, r_n$, and*

$$df_n = r_n \phi_n(r_1, \dots, r_{n-1}), \quad n \geqslant 1,$$

with the understanding that $df_1 = r_1 \phi_1$ for some $\phi_1 \in X$.

Proof. The labelling of the atoms of $\mathscr{F}_n$ discussed below Definition 3.1.8 defines a Rademacher sequence $(r_n)_{n\geqslant 1}$ by declaring $r_n :\equiv \pm 1$ on the atom F in $\mathscr{F}_n$ according to whether the last sign $s_n = \pm$ in the representation $F = \Omega^{s_1 \cdots s_n}$ is a plus or a minus. Clearly, for $n \geqslant 1$ we have $\mathscr{F}_n = \sigma(r_1, \dots, r_n)$.

Denoting by $\mathscr{F}_n^*$ the set of all atoms in $\mathscr{F}_n$, for $n \geqslant 1$ we may write

$$\begin{aligned} df_n &= \sum_{A \in \mathscr{F}_{n-1}^*} \mathbf{1}_A \, df_n \\ &= \sum_{A \in \mathscr{F}_{n-1}^*} \mathbf{1}_A \sum_{\substack{A' \in \mathscr{F}_n^* \\ A' \subseteq A}} \mathbf{1}_{A'} \, df_n = \sum_{A \in \mathscr{F}_{n-1}^*} \mathbf{1}_A (\mathbf{1}_{A^+} + \mathbf{1}_{A^-}) \, df_n, \end{aligned}$$

using the notations introduced below Definition 3.1.8 in the last expression.

Since df_n is $\mathscr{F}_n$-measurable, it takes constant values, say x_{A^+} and x_{A^-}, on the atoms A^+ and A^-. Since $\mathbb{E}(df_n|\mathscr{F}_{n-1}) = 0$ and $\mu(A^\pm) = \frac{1}{2}\mu(A)$, it follows that $x_{A^+} = -x_{A^-}$. Thus

$$\begin{aligned} df_n &= \sum_{A\in\mathscr{F}^*_{n-1}} \mathbf{1}_A x_{A^+}(\mathbf{1}_{A^+} - \mathbf{1}_{A^-}) \\ &= \Big(\sum_{A\in\mathscr{F}^*_{n-1}} \mathbf{1}_A x_{A^+}\Big)\Big(\sum_{A\in\mathscr{F}^*_{n-1}} \mathbf{1}_{A^+} - \mathbf{1}_{A^-}\Big) = \Big(\sum_{A\in\mathscr{F}^*_{n-1}} \mathbf{1}_A x_{A^+}\Big) r_n. \end{aligned}$$

Since each indicator $\mathbf{1}_A$ with $A \in \mathscr{F}^*_{n-1}$ can be expressed as a function of $r_1, \ldots, r_{n-1}$, this proves that df_n is of the required form. □

A very special case of Paley–Walsh martingales consist of those where $f_0 = 0$ and each ϕ_n in the representation of Proposition 3.1.10 takes a constant value $x \in X$. In this case, the martingale is given simply by

$$f_n = \sum_{k=1}^n df_k = \sum_{k=1}^n r_k x_k.$$

Such *random sums* will pop up every once in a while throughout this volume, and develop into one of the main characters of Volume II.

3.1.c Stopped martingales

The notion of stopping time lies at the heart of several powerful arguments that we shall explore in this and the subsequent chapters. It formalises the intuitive idea of carrying out business as usual, as long as everything is well, and adjusting the strategy the first moment that something goes wrong (a 'stopping condition'). Whereas the existence of a 'first moment' may be a somewhat delicate matter in the context of continuous time stochastic processes, it is relatively straightforward in the realm of the discrete-parameter sequences of our primary concern, where the only issue is the possibility of the stopping time taking the value $-\infty$ (there are arbitrarily early occurrences of the stopping condition) or $+\infty$ (the stopping condition will never occur). To account for these possibilities, we consider the set of extended integers $\overline{\mathbb{Z}} = \mathbb{Z} \cup \{-\infty, \infty\}$. The heuristic idea of a stopping time is then formalised as follows:

Definition 3.1.11 (Stopping time). *A mapping $\tau : S \to \overline{\mathbb{Z}}$ is called a* stopping time *with respect to the filtration $(\mathscr{F}_n)_{n\in\mathbb{Z}}$ if for all $n \in \mathbb{Z}$ we have $\{\tau \leqslant n\} \in \mathscr{F}_n$.*

Clearly, τ is a stopping time if and only if $\{\tau = n\} \in \mathscr{F}_n$ for all $n \in \mathbb{Z}$, or equivalently, if and only if $\{\tau \geqslant n + 1\} \in \mathscr{F}_n$ for all $n \in \mathbb{Z}$.

If τ_1 and τ_2 are both stopping times, then so are $\tau_1 \wedge \tau_2$ and $\tau_1 \vee \tau_2$:

$$\{\tau_1 \wedge \tau_2 \leqslant n\} = \{\tau_1 \leqslant n\} \cup \{\tau_2 \leqslant n\} \in \mathscr{F}_n,$$

$$\{\tau_1 \vee \tau_2 \leqslant n\} = \{\tau_1 \leqslant n\} \cap \{\tau_2 \leqslant n\} \in \mathscr{F}_n.$$

Example 3.1.12 (First hitting time). Let $(f_n)_{n\in\mathbb{Z}}$ be a sequence in $L^0(S;X)$ which is adapted to the filtration $(\mathscr{F}_n)_{n\in\mathbb{Z}}$. Let $B \subseteq X$ be a Borel set. Let $\tau : S \to \overline{\mathbb{Z}}$ be defined by

$$\tau := \inf\{n \in \mathbb{Z} : f_n \in B\},$$

where $\tau = \infty$ if the infimum is taken over the empty set and $\tau = -\infty$ if $f_n \in B$ for all $n \in \mathbb{Z}$. Here it is understood that we work with strongly $\mathscr{F}_n$-measurable representatives of the f_n. From

$$\{\tau \leqslant n\} = \bigcup_{m \leqslant n} \{f_m \in B\}$$

and noting that $\{f_m \in B\} \in \mathscr{F}_m \subseteq \mathscr{F}_n$ we see that $\{\tau \leqslant n\} \in \mathscr{F}_n$. The stopping time τ is called *the first hitting time* of B. Of course, τ depends on the particular choice of representatives of f_n, but different choices will always produce the same stopping time up to a μ-null set $N \in \mathscr{A}$. Thus in order to be completely rigorous, we would have to pass to equivalence classes of hitting times. As in the case of functions, we shall refrain from making the distinction between τ and its equivalence class, and leave it to the reader to check this slight ambiguity never leads to any problems.

If τ is a stopping time with values in $\mathbb{Z}$ and $(f_n)_{n\in\mathbb{Z}}$ is a sequence in $L^0(S;X)$ we define $f_\tau \in L^0(S;X)$ by

$$f_\tau := f_n \text{ on } \{\tau = n\}, \quad n \in \mathbb{Z}.$$

When τ is a stopping time with values in $\overline{\mathbb{Z}} := \mathbb{Z} \cup \{-\infty, +\infty\}$ and the limit $f_{\pm\infty}(s) := \lim_{n\to\pm\infty} f_n(s)$ exists almost everywhere on $\{\tau = \pm\infty\}$, we set $f_\tau := f_{\pm\infty}$ on this set.

In the rest of this subsection, unless stated otherwise, stopping times always take values in $\overline{\mathbb{Z}}$.

Definition 3.1.13 (Stopped sequence). *Let τ be a stopping time and let $f = (f_n)_{n\in\mathbb{Z}}$ be an adapted sequence in $L^0(S;X)$ with difference sequence $df = (df_n)_{n\in\mathbb{Z}}$. When the limit $f_{-\infty} = \lim_{n\to-\infty} f_n$ exists almost everywhere on $\{\tau = -\infty\}$, the* stopped sequence *$f^\tau = (f_n^\tau)_{n\in\mathbb{Z}}$ is defined by*

$$f_n^\tau = f_{\tau\wedge n}, \qquad n \in \mathbb{Z}.$$

Likewise, for a stopping time σ we may define the started sequence *${}^\sigma f = ({}^\sigma f_n)_{n\in\mathbb{Z}}$ by*

$${}^\sigma f_n = f_n - f_{\sigma\wedge n} = f_n - f_n^\sigma.$$

The difference sequences of f^τ and ${}^\sigma f$ are given by

$$df_n^\tau = \begin{cases} df_n \text{ on } \{n \leqslant \tau\} \\ 0 \quad \text{ on } \{n > \tau\} \end{cases} = \mathbf{1}_{\{n\leqslant\tau\}} df_n$$

and

$$d^\sigma f_n = \begin{cases} 0 \quad \text{ on } \{n \leqslant \sigma\} \\ df_n \text{ on } \{n > \sigma\} \end{cases} = \mathbf{1}_{\{n>\sigma\}} df_n.$$

We further define ${}^\sigma f^\tau := {}^\sigma(f^\tau) = ({}^\sigma f)^\tau = f^\tau - f^{\tau\wedge\sigma}$, that is,

$${}^\sigma f_n^\tau := \mathbf{1}_{\{\sigma<\tau\}}(f_{\tau\wedge n} - f_{\sigma\wedge n}) = \mathbf{1}_{\{\sigma\leqslant\tau\}}(f_{\tau\wedge n} - f_{\sigma\wedge n}) = f_{\tau\wedge n} - f_{\tau\wedge\sigma\wedge n}.$$

Its difference sequence is given by

$$d^\sigma f_n^\tau = \begin{cases} d_n \text{ on } \{\sigma < n \leqslant \tau\} \\ 0 \quad \text{ on } \{n \leqslant \sigma\} \cup \{\tau < n\} \end{cases} = \mathbf{1}_{\{\sigma<n\leqslant\tau\}} df_n.$$

Proposition 3.1.14 (Stopped martingales). *Let $f = (f_n)_{n\in\mathbb{Z}}$ be an X-valued martingale and τ and σ be stopping times, all with respect to a σ-finite filtration $(\mathscr{F}_n)_{n\in\mathbb{Z}}$. Suppose that $f_{-\infty} := \lim_{n\to-\infty} f_n$ exists almost everywhere on $\{\tau = -\infty\} \cup \{\sigma = -\infty\}$.*

(1) *The sequences f^τ, ${}^\sigma f$, and ${}^\sigma f^\tau$ are martingales with respect to $(\mathscr{F}_n)_{n\in\mathbb{Z}}$.*
(2) *If $(f_n)_{n\in\mathbb{Z}}$ is an L^p-martingale for some $p \in [1,\infty)$, then f^τ, ${}^\sigma f$, and ${}^\sigma f^\tau$ are L^p-martingales, and for all $m \leqslant n$ we have*

$$\begin{aligned} \|f_m^\tau\|_{L^p(S;X)} &\leqslant \|f_n\|_{L^p(S;X)}, \\ \|{}^\sigma f_m\|_{L^p(S;X)} &\leqslant 2\|f_n\|_{L^p(S;X)}, \\ \|{}^\sigma f_m^\tau\|_{L^p(S;X)} &\leqslant 2\|f_n\|_{L^p(S;X)}. \end{aligned} \tag{3.1}$$

A sufficient condition for the almost sure existence of the limit $f_{-\infty} := \lim_{n\to-\infty} f_n$ will be given in Theorem 3.3.5.

Proof. The martingale property for f^τ follows from Proposition 2.6.31, using that $\{\tau \geqslant n\} = \complement\{\tau \leqslant n-1\} \in \mathscr{F}_{n-1}$:

$$\mathbb{E}(df_n^\tau|\mathscr{F}_{n-1}) = \mathbf{1}_{\{n\leqslant\tau\}}\mathbb{E}(df_n|\mathscr{F}_{n-1}) = 0.$$

It is an immediate consequence that ${}^\sigma f = f - f^\sigma$ and ${}^\sigma f^\tau = f^\tau - f^{\tau\wedge\sigma}$ are martingales with respect to $(\mathscr{F}_n)_{n\in\mathbb{Z}}$ as well.

We continue with the L^p-inequality. Fix $1 \leqslant p < \infty$ and $n \in \mathbb{Z}$. Then

$$f_{\tau\wedge n} = \mathbf{1}_{\{\tau=-\infty\}} f_{-\infty} + \sum_{k\leqslant n} \mathbf{1}_{\{\tau=k\}} f_k$$

has a strongly $\mathscr{F}_n$-measurable version. We claim that for all $k \leqslant n$ we have $f_{\tau\wedge k} \in L^p(S;X)$ and $\|f_{\tau\wedge k}\|_{L^p(S;X)} \leqslant \|f_n\|_{L^p(S;X)}$. Indeed, for $i \leqslant j \leqslant n$, the conditional Jensen inequality gives us

$$\begin{aligned}\int_{\{\tau=i\}} \|f_j\|^p \,\mathrm{d}\mu &= \int_{\{\tau=i\}} \|\mathbb{E}(f_n|\mathscr{F}_j)\|^p \,\mathrm{d}\mu \\ &\leqslant \int_{\{\tau=i\}} \mathbb{E}(\|f_n\|^p|\mathscr{F}_j) \,\mathrm{d}\mu = \int_{\{\tau=i\}} \|f_n\|^p \,\mathrm{d}\mu.\end{aligned} \tag{3.2}$$

Taking $i = -\infty$, in combination with Fatou's lemma this gives

$$\int_{\{\tau=-\infty\}} \|f_{-\infty}\|^p \,\mathrm{d}\mu \leqslant \liminf_{j\to-\infty} \int_{\{\tau=-\infty\}} \|f_j\|^p \,\mathrm{d}\mu \leqslant \int_{\{\tau=-\infty\}} \|f_n\|^p \,\mathrm{d}\mu. \tag{3.3}$$

Combining (3.2) (with $i = j$) and (3.3), we find that

$$\begin{aligned}\int_S \|f_{\tau\wedge k}\|^p \,\mathrm{d}\mu &= \int_{\{\tau=-\infty\}} \|f_{-\infty}\|^p \,\mathrm{d}\mu + \sum_{j\leqslant k} \int_{\{\tau=j\}} \|f_j\|^p \,\mathrm{d}\mu \\ &\leqslant \int_{\{\tau=-\infty\}} \|f_n\|^p \,\mathrm{d}\mu + \sum_{j\leqslant k} \int_{\{\tau=j\}} \|f_n\|^p \,\mathrm{d}\mu \leqslant \int_S \|f_n\|^p \,\mathrm{d}\mu.\end{aligned}$$

The corresponding L^p-estimates for ${}^\sigma f = f - f^\sigma$ and ${}^\sigma f^\tau = {}^\sigma(f^\tau)$ follow from this. □

3.2 Martingale inequalities

While the basic defining property of a martingale is the equality (or a sequence of equalities) $\mathbb{E}(f_n|\mathscr{F}_{n-1}) = f_{n-1}$, a major part of our forthcoming applications of martingales will deal with estimates or inequalities that can be inferred from this identity. In this section, we present a first collection of fundamental inequalities for martingales, and also illustrate their relation with some inequalities of classical analysis.

3.2.a Doob's maximal inequalities

Our first martingale inequality is an important maximal estimate due to Doob. It is a close cousin of maximal inequalities in other areas of Analysis, notably the Hardy–Littlewood maximal inequality. This connection will be elaborated in Subsection 3.2.d below.

For a sequence of X-valued functions $f := (f_n)_{n\in\mathbb{Z}}$ in $L^0(S;X)$ we define the maximal functions

$$f_n^\star := \sup_{m\leqslant n} \|f_m\|, \quad f^\star := \sup_{m\in\mathbb{Z}} \|f_m\|. \tag{3.4}$$

Our formulation of Doob's inequality is somewhat more general than usual; its proof is not more difficult, but its extra generality readily leads to some useful consequences.

We work with a fixed σ-finite filtration $(\mathscr{F}_n)_{n\in\mathbb{Z}}$ of a measure space $(S,\mathscr{A},\mu)$, and all (sub)martingales are understood to be with respect to this reference filtration. A non-negative function $w \in L^0(S)$ is a called a *weight* (for the filtration $(\mathscr{F}_n)_{n\in\mathbb{Z}}$) if w is σ-integrable over every $\mathscr{F}_n$. This guarantees the existence of the conditional expectations $w_n := \mathbb{E}(w|\mathscr{F}_n)$, which give rise to the non-negative martingale $(w_n)_{n\in\mathbb{Z}}$ by $w_n := \mathbb{E}(w|\mathscr{F}_n)$, and the maximal function $w^\star$ by (3.4). We identify the function w with its induced measure defined by

$$w(A) := \int_A w \,\mathrm{d}\mu, \quad A \in \mathscr{A}.$$

A key to the martingale maximal inequalities is the following lemma, which also serves as an illustration of the powerful stopping time techniques.

Lemma 3.2.1. *Let $(f_n)_{n\in\mathbb{Z}}$ be a non-negative submartingale and $w \in L^0(S)$ be a weight. Then*

$$\lambda\, w(f_n^\star > \lambda) \leqslant \int_{\{f_n^\star>\lambda\}} f_n\, w^\star \,\mathrm{d}\mu, \qquad \forall \lambda > 0. \tag{3.5}$$

Proof. Let $\lambda > 0$ be arbitrary and fixed. By monotone convergence it suffices to show that for every $m \geqslant 0$,

$$\lambda w\Big(\sup_{-m\leqslant k\leqslant n} f_k > \lambda\Big) \leqslant \int_{\{f_n^\star>\lambda\}} f_n w^\star \,\mathrm{d}\mu.$$

Therefore it suffices to prove (3.5) in the case $f_k = f_{-m}$ for all $k \leqslant -m-1$. Fix $\lambda > 0$ and define the stopping time $\tau : S \to \{-m, \dots, n+1\}$ by

$$\tau := \min\big\{k \in \{-m,\dots,n\} : f_k > \lambda\big\}$$

with the convention that $\min \varnothing := n+1$. Then $\{f_n^\star > \lambda\} = \{-m \leqslant \tau \leqslant n\}$. We now compute

$$\begin{aligned}
\lambda w(f_n^\star > \lambda) &= \lambda \sum_{k=-m}^{n} \int_{\{\tau=k\}} w \,\mathrm{d}\mu = \lambda \sum_{k=-m}^{n} \int_{\{\tau=k\}} w_k \,\mathrm{d}\mu \\
&\leqslant \sum_{k=-m}^{n} \int_{\{\tau=k\}} f_k w_k \,\mathrm{d}\mu \leqslant \sum_{k=-m}^{n} \int_{\{\tau=k\}} \mathbb{E}(f_n|\mathscr{F}_k) w_k \,\mathrm{d}\mu \\
&= \sum_{k=-m}^{n} \int_{\{\tau=k\}} \mathbb{E}(f_n w_k|\mathscr{F}_k) \,\mathrm{d}\mu = \sum_{k=-m}^{n} \int_{\{\tau=k\}} f_n w_k \,\mathrm{d}\mu \\
&\leqslant \sum_{k=-m}^{n} \int_{\{\tau=k\}} f_n w^\star \,\mathrm{d}\mu = \int_{\{f_n^\star>\lambda\}} f_n w^\star \,\mathrm{d}\mu,
\end{aligned}$$

where we used the submartingale property, applied Proposition 2.6.31, and used the definition of the conditional expectation. $\square$

We now prove the fundamental inequalities of Doob:

Theorem 3.2.2 (Doob's maximal inequality). *Let $(f_n)_{n\in\mathbb{Z}}$ be a non-negative submartingale in $L^p(S)$, or a martingale in $L^p(S;X)$. Then*

$$\begin{aligned} \mu(f_n^\star > \lambda) &\leqslant \frac{1}{\lambda}\|f_n\|_1, \quad \text{if} \quad p=1, \\ \|f_n^\star\|_p &\leqslant p'\|f_n\|_p, \quad \text{if} \quad p\in(1,\infty]. \end{aligned} \tag{3.6}$$

Proof. We may assume that $(f_n)_{n\in\mathbb{Z}}$ is a non-negative submartingale, since the martingale case follows by considering the non-negative submartingale $(\|f_n\|_X)_{n\in\mathbb{Z}}$ instead. For the special case $w\equiv 1$, (3.5) becomes

$$\lambda\,\mu(f_n^\star>\lambda) \leqslant \int_{\{f_n^\star>\lambda\}} f_n\,\mathrm{d}\mu \leqslant \|f_n\|_1, \qquad \forall\lambda>0, \tag{3.7}$$

which proves the first bound in (3.6). Moreover, another application of (3.7) implies

$$\begin{aligned} \|f_n^\star\|_p^p &= \int_0^\infty p\lambda^{p-1}\mu(f_n^\star>\lambda)\,\mathrm{d}\lambda \leqslant \int_0^\infty p\lambda^{p-2}\int_{\{f_n^\star>\lambda\}} f_n\,\mathrm{d}\mu\,\mathrm{d}\lambda \\ &= \int_S\Big(\int_0^{f_n^\star(s)} p\lambda^{p-2}\,\mathrm{d}\lambda\Big) f_n(s)\,\mathrm{d}\mu(s) = p'\int_S f_n(f_n^\star)^{p-1}\,\mathrm{d}\mu \\ &\leqslant p'\|f_n\|_p\|f_n^\star\|_p^{p-1}. \end{aligned} \tag{3.8}$$

Dividing by $\|f_n^\star\|_p^{p-1}$, if this quantity is finite, the estimate $\|f_n^\star\|_p \leqslant p'\|f_n\|_p$ follows. The *a priori* finiteness is guaranteed by monotone convergence if one first replaces $f_n^\star$ by $\max_{-m\leqslant k\leqslant n} f_k$ and then passes to the limit $m\to\infty$. □

The classical estimate of Theorem 3.2.2 can be extended to a more general weighted setting by invoking the full strength of Lemma 3.2.1. However, in this case the direct computation (3.8) also needs to be replaced by an application of the Marcinkiewicz interpolation theorem. We remind the reader of the notation $L^{p,\infty}$ for the weak L^p-spaces introduced in Section 2.2.b. For a non-negative function $w\in L^0(S)$ we define

$$L^p(S,w;X) = L^p(S,w\,\mathrm{d}\mu;X)$$

to be the Banach space of all $f\in L^0(S;X)$ such that $fw^{1/p}\in L^p(S;X)$. Using these notations, we can prove the following:

Theorem 3.2.3 (Doob's maximal inequality with weight). *Let $(f_n)_{n\in\mathbb{Z}}$ be a non-negative submartingale or an X-valued martingale, let $w\in L^0(S)$ be a weight and $w^\star := \sup_{n\in\mathbb{Z}}\mathbb{E}(w|\mathscr{F}_n)$. Then for all $n\in\mathbb{Z}$ we have*

$$\|f_n^\star\|_{L^{p,\infty}(S,w)} \leqslant \|f_n\|_{L^p(S,w^\star)}, \qquad p\in[1,\infty), \tag{3.9}$$

$$\|f_n^\star\|_{L^p(S,w)} \leqslant p'\|f_n\|_{L^p(S,w^\star)}, \qquad p\in(1,\infty], \tag{3.10}$$

$$\|f_n^\star\|_{L^{p,\infty}(S,w)} \leqslant p'\|f_n\|_{L^{p,\infty}(S,w^\star)}, \qquad p\in(1,\infty). \tag{3.11}$$

Proof. We may again concentrate on the non-negative submartingale case. Lemma 3.2.1 contains the result (3.9) for $p = 1$. For $p \in (1, \infty)$, we observe that $(f_n^p)_{n\in\mathbb{Z}}$ is also a submartingale by Jensen's inequality, so that applying Lemma 3.2.1 to this instead, we obtain

$$\|f_n^\star\|_{L^{p,\infty}(S,w)} = \|(f_n^p)^\star\|_{L^{1,\infty}(S,w)}^{1/p} \leqslant \|f_n^p\|_{L^1(S,w^\star)}^{1/p} = \|f_n\|_{L^p(S,w^\star)}.$$

For the other two estimates, we shall employ the special case of the Marcinkiewicz interpolation theorem given in Corollary 2.2.4. Fix $n \in \mathbb{Z}$ and consider the sub-linear mapping $g \mapsto T_n^\star g$,

$$T_n^\star g := \sup_{k \leqslant n} \mathbb{E}(|g||\mathscr{F}_k).$$

Since $g_k := \mathbb{E}(|g||\mathscr{F}_k)$ for $k \leqslant n$ defines a non-negative submartingale, it follows from (3.9) that

$$T_n^\star : L^1(S, \mathscr{F}_n, w^\star) \to L^{1,\infty}(S, \mathscr{F}_n, w)$$

is bounded with $\|T_n^\star f\|_{L^{1,\infty}} \leqslant \|f\|_{L^1}$. The boundedness of

$$T_n^\star : L^\infty(S, \mathscr{F}_n, w^\star) \to L^\infty(S, \mathscr{F}_n, w),$$

with $\|T_n^\star f\|_{L^\infty} \leqslant \|f\|_{L^\infty}$, is immediate from the monotonicity of conditional expectations. Therefore, by Corollary 2.2.4 and using that $0 \leqslant f_k \leqslant \mathbb{E}(f_n|\mathscr{F}_k)$ for all $k \leqslant n$, we obtain (3.10) and (3.11):

$$\begin{aligned}
\|f_n^\star\|_{L^p(S,w)} &\leqslant \|T_n^\star f_n\|_{L^p(S,w)} \leqslant p'\|f_n\|_{L^p(S,w^\star)}, \\
\|f_n^\star\|_{L^{p,\infty}(S,w)} &\leqslant \|T_n^\star f_n\|_{L^{p,\infty}(S,w)} \leqslant p'\|f_n\|_{L^{p,\infty}(S,w^\star)}.
\end{aligned}$$

□

The constant in Doob's inequality

It will be shown next that Doob's inequality in (3.10) (and in particular its classical version in the second part of (3.6)) is quantitatively optimal. As a consequence of this and the proof of Theorem 3.2.3 one can deduce that the constant in Corollary 2.2.4 is sharp as well.

Proposition 3.2.4. *For all $1 < p \leqslant \infty$ the best constant c_p in Doob's inequality $\|f^\star\|_p \leqslant c_p\|f\|_p$ is $c_p = p'$, i.e., this inequality is invalid for any $c_p < p'$. The same holds for the estimate $\|f^\star\|_{L^{p,\infty}} \leqslant c_p\|f\|_{L^{p,\infty}}$.*

We already know that the inequality is valid for $c_p = p'$. To see that it cannot hold for a smaller constant, our strategy is to show that:

(1) Doob's inequality implies a certain other estimate (Hardy's inequality), with the same constant;

(2) the constant $c_p = p'$ is optimal for Hardy's inequality, by an explicit example.

Lemma 3.2.5. *If Doob's inequality holds with a constant c_p, then so does Hardy's inequality*

$$\Big\| x \mapsto \frac{1}{x}\int_0^x f(y)\,\mathrm{d}y \Big\|_{L^p(0,\infty)} \leqslant c_p \|f\|_{L^p(0,\infty)}. \tag{3.12}$$

The same holds if $L^p(0,\infty)$ is replaced by $L^{p,\infty}(0,\infty)$.

Proof. For $\delta > 0$, consider the filtration on $(0,\infty)$ given by

$$\mathscr{F}^\delta_{-n} := \{A \cup B : A \in \{\varnothing, (0, \delta n]\}, B \in \mathscr{B}(\delta n, \infty)\}, \quad n \in \mathbb{N},$$

and let $f^\star_\delta := \sup_n \mathbb{E}(f|\mathscr{F}^\delta_{-n})$ be the associated Doob maximal function. Then for a non-negative function f we find (cf. Example 2.6.13)

$$\begin{aligned} f^\star_\delta(x) \geqslant \mathbb{E}(f|\mathscr{F}^\delta_{-n})(x) &= \frac{1}{n\delta}\int_0^{n\delta} f(y)\,\mathrm{d}y, && \text{if } x \leqslant n\delta, \\ &\geqslant \frac{1}{x+\delta}\int_0^{x} f(y)\,\mathrm{d}y, && \text{if } (n-1)\delta < x \leqslant n\delta. \end{aligned}$$

Thus, from monotone convergence and Doob's inequality we deduce that

$$\begin{aligned} \Big\| x \mapsto \frac{1}{x}\int_0^x f(y)\,\mathrm{d}y \Big\|_{L^p(0,\infty)} &= \lim_{\delta\to 0} \Big\| x \mapsto \frac{1}{x+\delta}\int_0^x f(y)\,\mathrm{d}y \Big\|_{L^p(0,\infty)} \\ &\leqslant \limsup_{\delta\to 0} \|f^\star_\delta\|_{L^p(0,\infty)} \leqslant c_p\|f\|_{L^p(0,\infty)}. \end{aligned}$$

Similarly, for the weak L^p-norm we find

$$\begin{aligned} \lambda\,\mu\Big(\Big\{x : \frac{1}{x}\int_0^x f(y)\,\mathrm{d}y > \lambda\Big\}\Big)^{1/p} &= \lim_{\delta\to 0} \lambda\,\mu\Big(\Big\{x : \frac{1}{x+\delta}\int_0^x f(y)\,\mathrm{d}y > \lambda\Big\}\Big)^{1/p} \\ &\leqslant \limsup_{\delta\to 0} \|f^\star_\delta\|_{L^{p,\infty}(0,\infty)} \leqslant c_p\|f\|_{L^{p,\infty}(0,\infty)}. \end{aligned}$$

Now the result follows by taking the supremum over all $\lambda > 0$. $\square$

Lemma 3.2.6. *The best constant in Hardy's inequality (3.12) is $c_p = p'$. The same holds if $L^p(0,\infty)$ is replaced by $L^{p,\infty}(0,\infty)$.*

Proof. From Doob's inequality and Lemma 3.2.5 we already know that (3.12) holds with $c_p = p'$. We now show that it cannot hold for any smaller constant. This is done by investigating (3.12) with the particular functions $f_\alpha(x) := \mathbf{1}_{(0,1)}(x)\cdot x^{-\alpha}$ with $\alpha \in (0,1)$.

For all $x > 0$ one has

$$\frac{1}{x}\int_0^x f_\alpha(y)\,\mathrm{d}y \geqslant \frac{\mathbf{1}_{(0,1)}(x)}{x}\int_0^x y^{-\alpha}\,\mathrm{d}y = \frac{\mathbf{1}_{(0,1)}(x)}{x}\,\frac{x^{1-\alpha}}{1-\alpha} = \frac{f_\alpha(x)}{1-\alpha},$$

so (3.12) implies that $\|f_\alpha/(1-\alpha)\|_p \leqslant c_p\|f_\alpha\|_p$ and thus $1/(1-\alpha) \leqslant c_p$, *provided that* $f_\alpha \in L^p(0,\infty)$. This last condition holds if and only if $\alpha p < 1$, and thus we conclude that

$$c_p \geqslant \sup_{\alpha<1/p} \frac{1}{1-\alpha} = \frac{1}{1-1/p} = \frac{p}{p-1} = p'.$$

The same argument holds with L^p replaced by $L^{p,\infty}$. □

Doob's maximal inequality in ℓ^q

Doob's L^p-maximal inequality admits a non-trivial ℓ^q-valued extension. This result is the key ingredient in the proof, given in the next subsection, of the Fefferman–Stein maximal inequality (Theorem 3.2.28).

Suppose that for each $j \geqslant 1$ a sequence of functions $(f_n^{(j)})_{n\in\mathbb{Z}}$ is given. Its 'indexwise' maximal function is then defined as

$$f_n^{(j)\star} := \sup_{k\leqslant n} |f_k^{(j)}|, \quad j \geqslant 1.$$

Theorem 3.2.7. *Let $p \in (1,\infty)$ and $q \in (1,\infty]$. For each $j \geqslant 1$, let $(f_n^{(j)})_{n\in\mathbb{Z}}$ be a non-negative submartingale. If $(f_n^{(j)})_{j\in\mathbb{Z}} \in L^p(S;\ell^q)$ for all $n \in \mathbb{Z}$, then $(f_n^{(j)\star})_{j\in\mathbb{Z}} \in L^p(S;\ell^q)$ for all $n \in \mathbb{Z}$ and*

$$\|(f_n^{(j)\star})_{j\in\mathbb{Z}}\|_{L^p(S;\ell^q)} \leqslant C_{p,q}\|(f_n^{(j)})_{j\in\mathbb{Z}}\|_{L^p(S;\ell^q)}, \quad n \in \mathbb{Z},$$

where $C_{p,q}$ is a constant which only depends on p and q.

It will follow from Example 3.2.29 that this result does not hold for $p = \infty$ and $q \in (1,\infty)$. For $p = q = \infty$ the result holds trivially.

Proof. An easy argument involving Fubini's theorem shows that the case $p = q$ readily follows from Theorem 3.2.3.

For $p \in (1,\infty)$ and $q = \infty$, note that $(\|(f_n^{(j)})_{j\in\mathbb{Z}}\|_{\ell^\infty})_{n\in\mathbb{Z}}$ is a submartingale by Lemma 2.6.17. Indeed, for each $j \geqslant 1$ we have $f_{n-1}^{(j)} \leqslant \mathbb{E}(f_n^{(j)}|\mathscr{F}_{n-1}) \leqslant \mathbb{E}(\|f_n\|_{\ell^\infty}|\mathscr{F}_{n-1})$ and the result follows by taking the supremum over $j \geqslant 1$. Therefore, Theorem 3.2.3 yields

$$\begin{aligned}\|(f_n^{(j)\star})_{j\in\mathbb{Z}}\|_{L^p(S;\ell^\infty)}^p &= \int_S \sup_{j\geqslant 1}\sup_{k\leqslant n} |f_k^{(j)}|^p \,\mathrm{d}\mu \\ &\leqslant \int_S \sup_{k\leqslant n} \|(f_k^{(j)})_{j\in\mathbb{Z}}\|_{\ell^\infty}^p \,\mathrm{d}\mu \leqslant (p')^p \|(f_n^{(j)})_{j\in\mathbb{Z}}\|_{L^p(S;\ell^\infty)}^p.\end{aligned}$$

Turning to the case $1 < p, q < \infty$, by approximation we may assume that for all $n \in \mathbb{Z}$ we have $f_n^{(j)} = 0$ for all $j \geqslant j_0$ and that for each $j \geqslant 1$ the sequence $(f_n^{(j)})_{n\in\mathbb{Z}}$ becomes constant for $n \leqslant n_1$ (replace $f_n^{(j)}$ and $\mathscr{F}_n$ by $f_{n_1}^{(j)}$

and $\mathscr{F}_{n_1}$ if necessary). This will avoid convergence issues in the proof; at the end we may pass to the limits $j_0 \to \infty$ and $n_1 \to -\infty$.

First we consider the case $1 < q < p < \infty$. Set $r := \frac{p}{p-q}$ (so that $r' = p/q$). Choose a non-negative $w \in L^r(S)$ of norm at most one in such a way that

$$\Big(\int_S \|f_n^\star\|_{\ell^q}^p \,\mathrm{d}\mu\Big)^{q/p} = \int_S \|f_n^\star\|_{\ell^q}^q w \,\mathrm{d}\mu.$$

Then, by the weighted and unweighted Doob's maximal inequality and Hölder's inequality,

$$\begin{aligned}
\Big(\int_S \|(f_n^{(j)\star})_{j\in\mathbb{Z}}\|_{\ell^q}^p \,\mathrm{d}\mu\Big)^{q/p} &= \sum_{j\geqslant 1} \int_S |\sup_{k\leqslant n} f_k^{(j)}|^q w \,\mathrm{d}\mu \\
&\leqslant (q')^q \sum_{j\geqslant 1} \int_S |f_n^{(j)}|^q w^\star \,\mathrm{d}\mu \\
&= (q')^q \int_S \sum_{j\geqslant 1} |f_n^{(j)}|^q w^\star \,\mathrm{d}\mu \\
&\leqslant (q')^q \Big(\int_S \Big(\sum_{j\geqslant 1} |f_n^{(j)}|^q\Big)^{p/q} \,\mathrm{d}\mu\Big)^{q/p} \|w^\star\|_{L^r(S)} \\
&\leqslant (q')^q r' \Big(\int_S \|(f_n^{(j)})_{j\in\mathbb{Z}}\|_{\ell^q}^p \,\mathrm{d}\mu\Big)^{q/p},
\end{aligned}$$

using that $\|w^\star\|_{L^r(S)} \leqslant r'\|w\|_{L^r(S)} \leqslant r'$ in the last step. Therefore,

$$\|(f_n^{(j)\star})_{j\in\mathbb{Z}}\|_{L^p(S;\ell^q)} \leqslant q'\big(\tfrac{p}{q}\big)^{1/q} \|(f_n^{(j)})_{j\in\mathbb{Z}}\|_{L^p(S;\ell^q)}.$$

The case $1 < p < q < \infty$ will be reduced to the previous case by another duality argument. Fix some $r \in (1,p)$ and set $a := p/r$ and $b := q/r$. Let $\varepsilon > 0$ be arbitrary. By Proposition 1.3.1 we can find a non-negative $g = (g^{(j)})_{j\geqslant 1} \in L^{a'}(S;\ell^{b'})$ of norm $\leqslant 1+\varepsilon$ and such that

$$\Big(\int_S \|(f_n^{(j)\star})_{j\in\mathbb{Z}}\|_{\ell^q}^p \,\mathrm{d}\mu\Big)^{r/p} = \big\|(|f_n^{(j)\star}|^r)_{j\in\mathbb{Z}}\big\|_{L^a(S;\ell^b)} \leqslant \sum_{j\geqslant 1} \int_S |(f_n^{(j)\star})|^r g^{(j)} \,\mathrm{d}\mu,$$

which may be estimated, using Theorem 3.2.3 with $w = g^{(j)}$, by

$$\begin{aligned}
&\leqslant (r')^r \sum_{j\geqslant 1} \int_S |f_n^{(j)}|^r (g^{(j)\star}) \,\mathrm{d}\mu \\
&\leqslant (r')^r \Big(\int_S \Big(\sum_{j\geqslant 1} |f_n^{(j)}|^q\Big)^{p/q} d\mu\Big)^{r/p} \|(g^{(j)\star})_{j\in\mathbb{Z}}\|_{L^{a'}(S;\ell^{b'})} \\
&\leqslant (1+\varepsilon)(r')^r b\big(\tfrac{a'}{b'}\big)^{1/b'} \Big(\int_S \|f_n\|_{\ell^q}^p \,\mathrm{d}\mu\Big)^{r/p},
\end{aligned}$$

where in the last step we applied the previous case with $b' < a'$. □

The dual version of Doob's inequality

Notice that Doob's L^p-inequality may been viewed as the L^p-to-$L^p(\ell^\infty)$ boundedness of the linear operator $f \mapsto (\mathbb{E}(f|\mathscr{F}_n))_{n\in\mathbb{Z}}$. Dualising with $L^{p'}(\ell^1)$, this is seen to be equivalent with the following estimate. Here we provide a different proof of this dual estimate, which has independent interest.

Proposition 3.2.8. *Let $p \in [1,\infty)$. For every non-negative summable sequence $(g_n)_{n\in\mathbb{Z}}$ of measurable functions in $L^p(S)$:*

$$\Big\| \sum_{n\in\mathbb{Z}} \mathbb{E}(g_n|\mathscr{F}_n)\Big\|_p \leqslant p\Big\| \sum_{n\in\mathbb{Z}} g_n\Big\|_p.$$

We emphasise that the sequence $(g_n)_{n\in\mathbb{Z}}$ is not assumed to be adapted.

Proof. It will be convenient to introduce the shorthand notation $\mathbb{E}_n$ for the conditional expectation with respect to $\mathscr{F}_n$. By an approximation argument it suffices to consider the case where only finitely many g_n are non-zero. Observe that by non-negativity $\mathbb{E}_j g_j \geqslant 0$ (Lemma 2.6.17). To prove the result we use the following simple consequence of the mean value theorem:

$$y^p - x^p \leqslant p(y-x)y^{p-1}, \quad y \geqslant x \geqslant 0. \tag{3.13}$$

Put $G_\infty := \sum_{j\in\mathbb{Z}} g_j$ and let the conditioned partial sums be denoted by $C_n := \sum_{j\leqslant n} \mathbb{E}_j(g_j)$ for each n. By (3.13), we have $C_n^p - C_{n-1}^p \leqslant p\mathbb{E}_n(g_n)C_n^{p-1}$. Integrating both sides of the estimate with respect to μ, the self-adjointness of the conditional expectation (Proposition 2.6.32) gives

$$\int_S C_n^p - C_{n-1}^p \,\mathrm{d}\mu \leqslant p\int_S \mathbb{E}_n(g_n)C_n^{p-1}\,\mathrm{d}\mu = p\int_S g_n C_n^{p-1}\,\mathrm{d}\mu \leqslant p\int_S g_n C_\infty^{p-1}\,\mathrm{d}\mu.$$

Summing over $|n| \leqslant N$ and sending $N \to \infty$ we obtain

$$\begin{aligned}\|C_\infty\|_p^p &= \sum_{n\in\mathbb{Z}} \int_S C_n^p - C_{n-1}^p \,\mathrm{d}\mu \\ &\leqslant \sum_{n\in\mathbb{Z}} p\int_S g_n C_\infty^{p-1}\,\mathrm{d}\mu = p\int_S G_\infty C_\infty^{p-1}\,\mathrm{d}\mu \leqslant p\|G_\infty\|_p\|C_\infty\|_p^{p-1},\end{aligned}$$

applying Hölder's inequality in the last step. Now the result follows upon dividing both sides by $\|C_\infty\|_p^{p-1}$. □

3.2.b Rademacher variables and contraction principles

We now turn to an estimate of a rather different nature. It concerns *random sums* of the form

$$\sum_{n=1}^N \varepsilon_n x_n, \tag{3.14}$$

where the x_n are elements of a Banach space with scalar field $\mathbb{K}$, and $(\varepsilon_n)_{n=1}^N$ is a $\mathbb{K}$-*valued Rademacher sequence.* We have already encountered the real-valued version of these random variables in the representation of Paley–Walsh martingales in Proposition 3.1.10; the general definition extends the real case as follows:

Definition 3.2.9. *A* Rademacher variable *is a random variable (i.e., a measurable function)* $\varepsilon : \Omega \to \mathbb{K}$*, defined on some probability space* $(\Omega, \mathscr{F}, \mathbb{P})$*, which is uniformly distributed over* $S_{\mathbb{K}} := \{z \in \mathbb{K} : |z| = 1\}$*. A* Rademacher sequence *is a sequence* $(\varepsilon_i)_{i\in I}$ *of independent Rademacher random variables.*

It is implicit in the definition that the notion of a Rademacher variable depends on the choice of scalar field $\mathbb{K}$. When we wish to make the distinction, we speak of *real* or *complex Rademacher variables* according to the choice of $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Sometimes, as in Proposition 3.1.10, we explicitly insist on the use of the real Rademacher variables (even in a possibly complex Banach space X), and denote them by the symbol r. Complex Rademacher variables are sometimes called *Steinhaus random variables.* The simplest example of a complex Rademacher variable is obtained by taking $\varepsilon(\omega) = e^{2\pi i\omega}$ for $\omega \in \Omega = [0, 1]$; Rademacher sequences can be defined on products of $[0, 1]$.

The uniform distribution of ε_i over $\{z \in \mathbb{K} : |z| = 1\}$ implies in particular the vanishing average $\mathbb{E}\varepsilon_i = \int_\Omega \varepsilon_i \,\mathrm{d}\mathbb{P} = 0$. If $(\varepsilon_n)_{n\geqslant 1}$ is a Rademacher sequence, by independence it follows that $f_n := \sum_{k=1}^n \varepsilon_k x_k$ defines a martingale with respect to the filtration given by $\mathscr{F}_n := \sigma(\varepsilon_1, \dots, \varepsilon_n)$; in fact,

$$\mathbb{E}(f_n|\mathscr{F}_{n-1}) = f_{n-1} + \mathbb{E}(\varepsilon_n|\mathscr{F}_{n-1})x_n = f_{n-1} + (\mathbb{E}\varepsilon_n)x_n = f_{n-1},$$

where the independence of ε_n and $\mathscr{F}_{n-1}$, together with an application of Proposition 2.6.35, was used to identify $\mathbb{E}(\varepsilon_n|\mathscr{F}_{n-1}) = \mathbb{E}\varepsilon_n$.

Random sums like these are martingales of a very special form which play a distinguished role in Analysis in Banach spaces. They will be studied more thoroughly in Volume II. While many of their particular properties are not shared by general martingales, the following basic *contraction principle* foreshadows the important notion of *martingale transforms* that will be taken up in Section 3.5. In the rest of this subsection, the martingale nature of the random sums (3.14) is essentially invisible, it will be explicitly exploited in the following Subsection 3.2.c.

Proposition 3.2.10 (Kahane's contraction principle). *Let* $(\varepsilon_n)_{n=1}^N$ *be a Rademacher sequence on a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$*. For all sequences* $(a_n)_{n=1}^N$ *in* $\mathbb{K}$*, all finite sequences* $(x_n)_{n=1}^N$ *in* X*, and all* $1 \leqslant p \leqslant \infty$ *we have*

$$\Big\| \sum_{n=1}^N a_n \varepsilon_n x_n \Big\|_{L^p(\Omega;X)} \leqslant \max_{1\leqslant n\leqslant N} |a_n| \, \Big\| \sum_{n=1}^N \varepsilon_n x_n \Big\|_{L^p(\Omega;X)}.$$

Moreover, if $(r_n)_{n=1}^N$ *is a real Rademacher sequence on* $(\Omega, \mathscr{F}, \mathbb{P})$ *and* X *is a complex Banach space, then for all sequences* $(a_n)_{n=1}^N$ *in* $\mathbb{C}$ *and* $(x_n)_{n=1}^N$ *in* X *we have*

$$\Big\|\sum_{n=1}^{N}a_n r_n x_n\Big\|_{L^p(\Omega;X)} \leqslant \frac{\pi}{2}\max_{1\leqslant n\leqslant N}|a_n|\,\Big\|\sum_{n=1}^{N}r_n x_n\Big\|_{L^p(\Omega;X)}. \tag{3.15}$$

The constant $\pi/2$ in the second estimate is optimal.

As it turns out, this is an easy consequence of the following non-random version:

Lemma 3.2.11. *For all sequences $(a_n)_{n=1}^N$ in $\mathbb{K}$ and $(x_n)_{n=1}^N$ in X we have*

$$\Big\|\sum_{n=1}^{N}a_n x_n\Big\|_{X} \leqslant \max_{1\leqslant n\leqslant N}|a_n| \sup_{\epsilon\in(S_{\mathbb{K}})^N}\Big\|\sum_{n=1}^{N}\epsilon_n x_n\Big\|_{X}, \tag{3.16}$$

where $S_{\mathbb{K}}=\{z\in\mathbb{K}:|z|=1\}$ and

$$\Big\|\sum_{n=1}^{N}a_n x_n\Big\|_{L^p(\Omega;X)} \leqslant \frac{\pi}{2}\max_{1\leqslant n\leqslant N}|a_n| \sup_{\epsilon\in\{-1,1\}^N}\Big\|\sum_{n=1}^{N}\epsilon_n x_n\Big\|_{L^p(\Omega;X)}. \tag{3.17}$$

Proof of Proposition 3.2.10, assuming Lemma 3.2.11. We apply Lemma 3.2.11 with $L^p(\Omega;X)$ in place of X and $\varepsilon_n x_n$ in (3.16) and $r_n x_n$ in place of x_n in (3.17). Then observe that for all choices of $(\epsilon_n)_{n=1}^N\in(S_{\mathbb{K}})^N$, the sequences $(\epsilon_n\varepsilon_n)_{n=1}^N$ and $(\varepsilon_n)_{n=1}^N$ have equal distribution, and similarly, for all choices of $(\epsilon_n)_{n=1}^N\in\{-1,1\}^N$, the sequences $(\epsilon_n r_n)_{n=1}^N$ and $(r_n)_{n=1}^N$ have equal distribution. Thus, with the choice of quantities that we made, the two inequalities of Lemma 3.2.11 reduce precisely to those of Proposition 3.2.10.

To see the sharpness of $\pi/2$, we take $p=\infty$ (or let $p\to\infty$), $X=\mathbb{C}$, $x_n=e^{2\pi in/N}$ and $a_n=e^{-2\pi in/N}$. On the one hand, we have

$$N=\Big\|\sum_{n=1}^{N}r_n\Big\|_{L^\infty(\Omega)}=\Big\|\sum_{n=1}^{N}a_n x_n r_n\Big\|_{L^\infty(\Omega)}. \tag{3.18}$$

On the other hand,

$$\begin{aligned}\Big\|\sum_{n=1}^{N}x_n r_n\Big\|_{L^\infty(\Omega)} &= \sup_{\epsilon\in\{-1,1\}^N}\Big|\sum_{n=1}^{N}e^{2\pi in/N}\epsilon_n\Big| \\ &= \sup_{t\in[0,1)}\sum_{n=1}^{N}|\cos(2\pi(n/N+t))|,\end{aligned} \tag{3.19}$$

where the last step is obtained as in (3.22). For any t, upon dividing both sides by N, on the right-hand side we recognise a Riemann sum of

$$\int_0^1|\cos(2\pi u)|\,\mathrm{d}u=\frac{2}{\pi}.$$

So indeed we conclude that the ratio of (3.18) and (3.19) tends to $\pi/2$ as $N\to\infty$. □

As we shall see, the case of Lemma 3.2.11 (and hence of Proposition 3.2.10), which involves a fixed scalar field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ throughout (both a_n and ε_n being either real or complex), follows from a relatively simple convexity argument. The hybrid case, involving real r_n but complex a_n, would be a trivial consequence with constant 2 (forget the complex structure of X and consider real and imaginary parts of a_n separately). The surprising fact that the inequality holds with the sharp constant $\pi/2$ is somewhat more subtle; its proof introduces a tool which will also have another application later in this chapter.

As a preparation for the proof, we introduce:

Definition 3.2.12. *Let T be a subset of a vector space V.*

(i) *The* convex hull *of T, denoted by* conv(T)*, is the set of all vectors of the form $\sum_{j=1}^k \lambda_j x_j$ with $x_j \in T$, $\lambda_j \in [0,1]$, $k \geqslant 1$, and $\sum_{j=1}^k \lambda_j = 1$.*
(ii) *The* absolute convex hull *of T, denoted by* abco(T)*, is the set of all vectors of the form $\sum_{j=1}^k \lambda_j x_j$ with $x_j \in T$, $\lambda_j \in \mathbb{K}$, $k \geqslant 1$, and $\sum_{j=1}^k |\lambda_j| \leqslant 1$.*

We shall write $\mathrm{abco}_{\mathbb{K}}(T)$ if we wish to emphasise the choice of scalar field $\mathbb{K}$. Recall the notation for the closed unit ball and the unit sphere

$$\overline{B}_X := \{x \in X : \|x\| \leqslant 1\}, \quad S_X := \{x \in X : \|x\| = 1\}.$$

It is immediate that

$$\mathrm{abco}(T) = \mathrm{conv}(\overline{B}_{\mathbb{K}} T), \quad \text{with } \overline{B}_{\mathbb{K}} T := \{cx : \ c \in \overline{B}_{\mathbb{K}}, \ x \in T\}.$$

This is a useful relation for reducing considerations of the more complicated absolute convex hull to that of the simpler convex hull.

Lemma 3.2.13. *Let $V_1, \dots, V_N$ be vector spaces and let $T_n \subseteq V_n$ for every $1 \leqslant n \leqslant N$. Then*

$$\mathrm{conv}(T_1 \times \cdots \times T_N) = \mathrm{conv}(T_1) \times \cdots \times \mathrm{conv}(T_N).$$

Proof. By iteration it suffices to prove the case $N = 2$. The inclusion '$\subseteq$' is obvious. In the opposite direction let

$$\Big(\sum_{i=1}^m \lambda_i x_i, \sum_{j=1}^n \mu_j y_j\Big) \in \mathrm{conv}(T_1) \times \mathrm{conv}(T_2)$$

with $\sum_{i=1}^m \lambda_i = \sum_{j=1}^n \mu_j = 1$ and $x_1, \dots, x_m \in T_1$, $y_1, \dots, y_n \in T_2$. Since $\sum_{i=1}^m \sum_{j=1}^n \lambda_i \mu_j = 1$ one has

$$\Big(\sum_{i=1}^m \lambda_i x_i, \sum_{j=1}^n \mu_j y_j\Big) = \sum_{i=1}^m \sum_{j=1}^n \lambda_i \mu_j (x_i, y_j) \in \mathrm{conv}(T_1 \times T_2).$$

This proves the inclusion '$\supseteq$'. $\square$

The analogous statement for the absolute convex hulls is false in general: with $V_1 = V_2 = \mathbb{K} = \mathbb{R}$, and $T_i = \{x_i\}$ for $i = 1, 2$, it follows that $\mathrm{abco}(T_1 \times T_2)$ is the closed line segment from $(-x_1, -x_2)$ to (x_1, x_2), whereas $\mathrm{abco}(T_1) \times \mathrm{abco}(T_2)$ is the rectangle $[-x_1, x_1] \times [-x_2, x_2]$.

Proposition 3.2.14. *We have*

$$\begin{aligned} \mathrm{conv}(S_{\mathbb{R}}^n) &= (\overline{B}_{\mathbb{R}})^n = \mathrm{abco}_{\mathbb{R}}(\{-1,1\}^n), \\ \mathrm{conv}(S_{\mathbb{C}}^n) &= (\overline{B}_{\mathbb{C}})^n \subseteq \frac{\pi}{2}\,\mathrm{abco}_{\mathbb{C}}(\{-1,1\}^n). \end{aligned} \tag{3.20}$$

The identity $(\overline{B}_{\mathbb{K}})^n = \mathrm{conv}(S_{\mathbb{K}}^n)$ follows from Lemma 3.2.13 and the immediate observation that $\overline{B}_{\mathbb{K}} = \mathrm{conv}(S_{\mathbb{K}})$.

The identity $\mathrm{conv}(S_{\mathbb{R}}^n) = \mathrm{abco}_{\mathbb{R}}(\{-1,1\}^n)$ in the real case of (3.20) is clear by noting that $\mathrm{conv}(S_{\mathbb{R}}^n)$ is absolutely convex. The complex case of (3.20) lies somewhat deeper. It relies on the following functional analytic observation:

Lemma 3.2.15. *Let X and Y be Banach spaces and $T \in \mathscr{L}(X,Y)$ be bounded from below in the sense that there exists a $\delta > 0$ such that $\|Tx\| \geqslant \delta\|x\|$ for all $x \in X$. Then*

$$\delta \overline{B}_{X^*} \subseteq T^* \overline{B}_{Y^*}.$$

Proof. We may assume that $\dim(X) \geqslant 1$, the result being trivially true in the remaining cases. Let Y_0 denote the range of T, which is a closed subspace of Y. Fix an arbitrary $x_0^* \in X^*$ of norm one and define $y_0^* : Y_0 \to \mathbb{K}$ by $\langle Tx, y_0^*\rangle := \langle x, x_0^*\rangle$. Then $|\langle Tx, y_0^*\rangle| \leqslant \|x\|\,\|x_0^*\| \leqslant \delta^{-1}\|Tx\|$, so $y_0^* \in Y_0^*$ and $\|y_0^*\| \leqslant \delta^{-1}$. By the Hahn–Banach theorem, the functional $y_0^* \in Y_0^*$ has an extension to a functional $y^* \in Y^*$ with $\|y^*\| \leqslant \delta^{-1}$. It satisfies

$$\langle x, T^* y^*\rangle = \langle Tx, y^*\rangle = \langle Tx, y_0^*\rangle = \langle x, x_0^*\rangle$$

for all $x \in X$, and therefore $\delta x_0^* = T^*(\delta y^*) \in T^* \overline{B}_{Y^*}$. □

Complex case of (3.20). In this proof we work over the complex scalar field. Consider the operator

$$T : \ell_n^1 \to \ell^\infty(\{-1,1\}^n), \qquad (Tx)(\epsilon) = \sum_{j=1}^n \epsilon_j x_j,$$

with adjoint

$$T^* : \ell^1(\{-1,1\}^n) \to \ell_n^\infty, \qquad (T^*\lambda)_j = \sum_{\epsilon \in \{-1,1\}^n} \lambda(\epsilon)\epsilon_j.$$

Then (3.20) asserts that $\overline{B}_{\ell_n^\infty} \subseteq \frac{\pi}{2} T^* \overline{B}_{\ell^1(\{-1,1\}^n)}$, which by Lemma 3.2.15 would follow from

$$\|Tx\|_{\ell^\infty(\{-1,1\}^n)} \geqslant \frac{2}{\pi}\|x\|_{\ell_n^1}. \tag{3.21}$$

Let us check (3.21) to complete the proof. Let $x_j = |x_j|e^{2\pi i t_j}$. Then

$$\begin{aligned}
\|Tx\|_{\ell^\infty(\{-1,1\}^n)} &= \sup_{\epsilon\in\{-1,1\}^n} \sup_{t\in[0,1)} \Big|\Re\Big(e^{2\pi i t}\sum_{j=1}^n \epsilon_j x_j\Big)\Big| \\
&= \sup_{t\in[0,1)} \sup_{\epsilon\in\{-1,1\}^n} \Big|\sum_{j=1}^n \epsilon_j |x_j| \cos(2\pi(t_j+t))\Big| \\
&= \sup_{t\in[0,1)} \sum_{j=1}^n |x_j||\cos(2\pi(t_j+t))| \qquad (3.22)\\
&\geqslant \int_0^1 \sum_{j=1}^n |x_j||\cos(2\pi(t_j+t))|\,\mathrm{d}t \\
&= \frac{2}{\pi}\sum_{j=1}^n |x_j| = \frac{2}{\pi}\|x\|_{\ell_n^1}.
\end{aligned}$$

□

Now we have all the required tools for:

Proof of Lemma 3.2.11. We may assume that $\sup_{1\leqslant n\leqslant N}|a_n| = 1$, so that

$$(a_n)_{n=1}^N \in \overline{B}_{\mathbb{K}}^N = \operatorname{conv}(S_{\mathbb{K}}^N) \subseteq \frac{\pi}{2}\operatorname{abco}_{\mathbb{C}}(\{-1,1\}^N) \qquad (3.23)$$

by Proposition 3.2.14.

Case (3.16)*:* By (3.23), there exist $\lambda_j \in [0,1]$ satisfying $\sum_{j=1}^k \lambda_j \leqslant 1$ such that

$$(a_n)_{n=1}^N = \sum_{j=1}^k \lambda_j (\epsilon_n^j)_{n=1}^N,$$

where all $\epsilon_n^j \in \mathbb{K}$ have modulus one. Thus

$$\begin{aligned}
\Big\|\sum_{n=1}^N a_n x_n\Big\|_X &\leqslant \sum_{j=1}^k \lambda_j \Big\|\sum_{n=1}^N \epsilon_n^j x_n\Big\|_X \\
&\leqslant \sum_{j=1}^k \lambda_j \sup_{\epsilon\in(S_{\mathbb{K}})^N}\Big\|\sum_{n=1}^N \epsilon_n x_n\Big\|_X \leqslant \sup_{\epsilon\in(S_{\mathbb{K}})^N}\Big\|\sum_{n=1}^N \epsilon_n x_n\Big\|_X.
\end{aligned}$$

Case (3.17)*:* By (3.23), there exist $\lambda_j \in \mathbb{C}$ satisfying $\sum_{j=1}^k |\lambda_j| \leqslant \pi/2$ such that

$$(a_n)_{n=1}^N = \sum_{j=1}^k \lambda_j (\epsilon_n^j)_{n=1}^N,$$

where all $\epsilon_n^j \in \{-1,1\}$. Thus

$$\Big\|\sum_{n=1}^N a_n x_n\Big\|_X \leqslant \sum_{j=1}^k |\lambda_j| \Big\|\sum_{n=1}^N \epsilon_n^j x_n\Big\|_X \leqslant \sum_{j=1}^k |\lambda_j| \sup_{\epsilon\in\{-1,1\}^N} \Big\|\sum_{n=1}^N \epsilon_n x_n\Big\|_X$$
$$\leqslant \frac{\pi}{2} \sup_{\epsilon\in\{-1,1\}^N} \Big\|\sum_{n=1}^N \epsilon_n x_n\Big\|_X.$$

□

Remark 3.2.16. The sharpness of $\pi/2$ for Proposition 3.2.10 also implies its sharpness for Lemma 3.2.11 and for Proposition 3.2.14.

3.2.c John–Nirenberg and Kahane–Khintchine inequalities

The inequality of John and Nirenberg is an early example of powerful self-improvement phenomena of certain estimates that hold uniformly 'over all scales'. Here we study a martingale version of this inequality, where the 'scales' are formalised in terms of the levels $\mathscr{F}_n$ of the underlying filtration. As a corollary in the context of the random sums studied in Subsection 3.2.b, we prove the Kahane–Khintchine inequality, which says the L^p-norms of these random sums are equivalent for all finite exponents $p \in (0, \infty)$. This tool has widespread applications ranging over different branches of Analysis in Banach spaces, and we shall revisit it with a different approach in Volume II.

Throughout this section we work on a fixed measure space $(S, \mathscr{A}, \mu)$. Let $\phi = (\phi_n)_{n\in\mathbb{Z}}$ be an adapted sequence of functions, not necessarily a martingale. For such sequences, we introduce the following (semi-, quasi-) norms measuring the oscillation between different functions ϕ_n of this sequence:

$$\begin{aligned} \|\phi\|_{\star,q} &:= \sup_{\substack{k,n\in\mathbb{Z}\\ k\leqslant n}} \sup_{F\in\mathscr{F}_k^+} \Big(\fint_F \|\phi_n - \phi_{k-1}\|^q \,\mathrm{d}\mu\Big)^{1/q}, \\ \|\phi\|_{\star\star,q} &:= \sup_{k\in\mathbb{Z}} \sup_{F\in\mathscr{F}_k^+} \Big(\fint_F \sup_{n\geqslant k} \|\phi_n - \phi_{k-1}\|^q \,\mathrm{d}\mu\Big)^{1/q} \\ &=: \sup_{k\in\mathbb{Z}} \sup_{F\in\mathscr{F}_k^+} \Big(\fint_F ({}^{k-1}\phi^\star)^q \,\mathrm{d}\mu\Big)^{1/q}, \end{aligned} \tag{3.24}$$

where

$$\mathscr{F}_k^+ := \{F \in \mathscr{F}_k : 0 < \mu(F) < \infty\}$$

is the sub-collection of $\mathscr{F}_k$ for which the average integral $\fint_F = \frac{1}{|F|}\int_F$ is meaningful, and

$${}^{k-1}\phi^\star := \sup_{n\geqslant k} \|\phi_n - \phi_{k-1}\| = \sup_{n\in\mathbb{Z}} \|{}^{k-1}\phi_n\|$$

is the maximal function of the sequence ϕ started at $k-1$.

The prefix 'semi-' refers to the fact that these 'norms' may vanish even if ϕ is not a zero sequence; indeed, this happens for any sequence ϕ_n that is independent of n. The prefix 'quasi-' refers to the fact that the triangle inequality may hold only up to a constant, which happens for exponents $q \in (0,1)$, a case that we do not exclude, as the possibility of going 'below L^1' is important to some applications. Despite these shortcomings, we shall liberally refer to all $\|\cdot\|_{\star,q}$ and $\|\cdot\|_{\star\star,q}$ simply as 'norms'.

The surprising content of the John–Nirenberg inequality is the following:

Theorem 3.2.17 (John–Nirenberg inequality for adapted sequences). *On the class of adapted sequences $\phi = (\phi_n)_{n\in\mathbb{Z}}$, the norms defined in (3.24) are equivalent for all $q \in (0,\infty)$. More precisely, we have*

$$\|\phi\|_{\star,p} \leqslant \|\phi\|_{\star\star,p} \leqslant 2^{1+1/q} e\Big(1 + 2\frac{\max(p,1)}{q}\Big)\|\phi\|_{\star,q} \quad \forall p,q \in (0,\infty). \tag{3.25}$$

An adapted sequence ϕ for which one and hence all of these norms are finite, is said to be of *bounded mean oscillation*, abbreviated as BMO.

Remark 3.2.18. A reader familiar with the classical theory of the BMO spaces may be inclined to think that the ϕ_{k-1} appearing in (3.24) should rather be replaced by ϕ_k, leading to

$$\|\phi\|'_{\star,q} := \sup_{\substack{k,n\in\mathbb{Z}\\ k\leqslant n}} \sup_{F\in\mathscr{F}_k^+} \Big(\fint_F \|\phi_n - \phi_k\|^q \,\mathrm{d}\mu\Big)^{1/q}, \tag{3.26}$$

and a similar modification $\|\phi\|'_{\star\star,q}$ of $\|\phi\|_{\star\star,q}$. Since

$$\begin{aligned}\Big(\fint_F \|\phi_n - \phi_k\|^q \,\mathrm{d}\mu\Big)^{1/q} &= \Big(\fint_F \|[\phi_n - \phi_{k-1}] - [\phi_k - \phi_{k-1}]\|^q \,\mathrm{d}\mu\Big)^{1/q} \\ &\leqslant 2^{(1/q-1)_+}\Big[\Big(\fint_F \|\phi_n - \phi_{k-1}\|^q \,\mathrm{d}\mu\Big)^{1/q} + \Big(\fint_F \|\phi_k - \phi_{k-1}\|^q \,\mathrm{d}\mu\Big)^{1/q}\Big],\end{aligned}$$

it follows at once that

$$\|\phi\|'_{\star,q} \leqslant 2^{(1/q-1)_+}\|\phi\|_{\star,q}.$$

Conversely, suppose that the filtration is *regular* in the following sense: for some constant $K < \infty$, for every $k \in \mathbb{Z}$ and every $F \in \mathscr{F}_k^+$, there exists $\widehat{F} \in \mathscr{F}_{k-1}^+$ such that $F \subseteq \widehat{F}$ and $\mu(\widehat{F}) \leqslant K\mu(F)$. In this case,

$$\fint_F \|\phi_n - \phi_{k-1}\|^q \,\mathrm{d}\mu \leqslant K \fint_{\widehat{F}} \|\phi_n - \phi_{k-1}\|^q \,\mathrm{d}\mu,$$

from which it follows that

$$\|\phi\|_{\star,q} \leqslant K^{1/q}\|\phi\|'_{\star,q}.$$

Similar estimates can be deduced for "$\star\star$ versions", and hence one could equally well use the norms (3.26) under the regularity assumption.

However, in the generality that we consider (without additional assumptions on the filtration), the definition (3.24) is the 'only correct one', namely the one that supports the John–Nirenberg inequality. Indeed, consider the essentially trivial situation with $\mathscr{F}_k = \{\varnothing, \Omega\}$ for $k \leqslant 0$ and $\mathscr{F}_k = \mathscr{F}$ for $k \geqslant 1$ on a probability space $(\Omega, \mathscr{F}, \mu)$. Moreover, we let $\phi_k = 0$ for $k \leqslant 0$ and $\phi_k = \phi_1$ for $k \geqslant 1$. Considering $\phi_n - \phi_k$, this is only non-zero if $k \leqslant 0 < n$, in which case it is ϕ_1 and $F = \Omega$ is the unique element of $\mathscr{F}_k^+$. Thus it follows that $\|\phi\|'_{\star,q}$ would be simply $\|\phi_1\|_q$ in this case. The prospective John–Nirenberg inequality for this norm would then assert the equivalence of all L^q norms on a probability space, which is absurd. This should serve as a sufficient apology to using ϕ_{k-1} instead of ϕ_k in (3.24).

The first bound in (3.25) is completely elementary, as it amounts to comparing a supremum outside and inside an integral. Note that for $q = p$, the bound (3.25) shows in particular the comparability of $\|\phi\|_{\star,p}$ and $\|\phi\|_{\star\star,p}$, and the constants are uniformly bounded for all $p \in [1, \infty)$. Our proof of Theorem 3.2.17 will feature further oscillatory bounds equivalent to the finiteness of the norms defined in (3.24). To begin with, we record:

Lemma 3.2.19. *If an adapted sequence satisfies $\|\phi\|_{\star,q} < \infty$, then*

$$\mu(F \cap \{\|\phi_n - \phi_{k-1}\| > \alpha'\}) \leqslant \Big(\frac{\|\phi\|_{\star,q}}{\alpha'}\Big)^q \cdot \mu(F) \qquad \forall k \leqslant n,\ F \in \mathscr{F}_k \tag{3.27}$$

for any $\alpha' > 0$.

Proof. We have

$$\mu(F \cap \{\|\phi_n - \phi_{k-1}\| > \alpha'\}) \leqslant \int_F \Big(\frac{\|\phi_n - \phi_{k-1}\|}{\alpha'}\Big)^q \,\mathrm{d}\mu \leqslant \mu(F)\Big(\frac{\|\phi\|_{\star,q}}{\alpha'}\Big)^q.$$

□

The bound (3.27) is the weakest of all oscillatory bounds that we consider, and the rest of the argument consists of subsequent self-improvements of this basic bound. We first establish a stopped version:

Lemma 3.2.20. *If an adapted sequence satisfies*

$$\mu(F \cap \{\|\phi_n - \phi_{k-1}\| > \alpha'\}) \leqslant \eta' \cdot \mu(F) \qquad \forall k \leqslant n,\ F \in \mathscr{F}_k, \tag{3.28}$$

for some $\alpha' > 0$ and $\eta' > 0$, then it also satisfies

$$\mu(F \cap \{\nu < \infty\} \cap \{\|\phi_\nu - \phi_{k-1}\| > \alpha\}) \leqslant \eta \cdot \mu(F) \quad \forall k \in \mathbb{Z},\ F \in \mathscr{F}_k, \tag{3.29}$$

whenever ν is stopping time such that $\nu \geqslant k$ on F, and $\alpha = 2\alpha'$, $\eta = 2\eta'$.

Proof. We first compute

$$\begin{aligned}
&\mu(F\cap\{\nu<\infty\}\cap\{\|\phi_\nu-\phi_{k-1}\|>\alpha\})\\
&\quad=\sum_{n=k}^{\infty}\mu(F\cap\{\nu=n\}\cap\{\|\phi_n-\phi_{k-1}\|>\alpha\})\\
&\quad\leqslant\lim_{N\to\infty}\Big(\sum_{n=k}^{N-1}\mu(F\cap\{\nu=n\}\cap\{\|\phi_n-\phi_N\|>\alpha'\})\\
&\qquad\qquad+\sum_{n=k}^{N-1}\mu(F\cap\{\nu=n\}\cap\{\|\phi_{k-1}-\phi_N\|>\alpha'\})\Big)\\
&\quad\leqslant\lim_{N\to\infty}\sum_{n=k}^{N-1}\eta'\cdot\mu(F\cap\{\nu=n\})+\mu(F\cap\{\|\phi_{k-1}-\phi_N\|>\alpha'\})\\
&\quad\leqslant\eta'\cdot\mu(F)+\eta'\cdot\mu(F)=\eta\cdot\mu(F),
\end{aligned}$$

where we used the assumption (3.28) both as stated, and with $F\cap\{\nu=n\}\in\mathscr{F}_n\subseteq\mathscr{F}_{n+1}$ in place of F and $n+1$ in place of k, as well as the disjointness of the sets $\{\nu=n\}$. ☐

In the following lemma, the maximal function ${}^{k-1}\phi^\star$ appears:

Lemma 3.2.21. *If an adapted sequence satisfies* (3.29) *for some* $\alpha,\eta>0$, *then for all* $\lambda>0$,

$$\mu(F\cap\{{}^{k-1}\phi^\star>\lambda+\alpha\})\leqslant\eta\cdot\mu(F\cap\{{}^{k-1}\phi^\star>\lambda\})\quad\forall k\in\mathbb{Z},\ F\in\mathscr{F}_k. \tag{3.30}$$

Proof. Consider the stopping times

$$\begin{aligned}
\rho&:=\inf\{n\geqslant k:\|\phi_n-\phi_{k-1}\|>\lambda\},\\
\nu&:=\inf\{n\geqslant k:\|\phi_n-\phi_{k-1}\|>\lambda+\alpha\}.
\end{aligned}$$

Then $k\leqslant\rho\leqslant\nu$ and the claim reads as $\mu(F\cap\{\nu<\infty\})\leqslant\eta\cdot\mu(F\cap\{\rho<\infty\})$.

Fix an $n\geqslant k$. On the set $\{\rho=n\leqslant\nu<\infty\}$, we have

$$\|\phi_\nu-\phi_{n-1}\|\geqslant\|\phi_\nu-\phi_{k-1}\|-\|\phi_{n-1}-\phi_{k-1}\|>(\lambda+\alpha)-\lambda=\alpha.$$

Thus, using (3.29) with $k=n$ and $F\cap\{\rho=n\}\in\mathscr{F}_n$ in place of F,

$$\begin{aligned}
&\mu(F\cap\{\rho=n\}\cap\{\nu<\infty\})\\
&\quad=\mu(F\cap\{\rho=n\}\cap\{\nu<\infty\}\cap\{\|\phi_\nu-\phi_{n-1}\|>\alpha\})\\
&\quad\leqslant\eta\cdot\mu(F\cap\{\rho=n\}).
\end{aligned}$$

Now summing over all $n\geqslant 0$, the result follows. ☐

The following lemma shows how to exploit bounds of the form (3.30).

Lemma 3.2.22. *If a non-negative function f, supported on a set $F \in \mathscr{F}$, satisfies*

$$\mu(f > \lambda + \alpha) \leqslant \eta \cdot \mu(f > \lambda) \quad \forall \lambda > 0$$

for some $\eta \in (0,1)$ and some $\alpha > 0$, then it also satisfies

$$\|f\|_p \leqslant \frac{1+\eta^{1/p}}{1-\eta^{1/p}} \alpha \mu(F)^{1/p} \qquad \forall p \in [1,\infty).$$

In particular, if an adapted sequence satisfies (3.29), then this estimate holds with $f = {}^{k-1}\phi^\star \mathbf{1}_F$ for any $F \in \mathscr{F}_k$ and $k \in \mathbb{Z}$.

Proof. We may assume that $\mu(F) < \infty$, for otherwise there is nothing to prove.

If f satisfies the assumption, then so does $f \wedge r$ for every $r > 0$. And if $f \wedge r$ satisfies the conclusion for every $r > 0$, then so does f by monotone convergence. We may hence assume that f is bounded, so that the left-hand side of the claim is *a priori* finite, making an absorption argument available to us at a point below. We compute:

$$\begin{aligned}
\|f\|_p^p &= \int_0^\infty p\lambda^{p-1}\mu(f > \lambda)\,\mathrm{d}\lambda \\
&\leqslant \int_0^\alpha p\lambda^{p-1}\mu(F)\,\mathrm{d}\lambda + \int_0^\infty p(\lambda+\alpha)^{p-1}\mu(f > \lambda+\alpha)\,\mathrm{d}\lambda \\
&\leqslant \alpha^p\mu(F) + \int_0^\infty p(\lambda+\alpha)^{p-1}\eta\cdot\mu(f>\lambda)\,\mathrm{d}\lambda \\
&= \alpha^p\mu(F) + \eta\|f + \alpha\mathbf{1}_F\|_p^p.
\end{aligned}$$

We take the pth roots and use the triangle inequality in L^p to the result that

$$\|f\|_p \leqslant \alpha\mu(F)^{1/p} + \eta^{1/p}\|f\|_p + \eta^{1/p}\alpha\mu(F)^{1/p},$$

and the proof is concluded by absorbing $\eta^{1/p}\|f\|_p$ to the left. □

Proof of Theorem 3.2.17. Combining all lemmas following the statement of Theorem 3.2.17, we find that if $\|\phi\|_{\star,q} < \infty$ for some $q \in (0,\infty)$, then for all $p \in [1,\infty)$, $k \in \mathbb{Z}$, and $F \in \mathscr{F}_k$ we have

$$\|{}^{k-1}\phi^\star\mathbf{1}_F\|_p \leqslant c(\eta,p)\alpha\mu(F)^{1/p}, \quad c(\eta,p) := \frac{1+\eta^{1/p}}{1-\eta^{1/p}}, \tag{3.31}$$

provided that $\alpha = 2\alpha' > 0$ is chosen so that

$$\eta = 2\eta' = 2\Big(\frac{\|\phi\|_{\star,q}}{\alpha'}\Big)^q = 2\Big(\frac{2\|\phi\|_{\star,q}}{\alpha}\Big)^q$$

is smaller than 1, which requires $\alpha > 2^{1+1/q}\|\phi\|_{\star,q}$. Then it follows, by combining (3.31) and the definition of $\|\phi\|_{\star\star,p}$ in (3.24), that

$$\|\phi\|_{\star\star,p} \leqslant c(\eta,p)\alpha = c\big(\eta(\alpha,q),p\big)\alpha, \qquad p \in [1,\infty),$$

and it remains to make an efficient choice of α in this estimate. We choose

$$\alpha = 2^{1+1/q}\|\phi\|_{\star,q}\Big(\frac{p+q}{p}\Big)^{p/q},$$

to give

$$\eta = \Big(\frac{p}{p+q}\Big)^{p}, \qquad c(\eta,p) = \frac{2p+q}{q} = 1 + 2\cdot\frac{p}{q}.$$

Together with the elementary inequality

$$\Big(\frac{p+q}{p}\Big)^{p/q} = \Big(1+\frac{q}{p}\Big)^{p/q} \leqslant e,$$

these combine to give

$$\|\phi\|_{\star\star,p} \leqslant 2^{1+1/q}\cdot e\cdot\Big(1+2\cdot\frac{p}{q}\Big)\|\phi\|_{\star,q}, \qquad \forall p \in [1,\infty).$$

The proof is concluded by observing that for $0 < p < 1$, $\|\phi\|_{\star\star,p} \leqslant \|\phi\|_{\star\star,1}$ simply by Hölder's inequality. □

The Kahane–Khintchine inequality

As an application of the John–Nirenberg inequality, we show that the essence of Khintchine's inequality persists for Rademacher sums with values in an arbitrary Banach space. Our proof also shows that the somewhat complicated-looking norms $\|\phi\|_{\star,q}$ actually admit rather concrete expressions in some applications like the one at hand.

Theorem 3.2.23 (Kahane–Khintchine inequality). *Let X be a Banach space let $(\varepsilon_n)_{n\geqslant 1}$ be a Rademacher sequence defined on a probability space $(\Omega,\mathscr{F},\mathbb{P})$. For all $p,q \in (0,\infty)$ there exists a constant $\kappa_{p,q}$, depending only on the parameters p and q, such that for all choices of $x_1,\dots,x_N \in X$ we have*

$$\Big\|\sum_{n=1}^{N}\varepsilon_n x_n\Big\|_{L^p(\Omega;X)} \leqslant \kappa_{p,q}\Big\|\sum_{n=1}^{N}\varepsilon_n x_n\Big\|_{L^q(\Omega;X)}. \tag{3.32}$$

Proof. We consider the filtration given by $\mathscr{F}_k := \sigma(\{\varepsilon_j : j \leqslant k\})$ and the adapted sequence $\phi_n = \sum_{j\leqslant n}\varepsilon_j x_j$. The main point is computing the norm $\|\phi\|_{\star,q}$. For this purpose, fix numbers $k \leqslant n$ and some $F \in \mathscr{F}_k$. We need to compute

$$\|\mathbf{1}_F\cdot(\phi_n-\phi_{k-1})\|_q^q = \mathbb{E}\Big(\mathbf{1}_F\Big\|\sum_{j=k}^{n}\varepsilon_j x_j\Big\|^q\Big).$$

Since $1 = |\varepsilon_k|^2 = \varepsilon_k\bar{\varepsilon}_k$, we have

$$\Big\|\sum_{j=k}^{n}\varepsilon_j x_j\Big\| = \Big\|\varepsilon_k\Big(x_k + \sum_{j=k+1}^{n}\bar\varepsilon_k\varepsilon_j x_j\Big)\Big\| = \Big\|x_k + \sum_{j=k+1}^{n}\bar\varepsilon_k\varepsilon_j x_j\Big\|. \tag{3.33}$$

The key observation is that the sequence

$$(\varepsilon_j')_{1\leqslant j\leqslant N} := (\varepsilon_1,\dots,\varepsilon_k,\bar\varepsilon_k\varepsilon_{k+1},\dots\bar\varepsilon_k\varepsilon_N)$$

is another Rademacher sequence, and hence

$$\begin{aligned}\mathbb{E}\Big(\mathbf{1}_F\Big\|\sum_{j=k}^{n}\varepsilon_j x_j\Big\|^q\Big) &= \mathbb{E}\Big(\mathbf{1}_F\Big\|x_k + \sum_{j=k+1}^{n}\varepsilon_j' x_j\Big\|^q\Big)\\ &= \mathbb{E}(\mathbf{1}_F)\mathbb{E}\Big\|x_k + \sum_{j=k+1}^{n}\varepsilon_j' x_j\Big\|^q\end{aligned}$$

by independence. Finally, we have $\mathbb{E}(\mathbf{1}_F) = \mathbb{P}(F)$ and

$$\mathbb{E}\Big\|x_k + \sum_{j=k+1}^{n}\varepsilon_j' x_j\Big\|^q = \mathbb{E}\Big\|\sum_{j=k}^{n}\varepsilon_j x_j\Big\|^q$$

by reading the identity (3.33) in the reverse direction. The conclusion of our computation is that

$$\|\mathbf{1}_F(\phi_n - \phi_{k-1})\|_q^q = \mathbb{P}(F)\|\phi_n - \phi_{k-1}\|_q^q \qquad \forall F\in\mathscr{F}_k.$$

With the σ-algebra $\mathscr{F}_{k,n} = \sigma(\{\varepsilon_j : k\leqslant j\leqslant n\})$ we see that for $q\in[1,\infty)$

$$\|\phi_n - \phi_{k-1}\|_q = \Big\|\sum_{j=k}^{n}\varepsilon_j x_j\Big\|_q = \Big\|\mathbb{E}\Big(\sum_{j=1}^{N}\varepsilon_j x_j\Big|\mathscr{F}_{k,n}\Big)\Big\|_q \leqslant \Big\|\sum_{j=1}^{N}\varepsilon_j x_j\Big\|_q.$$

Note that there is equality in the above estimate in the case that $k=1$, $n=N$ and $F=\Omega$. Thus, taking the relevant suprema as in the definition of $\|\phi\|_{\star,q}$, we deduce that

$$\|\phi\|_{\star,q} = \Big\|\sum_{j=1}^{N}\varepsilon_j x_j\Big\|_q, \qquad q\in[1,\infty),$$

from which the estimate (3.32) for $p,q\in[1,\infty)$ is immediate from Theorem 3.2.17.

For $q<1\leqslant p$, we write $1/p = \theta/q + (1-\theta)/2p$ with $\theta = q/(2p-q)$ to compute

$$\|\xi\|_p \leqslant \|\xi\|_q^\theta\|\xi\|_{2p}^{1-\theta} \leqslant \|\xi\|_q^\theta(\kappa_{2p,p}\|\xi\|_p)^{1-\theta}, \quad \xi := \sum_{j=1}^{N}\varepsilon_j x_j,$$

which shows that

$$\|\xi\|_p \leqslant \kappa_{2p,p}^{(1-\theta)/\theta}\|\xi\|_q = \kappa_{2p,p}^{2(p-q)/q}\|\xi\|_q, \qquad 0<q<1\leqslant p<\infty.$$

Finally, the case $p\in(0,1)$ follows trivially from the previous cases simply by starting with $\|\xi\|_p \leqslant \|\xi\|_1$. $\square$

From the previous proof and Theorem 3.2.17, one can extract the bound $\kappa_{p,q} \leqslant cp/q$ for $p \geqslant q \geqslant 1$. By different methods, this may be improved to $c\sqrt{p/q}$, as we shall see in Volume II.

Corollary 3.2.24 (Khintchine's inequality). *Let $(\varepsilon_n)_{n\geqslant 1}$ be a Rademacher sequence defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. For all $p \in (0,\infty)$ there exist constants $0 < A_p \leqslant B_p < \infty$ such that for all finite sequences $(h_n)_{n=1}^N$ in a Hilbert space H,*

$$A_p\Big(\sum_{n=1}^N \|h_n\|^2\Big)^{1/2} \leqslant \Big\|\sum_{n=1}^N h_n\varepsilon_n\Big\|_{L^p(\Omega;H)} \leqslant B_p\Big(\sum_{n=1}^N \|h_n\|^2\Big)^{1/2}. \qquad (3.34)$$

Proof. The independence and mean zero property of the functions ε_n implies their orthogonality in $L^2(\Omega)$, namely

$$\mathbb{E}(\varepsilon_m\overline{\varepsilon}_n) = \mathbb{E}(\varepsilon_m)\mathbb{E}(\overline{\varepsilon}_n) = 0 \cdot 0.$$

For $p = 2$, it follows that

$$\Big\|\sum_{n=1}^N h_n\varepsilon_n\Big\|_{L^2(\Omega;H)}^2 = \sum_{n=1}^N \|h_n\|_H^2,$$

so that (3.34) holds with equality and $A_2 = B_2 = 1$ in this case. The general case then follows from the equivalence of the different L^p-norms of the random sum provided by Theorem 3.2.23. $\square$

3.2.d Applications to inequalities on $\mathbb{R}^d$

The aim of this subsection is to provide a first illustration of the use of martingale methods in the context of classical analysis on the Euclidean space $\mathbb{R}^d$, a topic that will recur in deeper variants later in the book. For the present, we have two main goals in mind: first, to provide an alternative approach to the Hardy–Littlewood maximal operator, which we already studied by more classical methods, and second, to prove the classical version of the John–Nirenberg inequality for BMO functions. In both cases, the bridge from the abstract martingale theory to $\mathbb{R}^d$ will be provided by the *dyadic filtrations* of $\mathbb{R}^d$, and the following definition is basic to the theory:

Definition 3.2.25. *The* standard dyadic system $\mathscr{D}^0$ *in $\mathbb{R}^d$ is the collection of cubes*

$$\mathscr{D}^0 := \bigcup_{j\in\mathbb{Z}} \mathscr{D}_j^0, \qquad \mathscr{D}_j^0 := \{2^{-j}([0,1)^d + m) :\ m \in \mathbb{Z}^d\}.$$

For a vector $\alpha \in \{0, \frac{1}{3}, \frac{2}{3}\}^d$ the shifted dyadic system $\mathscr{D}^\alpha$ *in $\mathbb{R}^d$ is the collection of cubes*

$$\mathscr{D}^\alpha := \bigcup_{j\in\mathbb{Z}} \mathscr{D}_j^\alpha, \qquad \mathscr{D}_j^\alpha := \{2^{-j}([0,1)^d + m + (-1)^j\alpha) :\ m \in \mathbb{Z}^d\}.$$

Note that for any $\alpha \in \{0, \frac{1}{3}, \frac{2}{3}\}^d$, the atomic σ-algebras $\sigma(\mathscr{D}_j^\alpha)$ for $j \in \mathbb{Z}$, form a filtration. Indeed, it is enough to check that $D = 2^{-j}([0,1)^d + m + (-1)^j\alpha) \in \mathscr{D}_j^\alpha$ can be written as the union of cubes in $\mathscr{D}_{j+1}^\alpha$. Because of the grid structure it suffices to argue for each coordinate separately. If $\alpha = 0$, the result is obvious. If $\alpha = \frac{1}{3}$ the result follows from the identity

$$2^{-(j+1)}\big(2m + (-1)^j + (-1)^{j+1}\tfrac{1}{3}\big) = 2^{-j}\big(m + (-1)^j\tfrac{1}{3}\big).$$

The case $\alpha = \frac{2}{3}$ is similar; the $(-1)^j$ on the left-hand side needs to be replaced by $2(-1)^j$.

We shall need the following covering lemma:

Lemma 3.2.26 (Covering lemma). *For every cube $Q = I_1 \times \ldots \times I_d \subseteq \mathbb{R}^d$, where the I_j are finite intervals of equal length $\ell(Q)$, there exist a vector $\alpha \in \{0, \frac{1}{3}, \frac{2}{3}\}^d$ and a dyadic cube $D \in \mathscr{D}^\alpha$ such that*

$$\frac{3}{2}\ell(Q) < \ell(D) \leqslant 3\ell(Q) \quad \textit{and} \quad Q \subseteq D,$$

where $\ell(Q)$ denotes the length of the cube Q.

Proof. First consider the case $d = 1$. Let $I = [a, b)$ with $-\infty < a < b < \infty$. Let $k \in \mathbb{Z}$ be the unique integer such that $\frac{3}{2}\ell(I) < 2^{-k} \leqslant 3\ell(I)$. It suffices to construct an interval $J \in \bigcup_{\alpha \in \{0,\frac{1}{3},\frac{2}{3}\}} \mathscr{D}_k^\alpha$ such that $I \subseteq J$. For $\alpha \in \{0, \frac{1}{3}, \frac{2}{3}\}$ let $E_k^\alpha = \bigcup\{\partial J : J \in \mathscr{D}_k^\alpha\}$ be the set of all endpoints of dyadic intervals in $\mathscr{D}_k^\alpha$. Let $E_k = \bigcup_{\alpha \in \{0,\frac{1}{3},\frac{2}{3}\}^d} E_k^\alpha$. Then $E_k = \{\frac{m}{3}2^{-k} : m \in \mathbb{Z}\}$. Therefore, any two consecutive points in E_k have distance $\frac{1}{3}2^{-k}$. In particular, $E_k \cap I$ consists at most of two points.

Since the sets E_k^α are disjoint for different α, there is at least one α such that $E_k^\alpha = \varnothing$. But then, picking an interval $J \in \mathscr{D}_k^\alpha$ that intersects I, this J in fact must contain I, and the proof in the case $d = 1$ is complete.

Next assume that $d \geqslant 2$. Fix a cube $Q = I_1 \times \ldots \times I_d \subseteq \mathbb{R}^d$ with equal side-lengths $\ell(Q)$. For every $1 \leqslant i \leqslant d$ we can apply Step 1 to find $\alpha_i \in \{0, \frac{1}{3}\}$ and $J_i \in \mathscr{D}^{\alpha_i}$ such that $\ell(I_i) < \ell(J_i) \leqslant 6\ell(I_i)$ and $I_i \subseteq J_i$ hold. Now the dyadic cube $D = J_1 \times \ldots \times J_d$ satisfies the required conditions. $\square$

For the Hardy–Littlewood maximal operator, we here adopt the definition

$$Mf(x) := \sup_{Q \ni x} \fint_Q \|f(y)\| \,\mathrm{d}y, \quad f \in L^1_{\mathrm{loc}}(\mathbb{R}^d; X), \tag{3.35}$$

where the supremum is taken over all cubes Q in $\mathbb{R}^d$ containing the point x and with sides parallel to the coordinate axes; the notation

$$\fint_Q := \frac{1}{|Q|}\int_Q$$

is used for the average over Q. As indicated earlier, this is equivalent to the definition involving balls by at most a multiplicative dimensional constant.

The *shifted dyadic maximal function* is defined by

$$M^\alpha f(x) := \sup_{Q^\alpha \ni x} \fint_{Q^\alpha} \|f(y)\| \,\mathrm{d}y, \qquad f \in L^1_{\mathrm{loc}}(\mathbb{R}^d; X),$$

where the supremum is taken over all dyadic cubes $Q^\alpha \in \mathscr{D}^\alpha$ containing the point x. As a consequence of the covering lemma, for $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d; X)$ we find, pointwise on $\mathbb{R}^d$,

$$Mf \leqslant 3^d \sup_{\alpha \in \{0, \frac13, \frac23\}^d} M^\alpha f \leqslant 3^d \sum_{\alpha \in \{0, \frac13, \frac23\}^d} M^\alpha f. \tag{3.36}$$

Accordingly it suffices to prove boundedness properties for each M^α.

The operators M^α have the following martingale interpretation. For $\alpha \in \{0, \frac13, \frac23\}^d$ and $j \in \mathbb{Z}$ let $\mathscr{F}_j^\alpha = \sigma(\mathscr{D}_j^\alpha)$. Introducing, for $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d; X)$, the non-negative martingale

$$f_j^\alpha = \mathbb{E}(\|f\| | \mathscr{F}_j^\alpha) = \sum_{Q \in \mathscr{D}_j^\alpha} \mathbf{1}_Q \fint_Q \|f(y)\| \,\mathrm{d}y, \quad j \in \mathbb{Z},$$

one sees that

$$M^\alpha f = (f^\alpha)^\star.$$

As a consequence, boundedness properties of M^α can be deduced from Doob's maximal inequalities. Indeed, let $w \in L^0(S)$ be a non-negative function. Recall that for $A \in \mathscr{A}$ we write $w(A) := \int_A w \,\mathrm{d}\mu$. From (3.9) and (3.10), we immediately obtain:

(1) If $f \in L^1(\mathbb{R}^d, M^\alpha w; X)$, then for all $r > 0$,

$$r\, w\big(\{M^\alpha f > r\}\big) \leqslant \|f\|_{L^1(\mathbb{R}^d, Mw; X)}.$$

(2) If $f \in L^p(\mathbb{R}^d, M^\alpha w; X)$ with $1 < p \leqslant \infty$, then

$$\|M^\alpha f\|_{L^p(\mathbb{R}^d, w)} \leqslant p' \|f\|_{L^p(\mathbb{R}^d, Mw; X)},$$

where $\frac1p + \frac1{p'} = 1$.

Combining these estimates with (3.36) yields the following version of the Hardy–Littlewood maximal Theorem 2.3.2:

Theorem 3.2.27 (Hardy–Littlewood maximal theorem revisited). *Let M be the Hardy–Littlewood maximal function as defined in* (3.35) *and let $w \in L^0(\mathbb{R}^d)$ be non-negative.*

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

$$r\, w\big(\{Mf > r\}\big) \leqslant 9^d \|f\|_{L^1(\mathbb{R}^d, Mw; X)}.$$

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

$$\|Mf\|_{L^p(\mathbb{R}^d,w)} \leqslant 9^d p' \|f\|_{L^p(\mathbb{R}^d,Mw;X)},$$

where $\frac{1}{p} + \frac{1}{p'} = 1$.

In particular, for $w \equiv 1$ one has $Mw \equiv 1$ and we obtain the classical Hardy–Littlewood maximal inequalities.

As a more substantial application of the dyadic techniques, we show next that for $X = \ell^q$ we can give more refined results. With a slight abuse of notation, for a function $f \in L^1_{\rm loc}(\mathbb{R}^d; \ell^q)$ we define Mf coordinatewise by

$$(Mf)^{(j)}(x) = Mf^{(j)}(x).$$

Theorem 3.2.28 (Fefferman–Stein). *Let* $p \in (1, \infty)$ *and* $q \in (1, \infty]$. *With the above notation, for all* $f \in L^p(\mathbb{R}^d; \ell^q)$ *we have* $Mf \in L^p(\mathbb{R}^d; \ell^q)$ *and*

$$\|Mf\|_{L^p(\mathbb{R}^d;\ell^q)} \leqslant C_{p,q,d} \|f\|_{L^p(\mathbb{R}^d;\ell^q)}.$$

Proof. By (3.36) it suffices to prove the boundedness of the shifted dyadic maximal functions M^α. This is immediate from Theorem 3.2.7 applied with the dyadic filtration $(\mathscr{F}_n^\alpha)_{n\in\mathbb{Z}}$ and to the submartingale $(\mathbb{E}(|f^{(j)}||\mathscr{F}_n^\alpha))_{n\in\mathbb{Z}}$. □

The above result does not hold for $p = \infty$ and $q \in (1, \infty)$, as is shown by the following example.

Example 3.2.29. Fix $q \in (1, \infty)$ and let $f \in L^\infty(\mathbb{R}^d; \ell^q)$ be defined by $f = \sum_{j=1}^\infty \mathbf{1}_{(2^{-j},2^{-j+1}]} e_j$. Then $\|f(x)\|_{\ell^q} = \mathbf{1}_{(0,1]}(x)$ for all $x \in \mathbb{R}^d$ and consequently $\|f\|_{L^\infty(\mathbb{R}^d;\ell^q)} = 1$. On the other hand, for all $t \leqslant a < b$ we have $M\mathbf{1}_{(a,b]}(t) = \frac{b-a}{b-t}$, and therefore

$$Mf(t) = \sum_{j=1}^\infty \frac{2^{-j+1} - 2^{-j}}{2^{-j+1} - t} e_j = \sum_{j=1}^\infty \frac{1}{2 - 2^j t} e_j, \quad t \leqslant 0.$$

Now, for fixed $j_0 \geqslant 1$ and all $t \in [-2^{-j_0}, 0]$, we find

$$\|Mf(t)\|_{\ell^q} \geqslant \frac{1}{3}\Big\| \sum_{j=1}^{j_0} e_j \Big\|_{\ell^q} = \frac{1}{3} j_0^{1/q}.$$

This implies that $Mf \notin L^\infty(\mathbb{R}^d; \ell^q)$.

The BMO *space and the classical John–Nirenberg inequality*

For a function $f \in L^1_{\rm loc}(\mathbb{R}^d; X)$ and exponents $p \in [1, \infty)$, consider the BMO (semi-)norms

$$\|f\|_{\mathrm{BMO}^p(\mathbb{R}^d;X)} := \sup_Q \Big(\fint_Q \|f(x) - \langle f\rangle_Q\|_X^p \,\mathrm{d}x \Big)^{1/p}, \tag{3.37}$$

where $\langle f\rangle_Q := \fint_Q f(x)\,\mathrm{d}x$.

Theorem 3.2.30 (John–Nirenberg inequality for functions on $\mathbb{R}^d$). *For $p \in [1,\infty)$, all* BMO *norms defined in* (3.37) *are equivalent, and in fact*

$$\|f\|_{\mathrm{BMO}^q(\mathbb{R}^d;X)} \leqslant \|f\|_{\mathrm{BMO}^p(\mathbb{R}^d;X)} \leqslant 8\cdot 6^d \cdot e \cdot \Big(1+\frac{2p}{q}\Big)\|f\|_{\mathrm{BMO}^q(\mathbb{R}^d;X)},$$
$$\textit{for}\quad 1\leqslant q\leqslant p<\infty.$$

Proof. The first estimate is immediate from Hölder's inequality. We will prove the second one by reducing it to the adapted version of Theorem 3.2.17.

For every $\alpha \in \{0,\frac13,\frac23\}^d$, we consider the norm

$$\|f\|_{\mathrm{BMO}^p_\alpha(\mathbb{R}^d;X)} := \sup_{Q\in\mathscr{D}^\alpha}\Big(\fint_Q \|f(x)-\langle f\rangle_{\widehat{Q}^\alpha}\|_X^p\,\mathrm{d}x\Big)^{1/p},$$

where $\widehat{Q}^\alpha$ is the dyadic parent of Q in $\mathscr{D}^\alpha$: the unique $R\in\mathscr{D}^\alpha$ with $R\supseteq Q$ and $\ell(R)=2\ell(Q)$.

Since

$$\Big(\fint_Q \|f(x)-\langle f\rangle_{\widehat{Q}^\alpha}\|_X^p\,\mathrm{d}x\Big)^{1/p} \leqslant \Big(2^d\fint_{\widehat{Q}^\alpha} \|f(x)-\langle f\rangle_{\widehat{Q}^\alpha}\|_X^p\,\mathrm{d}x\Big)^{1/p},$$

it is immediate that

$$\|f\|_{\mathrm{BMO}^p_\alpha} \leqslant 2^{d/p}\|f\|_{\mathrm{BMO}^p}.$$

On the other hand, given any cube Q, for some α we can find a dyadic $D\in\mathscr{D}^\alpha$ containing Q, with $\ell(D)\leqslant 3\ell(Q)$, and then, with any constant $c\in X$,

$$\begin{aligned}\Big(\fint_Q \|f-\langle f\rangle_Q\|_X^p\,\mathrm{d}x\Big)^{1/p} &= \Big(\fint_Q \|(f-c)-\langle f-c\rangle_Q\|_X^p\,\mathrm{d}x\Big)^{1/p}\\ &\leqslant 2\Big(\fint_Q \|f-c\|_X^p\,\mathrm{d}x\Big)^{1/p} \leqslant 2\Big(3^d\fint_D \|f-c\|_X^p\,\mathrm{d}x\Big)^{1/p}.\end{aligned}$$

Choosing $c=\langle f\rangle_{\widehat{D}^\alpha}$ this shows that

$$\|f\|_{\mathrm{BMO}^p} \leqslant 2\cdot 3^{d/p}\sup_\alpha \|f\|_{\mathrm{BMO}^p_\alpha}.$$

The next step is to identify BMO^p_α with the adapted BMO spaces considered previously. To this end, let

$$f^\alpha := (f^\alpha_n)_{n\in\mathbb{Z}} := (\mathbb{E}(f|\mathscr{F}^\alpha_n))_{n\in\mathbb{Z}}, \qquad \mathscr{F}^\alpha_k := \sigma(\mathscr{D}^\alpha_k),$$

be the martingale generated by the function f and the filtration corresponding to the dyadic system $\mathscr{D}^\alpha$. We claim that

$$\|f\|_{\mathrm{BMO}^p_\alpha} = \|f^\alpha\|_{\star,p},$$

where the right hand side is defined as in (3.24). Note that for $x\in Q\in\mathscr{D}^\alpha_k$, we have $f^\alpha_{k-1}(x)=\langle f\rangle_{\widehat{Q}^\alpha}$.

Since $f = \lim_{n\to\infty} f_n^\alpha$ almost everywhere by Lebesgue's differentiation theorem, it follows from Fatou's lemma that, for $Q \in \mathscr{D}_k^\alpha \subseteq (\mathscr{F}_k^\alpha)^+$,

$$\begin{aligned}\fint_Q \|f - \langle f\rangle_{\widehat{Q}^\alpha}\|_X^p \,\mathrm{d}x &= \fint_Q \lim_{n\to\infty} \|f_n^\alpha - f_{k-1}^\alpha\|_X^p \,\mathrm{d}x \\ &\leqslant \liminf_{n\to\infty} \fint_Q \|f_n^\alpha - f_{k-1}^\alpha\|_X^p \,\mathrm{d}x,\end{aligned}$$

from which it follows that $\|f\|_{\mathrm{BMO}_\alpha^p} \leqslant \|f^\alpha\|_{\star,p}$. On the other hand, any $F \in (\mathscr{F}_k^\alpha)^+$ is a disjoint union of some $Q_i \in \mathscr{D}_k^\alpha$, and hence for $n \geqslant k$,

$$\begin{aligned}\fint_F \|f_n^\alpha - f_{k-1}^\alpha\|_X^p \,\mathrm{d}x &= \frac{1}{|F|}\int \|\mathbb{E}(\mathbf{1}_F[f - f_{k-1}^\alpha]|\mathscr{F}_n^\alpha)\|_X^p \,\mathrm{d}x \\ &\leqslant \frac{1}{|F|}\int_F \|f - f_{k-1}^\alpha\|_X^p \,\mathrm{d}x \\ &= \frac{1}{|F|}\sum_i \int_{Q_i} \|f - \langle f\rangle_{\widehat{Q}_i^\alpha}\|_X^p \,\mathrm{d}x \\ &\leqslant \frac{1}{|F|}\sum_i |Q_i| \|f\|_{\mathrm{BMO}_\alpha^p}^p,\end{aligned}$$

from which it follows that $\|f^\alpha\|_{\star,p} \leqslant \|f\|_{\mathrm{BMO}_\alpha^p}$.

The proof is completed by combining the previous estimates with an application of Theorem 3.2.17 to the norms $\|f^\alpha\|_{\star,p}$:

$$\begin{aligned}\|f\|_{\mathrm{BMO}^p(\mathbb{R}^d;X)} &\leqslant 2\cdot 3^{d/p}\sup_\alpha \|f\|_{\mathrm{BMO}_\alpha^p(\mathbb{R}^d;X)} = 2\cdot 3^{d/p}\sup_\alpha \|f^\alpha\|_{\star,p} \\ &\leqslant 2\cdot 3^{d/p}\sup_\alpha 2^{1+1/q} e\Big(1 + \frac{2p}{q}\Big)\|f^\alpha\|_{\star,q} \\ &= 4\cdot 2^{1/q}\cdot 3^{d/p} e\Big(1 + \frac{2p}{q}\Big)\sup_\alpha \|f\|_{\mathrm{BMO}_\alpha^q(\mathbb{R}^d;X)} \\ &\leqslant 4\cdot 2^{1/q}\cdot 3^{d/p} e\Big(1 + \frac{2p}{q}\Big)2^{d/q}\|f\|_{\mathrm{BMO}^q(\mathbb{R}^d;X)} \\ &\leqslant 8\cdot 6^d\cdot e\Big(1 + \frac{2p}{q}\Big)\|f\|_{\mathrm{BMO}^q(\mathbb{R}^d;X)}.\end{aligned}$$

□

As in Theorem 3.2.17, a version of Theorem 3.2.30 is also valid for $p, q \in (0,\infty)$. In this case, the reduction to Theorem 3.2.17 would be somewhat more complicated, since the auxiliary adapted sequences would no longer be martingales.

The quantitative comparison of L^p-based BMO norms, as expressed in Theorem 3.2.30, also implies the following exponential integrability, which often goes under the name of John–Nirenberg inequality:

Corollary 3.2.31. *For* $0 < \epsilon \leqslant \left(48 \cdot 6^d \cdot e \cdot \|f\|_{\mathrm{BMO}^1(\mathbb{R}^d;X)}\right)^{-1}$, *we have*

$$\fint_Q \exp\Big(\epsilon \|f - \langle f\rangle_Q\|_X\Big)\,\mathrm{d}x \leqslant 2.$$

Proof. By expanding the exponential, we have

$$\begin{aligned}\fint_Q \exp\Big(\epsilon \|f - \langle f\rangle_Q\|_X\Big)\,\mathrm{d}x &= \fint_Q \sum_{k=0}^{\infty} \frac{\epsilon^k}{k!}\|f - \langle f\rangle_Q\|_X^k\,\mathrm{d}x \\ &\leqslant 1 + \sum_{k=1}^{\infty} \frac{\epsilon^k}{k!}\|f\|_{\mathrm{BMO}^k}^k.\end{aligned}$$

Using Theorem 3.2.30 with $p = k$, $q = 1$, and noting that $1 + 2k \leqslant 3k$ for $k \geqslant 1$, we have

$$\frac{\epsilon^k}{k!}(24 \cdot 6^d \cdot k \cdot \|f\|_{\mathrm{BMO}^1})^k \leqslant (\epsilon \cdot 24 \cdot 6^d \cdot e \cdot \|f\|_{\mathrm{BMO}^1})^k \leqslant 2^{-k},$$

using the elementary estimate $k^k/k! \leqslant e^k$ and the assumption on ϵ. The result follows by summing $\sum_{k=1}^{\infty} 2^{-k} = 1$. $\square$

3.3 Martingale convergence

Having defined a martingale as a sequence $(f_n)_{n\in\mathbb{Z}}$ indexed by integers to begin with, in this section we pursue a detailed study of conditions under which the end-point values f_∞ and $f_{-\infty}$ can be meaningfully augmented to the sequence. Among other applications, this will allow a more flexible use of the notion of the stopped sequence f_τ in the (not uncommon) situations where the stopping time τ takes values in the extended integers $\overline{\mathbb{Z}}$.

The relevant measure-theoretic setting for the end-point values $f_{\pm\infty}$ is provided by the extremal σ-algebras of this filtration, namely

$$\mathscr{F}_{-\infty} := \bigcap_{n\in\mathbb{Z}} \mathscr{F}_n, \qquad \mathscr{F}_\infty := \sigma\Big(\bigcup_{n\in\mathbb{Z}} \mathscr{F}_n\Big).$$

Notice that the intersection of σ-algebras is another σ-algebra, while the union may fail to be one; that is why we define $\mathscr{F}_\infty$ as the σ-algebra generated by the union.

Unless stated otherwise, X is an arbitrary Banach space and $(S, \mathscr{A}, \mu)$ is a measure space endowed with a σ-finite filtration $(\mathscr{F}_n)_{n\in\mathbb{Z}}$ throughout this section.

3.3.a Forward convergence

If $f_n = \mathbb{E}(f|\mathscr{F}_n)$ is the martingale generated by a function $f \in L^0(S;X)$ that is σ-integrable over every $\mathscr{F}_n$, the identity

$$\mathbb{E}(f|\mathscr{F}_n) = \mathbb{E}(\mathbb{E}(f|\mathscr{F}_\infty)|\mathscr{F}_n)$$

shows that this martingale is already generated by $\mathbb{E}(f|\mathscr{F}_\infty)$. In the converse direction, we now investigate how to recover the generating function $\mathbb{E}(f|\mathscr{F}_\infty)$ from the mere knowledge of the martingale $(f_n)_{n\in\mathbb{Z}}$. More elaborate convergence theorems, which produce a generating function for a martingale without assuming its *a priori* existence, typically require additional assumptions on the Banach space X, and will be dealt with in later sections.

Lemma 3.3.1. *For $p \in [1,\infty)$, $\bigcup_{n\in\mathbb{Z}} L^p(S,\mathscr{F}_n;X)$ is dense in $L^p(S,\mathscr{F}_\infty;X)$.*

Proof. Fix a function $f \in L^p(S,\mathscr{F}_\infty;X)$ and let $\varepsilon > 0$ be arbitrary. Since $\mathscr{F}_0$ is σ-finite we can choose a set $A \in \mathscr{F}_0$ of finite μ-measure such that $\|f - \mathbf{1}_A f\|_p < \varepsilon$. In view of $\mathbb{E}(\mathbf{1}_A f|\mathscr{F}_n) = \mathbf{1}_A\mathbb{E}(f|\mathscr{F}_n)$ for $n \in \mathbb{N}\cup\{\infty\}$, in the rest of the proof we may assume that the measure space $(S,\mathscr{A},\mu)$ is finite.

By Lemma A.1.2, applied to the algebra $\bigcup_{n\geqslant 0}\mathscr{F}_n$, for all $B \in \mathscr{F}_\infty$ we have $\mathbf{1}_B = \lim_{k\to\infty}\mathbf{1}_{B_k}$ in $L^p(S)$, where $B_k \in \mathscr{F}_{n_k}$ for suitable $n_k \geqslant 1$. It follows that every simple function of $L^p(S,\mathscr{F}_\infty;X)$ is contained in the closure of $\bigcup_{n\geqslant 1} L^p(S,\mathscr{F}_n;X)$ in $L^p(S,\mathscr{F}_\infty;X)$. As a consequence, all of $L^p(S,\mathscr{F}_\infty;X)$ is contained in the closure of $\bigcup_{n\geqslant 1} L^p(S,\mathscr{F}_n;X)$. □

Theorem 3.3.2 (Forward convergence of generated martingales). *Let $(\mathscr{F}_n)_{n\in\mathbb{Z}}$ be a σ-finite filtration of the measure space $(S,\mathscr{A},\mu)$.*

(1) *If $f \in L^0(S;X)$ is σ-integrable over $\mathscr{F}_n$ for all (or just sufficiently large) $n \in \mathbb{Z}$, then*
$$\lim_{n\to\infty}\mathbb{E}(f|\mathscr{F}_n) = \mathbb{E}(f|\mathscr{F}_\infty)$$
almost everywhere

(2) *If $f \in L^p(S;X)$ for some $p \in [1,\infty)$, then*
$$\lim_{n\to\infty}\mathbb{E}(f|\mathscr{F}_n) = \mathbb{E}(f|\mathscr{F}_\infty)$$
in the norm of $L^p(S;X)$.

Proof. Without loss of generality, in both (1) and (2) we may assume that $f = \mathbb{E}(f|\mathscr{F}_\infty)$.

Let us first prove part (2). Let $f = \mathbb{E}(f|\mathscr{F}_\infty) \in L^p(S,\mathscr{F}_\infty;X)$ be given and let $\varepsilon > 0$ be arbitrary. By Lemma 3.3.1 we can find an $N \in \mathbb{Z}$ and $g \in L^p(S,\mathscr{F}_N;X)$ such that $\|f-g\|_p < \varepsilon/2$. Then $\mathbb{E}(g|\mathscr{F}_n) = g$ for all $n \geqslant N$. Thus, for all $n \geqslant N$,

$$\|f - \mathbb{E}(f|\mathscr{F}_n)\|_p = \|f - g + \mathbb{E}(g-f|\mathscr{F}_n)\|_p \leqslant 2\|f-g\|_p \leqslant \varepsilon,$$

and (2) follows.

To prove part (1) we first assume that $f = \mathbb{E}(f|\mathscr{F}_\infty) \in L^1(S, \mathscr{F}_\infty; X)$. Let again $\varepsilon > 0$ be arbitrary and choose g as in the proof of part (2) (with $p = 1$). We have

$$\begin{aligned}\limsup_{n\to\infty} \|f - \mathbb{E}(f|\mathscr{F}_n)\| &= \limsup_{n\to\infty} \|f - g + \mathbb{E}(g - f|\mathscr{F}_n)\| \\ &\leqslant \|f - g\| + (g - f)^\star,\end{aligned}$$

where the last term is the Doob maximal function. Using Doob's weak type inequality to bound this term, we find that

$$\begin{aligned}\mu(\limsup_{n\to\infty} \|f - \mathbb{E}(f|\mathscr{F}_n)\| > \delta) &\leqslant \mu(\|f - g\| > \delta/2) + \mu((f - g)^\star > \delta/2) \\ &\leqslant \frac{2}{\delta}\|f - g\|_1 + \frac{2}{\delta}\|f - g\|_1 \leqslant \frac{2\varepsilon}{\delta}.\end{aligned}$$

The bound is valid for any $\varepsilon > 0$, so the left-hand side must vanish. Since this is true for any $\delta > 0$, it follows that $\limsup_{n\to\infty} \|f - \mathbb{E}(f|\mathscr{F}_n)\| = 0$ almost everywhere, that is, $\mathbb{E}(f|\mathscr{F}_n) \to f$ almost everywhere. This proves (1) in the special case that $f \in L^1(S, \mathscr{F}_\infty; X)$.

To prove the general case, recall that σ-integrability of $f = \mathbb{E}(f|\mathscr{F}_\infty) \in L^0(S, \mathscr{F}_\infty; X)$ over $\mathscr{F}_N$ means the existence of an exhausting sequence of sets $F_i \in \mathscr{F}_N$ such that $\mathbf{1}_{F_i} f \in L^1(S, \mathscr{F}_\infty; X)$ (see Definition 2.6.1). By the case already proved, we have $\mathbf{1}_{F_i}\mathbb{E}(f|\mathscr{F}_n) = \mathbb{E}(\mathbf{1}_{F_i} f|\mathscr{F}_n) \to \mathbf{1}_{F_i} f$ almost everywhere as $N \leqslant n \to \infty$. Thus $\mathbb{E}(f|\mathscr{F}_n) \to f$ almost everywhere on F_i and then, by exhaustion, almost everywhere on S. This completes the proof of (1). $\square$

Corollary 3.3.3. *Let $p \in [1, \infty)$ be given and let $(f_n)_{n\geqslant 0}$ be a martingale in $L^p(S; X)$ adapted to a σ-finite filtration $(\mathscr{F}_n)_{n\geqslant 0}$. Then $(f_n)_{n\geqslant 0}$ converges in $L^p(S; X)$ if and only if there exists an $f \in L^p(S; X)$ such that $f_n = \mathbb{E}(f|\mathscr{F}_n)$ for all $n \geqslant 0$. In this situation we have $\lim_{n\to\infty} f_n = \mathbb{E}(f|\mathscr{F}_\infty)$ in $L^p(S; X)$.*

Proof. The "if" part follows from Theorem 3.3.2. For the "only if" part let $f = \lim_{n\to\infty} f_n$ in $L^p(S; X)$. Then for all $A \in \mathscr{F}_n$ with $\mu(A) < \infty$ one has

$$\int_A f \,\mathrm{d}\mu = \lim_{m\to\infty} \int_A f_m \,\mathrm{d}\mu = \int_A f_n \,\mathrm{d}\mu,$$

which proves that $f_n = \mathbb{E}(f|\mathscr{F}_n)$. $\square$

3.3.b Backward convergence

Our next theorem is concerned with the backward convergence of two-sided martingales $(f_n)_{n\in\mathbb{Z}}$. Its proof relies on an elementary Hilbert space lemma.

Lemma 3.3.4. *Let $(H_n)_{n\geqslant 1}$ be a decreasing sequence of closed subspaces of a Hilbert space H_0 and put $H_\infty = \bigcap_{n\geqslant 1} H_n$. For each $n \in \mathbb{N} \cup \{\infty\}$ let $P_n : H_0 \to H_n$ denote the orthogonal projection onto H_n. Then for all $h \in H_0$ we have $\lim_{n\to\infty} P_n h = P_\infty h$.*

Proof. Let $H_n' := H_n \ominus H_{n+1}$ be the orthogonal complement of H_{n+1} in H_n. By iteration, it follows that $H_m = \bigoplus_{k=m}^n H_k' \oplus H_{n+1}$ for $m \leqslant n$. Let P_n' be the orthogonal projection of H_0 onto H_n'. Fix $h \in H_0$. Since the projections $P_{n'}$ are pairwise orthogonal, it follows that $\sum_{k=0}^\infty \|P_k' h\|^2 \leqslant \|h\|^2$, so in particular the series converges, and hence

$$\|P_m h - P_n h\|^2 = \Big\| \sum_{k=m}^n P_k' h \Big\|^2 = \sum_{k=m}^n \|P_k' h\|^2 \to 0.$$

By Cauchy's criterion, we have $\lim_{n\to\infty} P_n h \to y$ for some $y \in H_0$. From the containment of the spaces, we have $P_n P_m = P_m P_n = P_{\max(m,n)}$. Thus $y = \lim_{n\to\infty} P_n h = \lim_{n\to\infty} P_m P_n h$, so that $y \in H_m$ for all m, and thus $y \in H_\infty$. A similar reasoning then gives that $\lim_{n\to\infty} P_n h = y = P_\infty y = \lim_{n\to\infty} P_\infty P_n h = \lim_{n\to\infty} P_\infty h = P_\infty h$, and we are done. □

Theorem 3.3.5 (Backward convergence of martingales). *Suppose that $(f_n)_{n\in\mathbb{Z}}$ is an X-valued martingale adapted to a σ-finite filtration $(\mathscr{F}_n)_{n\in\mathbb{Z}}$.*

(1) *If μ is σ-finite on $\mathscr{F}_{-\infty}$ and each f_m is σ-integrable on $\mathscr{F}_{-\infty}$, then for all $m \in \mathbb{Z}$*

$$\lim_{n\to-\infty} f_n = \mathbb{E}(f_m|\mathscr{F}_{-\infty}) =: f_{-\infty} \quad \textit{almost everywhere.}$$

Furthermore, for any $p \in [1,\infty)$ the following assertions hold:

(2) *If μ is σ-finite on $\mathscr{F}_{-\infty}$ and each f_m belongs to $L^p(S;X)$, then for all $m \in \mathbb{Z}$*
$$\lim_{n\to-\infty} f_n = \mathbb{E}(f_m|\mathscr{F}_{-\infty}) =: f_{-\infty} \quad \textit{in } L^p(S;X).$$

(3) *If μ is purely infinite on $\mathscr{F}_{-\infty}$ and each f_m belongs to $L^p(S;X)$, then*

$$\lim_{n\to-\infty} f_n = 0 \quad \textit{almost everywhere.}$$

For $p \in (1,\infty)$ we also have convergence in the $L^p(S;X)$-norm.

Note that the conditional expectations in (1) and (2) exist by Theorem 2.6.20 and Corollary 2.6.30, respectively. Recall that a measure μ is *purely infinite* if it takes values in $\{0,\infty\}$.

The L^p-convergence in (3) is wrong for $p = 1$, as is evidenced by Remark 2.6.40. In the following example we show that the almost everywhere convergence in (3) fails for $p = \infty$ and $p = 0$.

Example 3.3.6. Let $S = [0,\infty)$ be the half-line endowed with the Lebesgue measure, and let $\mathscr{F}_n = \sigma([2^{-n}k, 2^{-n}(k+1)) : k \in \mathbb{N})$, $n \in \mathbb{Z}$. The Lebesgue measure is purely infinite on $\mathscr{F}_{-\infty} = \{\varnothing, [0,\infty)\}$.

Define $f \in L^\infty(0,\infty)$ by $f := (-1)^m$ on $[2^m, 2^{m+1})$ for $m = 0,1,2,\dots$ and $f := 0$ on $[0,1)$. For $n = 0,1,2,\dots$ one easily checks that $\mathbb{E}(f|\mathscr{F}_{-n}) = f$ on $[2^n,\infty)$. We claim that for $n = 1,2,\dots,$

$$\frac{1}{4} \leqslant (-1)^{n-1}\mathbb{E}(f|\mathscr{F}_{-n}) \leqslant \frac{1}{2} \ \text{ on } [0,2^n).$$

This follows by induction. Indeed, for $n = 1$ this follows from the identity $⨍_0^2 f(t)\,\mathrm{d}t = \frac{1}{2}$. Now assume the statement holds for some $n \geqslant 1$. From

$$A := (-1)^n ⨍_0^{2^n} f(t)\,\mathrm{d}t \in [-\tfrac{1}{2}, -\tfrac{1}{4}] \ \text{ and } \ B := (-1)^n ⨍_{2^n}^{2^{n+1}} f(t)\,\mathrm{d}t = 1,$$

we infer that $\frac{1}{2}(A+B) \in [\frac{1}{4}, \frac{1}{2}]$, and the claim follows. As a consequence, $(\mathbb{E}(f|\mathscr{F}_{-n}))_{n\geqslant 0}$ diverges almost everywhere as $n \to \infty$.

Proof of Theorem 3.3.5. Given any fixed $m \in \mathbb{Z}$, the limits in question depend only on $n \leqslant m$, and for these n the martingale property shows that $f_n = \mathbb{E}(f_m|\mathscr{F}_n)$. The tower property of conditional expectations shows that $\mathbb{E}(f_m|\mathscr{F}_{-\infty})$ is independent of m. Hence, without loss of generality, we may assume that $f_n = \mathbb{E}(f|\mathscr{F}_n)$ is generated by a function $f \in L^0(S;X)$ (in (1)), respectively by a function $f \in L^p(S;X)$ (in (2) and (3)); indeed, we may take $f = f_m$ for any fixed $m \in \mathbb{Z}$ and only consider $n \leqslant m$ in the sequel.

We begin with the proof of (2). Let first $f \in L^2(S)$ be a scalar-valued function. Let $H_n := L^2(S, \mathscr{F}_{m-n})$ and $P_n := \mathbb{E}(\cdot|\mathscr{F}_{m-n})$ for $n \in \mathbb{N}\cup\{\infty\}$. By Lemma 2.6.21, P_n is the orthogonal projection onto H_n, so that the L^2-convergence of $P_n f$ to $P_\infty f$ follows from Lemma 3.3.4.

We next consider the L^p-convergence for $p \in [1,\infty)$. By an approximation argument and the contractivity of the conditional expectation it suffices to consider μ-simple functions f. Then, by linearity and the fact that $\mathbb{E}(\mathbf{1}_A \otimes x|\mathscr{F}_n) = \mathbb{E}(\mathbf{1}_A|\mathscr{F}_n)\otimes x$, it is enough to prove the result for indicators $f = \mathbf{1}_A$ with $A \in \mathscr{A}$ and $\mu(A) < \infty$.

For $p = 2$, the result was already proved. For $p \in (2,\infty)$ it follows from the almost everywhere inequality $|f_n - f_{-\infty}| \leqslant 2$ that $\|f_n - f_{-\infty}\|_p^p \leqslant 2^{p-2}\|f_n - f_{-\infty}\|_2^2 \to 0$.

Let us next consider $p \in [1,2)$. Fix $\varepsilon > 0$. Using the σ-finiteness of $\mathscr{F}_{-\infty}$, choose $F \in \mathscr{F}_{-\infty}$ of finite measure such that $\|\mathbf{1}_{\complement F}(f_m - f_{-\infty})\|_p < \varepsilon$. From Jensen's inequality and the definition of a conditional expectation we see that for all $n \leqslant m$, $\|\mathbf{1}_{\complement F}(f_n - f_{-\infty})\|_p \leqslant \varepsilon$. Therefore, we have

$$\|f_n - f_{-\infty}\|_p \leqslant \varepsilon + \|\mathbf{1}_F(f_n - f_{-\infty})\|_p \leqslant \varepsilon + \mu(F)^{\frac{1}{p}-\frac{1}{2}}\|f_n - f_{-\infty}\|_2,$$

and the result follows since $\lim_{n\to\infty}\|f_n - f_{-\infty}\|_2 = 0$.

(1): Let $(S_j)_{j\geqslant 1}$ be a disjoint exhausting sequence for f in $\mathscr{F}_{-\infty}$ (cf. Definition 2.6.1). Fix $j \geqslant 1$ and define the martingale $(g_n)_{n\in\mathbb{Z}}$ by $g_n = \mathbf{1}_{S_j} f_n - \mathbf{1}_{S_j} f_{-\infty}$. By Doob's maximal inequality applied to the submartingale $\|g_n\|$, one obtains that for all $r > 0$,

$$\mu\big(\lim_{n\to-\infty} g_n^\star > r\big) = \lim_{n\to-\infty} \mu(g_n^\star > r) \leqslant \frac{1}{r} \lim_{n\to-\infty} \|g_n\|_1 = 0,$$

where we applied (2) to g_n. Therefore, $\lim_{n\to-\infty} g_n^\star = 0$ almost everywhere. This implies $\lim_{n\to-\infty} f_n = f_{-\infty}$ almost everywhere on S_j, and then by exhaustion almost everywhere on S.

(3): Let $g := \limsup_{n\to-\infty} \|f_n\|$. Choosing strongly $\mathscr{F}_n$-measurable versions of the f_n's, we find that g is an $\mathscr{F}_{-\infty}$-measurable $[0,\infty]$-valued function. The pointwise bound $g \leqslant f^\star$ is also clear. Consider the set $\{g > \varepsilon\} \in \mathscr{F}_{-\infty}$. Since μ is purely infinite on $\mathscr{F}_{-\infty}$, we have $\mu(g > \varepsilon) \in \{0, \infty\}$. On the other hand,

$$\mu(g > \varepsilon) \leqslant \mu(f^\star > \varepsilon) \leqslant \frac{1}{\varepsilon^p} \|f\|_p^p < \infty$$

by Doob's inequality. Hence $\mu(g > \varepsilon) = 0$ for any $\varepsilon > 0$, and thus $g = 0$ almost everywhere. This proves that $\lim_{n\to-\infty} f_n = 0$ almost everywhere.

If $p \in (1, \infty)$, then $\|f_n\| \leqslant f^\star \in L^p(S)$ by Doob's maximal inequality. The almost everywhere convergence just proved implies L^p-norm convergence, by the dominated convergence theorem. □

Corollary 3.3.7. *Let $1 \leqslant p < \infty$ be given. Then for any L^p-martingale $(f_n)_{n\in\mathbb{Z}}$ the limit $\lim_{n\to-\infty} f_n$ exists almost everywhere, and also in $L^p(S;X)$ if $1 < p < \infty$. In particular, this holds for $f_n = \mathbb{E}(f|\mathscr{F}_n)$, for any $f \in L^p(S;X)$.*

Proof. This follows by combining Theorem 3.3.5 with Proposition A.1.4. □

By combining the convergence results of this section, the following representation of a function in terms of its martingale differences is obtained:

Theorem 3.3.8 (Martingale convergence). *Let $f \in L^p(S;X)$ for some $p \in [1,\infty)$, let $(\mathscr{F}_n)_{n\in\mathbb{Z}}$ be a σ-finite filtration of $(S, \mathscr{A}, \mu)$, and let $f_n := \mathbb{E}(f|\mathscr{F}_n)$ for $n \in \mathbb{Z}$. Then the limits*

$$f_\infty = \lim_{n\to\infty} f_n = \mathbb{E}(f|\mathscr{F}_\infty), \qquad f_{-\infty} = \lim_{m\to-\infty} f_m$$

exist almost everywhere, and we have

$$f_\infty = f_{-\infty} + \sum_{n=-\infty}^{\infty} df_n, \qquad df_n := f_n - f_{n-1},$$

with convergence almost everywhere. The convergence holds in $L^p(S;X)$ if $p \in (1,\infty)$ or if μ is σ-finite on $\mathscr{F}_{-\infty}$. In the latter case, $f_{-\infty} = \mathbb{E}(f|\mathscr{F}_{-\infty})$.

Proof. By telescoping,

$$f_\infty - f_{-\infty} = \lim_{n\to\infty} f_n - \lim_{m\to-\infty} f_m = \lim_{\substack{n\to+\infty \\ m\to-\infty}} \sum_{k=m+1}^{n} (f_k - f_{k-1})$$

and the existence of the limit on the right is, by definition, the same as the convergence of the series in the assertion. □

We conclude this subsection with two beautiful applications of backward convergence. The first concerns the almost everywhere approximation of Bochner integrals by Riemann sums.

Theorem 3.3.9 (Jessen). *Let X be a Banach space and $(m_n)_{n\geqslant 1}$ be a sequence of positive integers tending to ∞, where m_{n+1} is divisible by m_n for every n. For all $f \in L^1(0,1;X)$ the Riemann sums*

$$\frac{1}{m_n}\sum_{k=1}^{m_n} f\Big(t + \frac{k}{m_n} \mod 1\Big) \tag{3.38}$$

converge to $\int_0^1 f(s)\,\mathrm{d}s$ at almost every $t \in [0,1)$.

Proof. For the convenience of notation, we understand that all additions on the unit interval are taken modulo 1, without specifying this explicitly.

To derive the theorem from martingale convergence, we define two (in a sense complementary) filtrations on $[0,1)$ by

$$\mathscr{F}_{-n} := \Big\{E \subseteq [0,1) \text{ measurable} : E + \frac{1}{m_n} = E\Big\}$$

and

$$\mathscr{G}_n := \sigma\Big(\Big\{J_{n,j} := \Big[\frac{j-1}{m_n}, \frac{j}{m_n}\Big) : j = 1,\dots,m_n\Big\}\Big).$$

It is easy to see that these are σ-algebras, and the fact that they are filtrations follows from the assumption that m_{n+1} is divisible by m_n.

We first claim that $\mathbb{E}(f|\mathscr{F}_{-n})$ is given by the expression (3.38). Indeed, if $E \in \mathscr{F}_{-n}$, then $E - k/m_n = E + (m_n - k)/m_n = E$, and

$$\begin{aligned}\int_E \frac{1}{m_n}\sum_{k=1}^{m_n} f\Big(t + \frac{k}{m_n}\Big)\,\mathrm{d}t &= \frac{1}{m_n}\sum_{k=1}^{m_n}\int_{E-k/m_n} f\,\mathrm{d}t \\ &= \frac{1}{m_n}\sum_{k=1}^{m_n}\int_E f\,\mathrm{d}t = \int_E f\,\mathrm{d}t,\end{aligned}$$

so it remains to check that the expression in (3.38) is $\mathscr{F}_{-n}$-measurable. If $f = \mathbf{1}_F$, then

$$\begin{aligned}\sum_{k=1}^{m_n} \mathbf{1}_F\Big(\cdot + \frac{k}{m_n}\Big) &= \sum_{k=1}^{m_n} \mathbf{1}_{F-k/m_n} = \sum_{k=1}^{m_n}\sum_{j=1}^{m_n} \mathbf{1}_{(F-k/m_n)\cap J_{n,j-k}} \\ &= \sum_{j=1}^{m_n} \mathbf{1}_{\bigcup_{k=1}^{m_n}(F-k/m_n)\cap J_{n,j-k}} =: \sum_{j=1}^{m_n} \mathbf{1}_{F_{n,j}},\end{aligned}$$

where it is easy to see that $F_{n,j} \in \mathscr{F}_{-n}$. Thus the expression in (3.38) is $\mathscr{F}_{-n}$-measurable for every indicator function f, hence for every simple f by linearity, and thus for every $f \in L^1(0,1;X)$ by approximation.

Our next goal is to identify $\mathscr{F}_{-\infty} := \bigcap_{n\leqslant -1} \mathscr{F}_n$ and $\mathbb{E}(f|\mathscr{F}_{-\infty})$. By definition, $\mathscr{F}_{-\infty}$ consists of measurable F such that $F + 1/m_n = F$ for every n.

We claim that $|F| \in \{0,1\}$ for every $F \in \mathscr{F}_{-\infty}$. Since $F \in \mathscr{F}_{-n}$, it follows that

$$|F \cap J_{n,j}| = \Big|\Big(F + \frac{j}{m_n}\Big) \cap \Big(J_{n,0} + \frac{j}{m_n}\Big)\Big| = \Big|(F \cap J_{n,0}) + \frac{j}{m_n}\Big| = |F \cap J_{n,0}|$$

by the translation invariance of the Lebesgue measure. Thus $\mathbb{E}(\mathbf{1}_F|\mathscr{G}_n)$, whose value on $J_{n,j}$ is $|F \cap J_{n,j}|/|J_{n,j}|$, is actually constant on $[0,1)$, and this constant must be equal to $|F|$. From the forward convergence of martingales it follows that $\mathbf{1}_F = \lim_{n\to\infty} \mathbb{E}(\mathbf{1}_F|\mathscr{G}_n) = |F|$ almost everywhere, and this is only possible with $|F| \in \{0,1\}$.

Thus every indicator of a set in $\mathscr{F}_{-\infty}$ is constant almost everywhere, and so is every $\mathscr{F}_{-\infty}$-simple function by linearity, and thus every $\mathscr{F}_{-\infty}$-measurable function by approximation. In particular, $\mathbb{E}(f|\mathscr{F}_{-\infty})$ is a constant, which must be equal to $\int_0^1 f(t)\,\mathrm{d}t$ by the defining property of conditional expectations. But then the assertion of the theorem is immediate from the almost everywhere convergence $\mathbb{E}(f|\mathscr{F}_{-n}) \to \mathbb{E}(f|\mathscr{F}_{-\infty})$ as $n \to \infty$. □

As a second illustration of the backward and forward convergence techniques we present a proof of the strong law of large numbers.

Theorem 3.3.10 (Strong law of large numbers). *Let X be a Banach space and let $(\xi_n)_{n\geqslant 1}$ be a sequence of independent and identically distributed random variables in $L^1(\Omega;X)$, where $(\Omega,\mathscr{A},\mathbb{P})$ is a probability space. Then*

$$\lim_{n\to\infty} \Big(\frac{1}{n}\sum_{k=1}^n \xi_k\Big) = \mathbb{E}(\xi_1)$$

in $L^1(\Omega;X)$ and almost surely in X.

Proof. The proof assumes familiarity with some elementary results concerning independence.

We introduce the filtrations

$$\mathscr{T}_{-n} := \sigma(\xi_{n+1},\xi_{n+2},\dots), \qquad n \geqslant 0;$$

$$\mathscr{F}_{-n} := \sigma(\overline{\xi}_n, \overline{\xi}_{n+1}, \dots) = \sigma(\overline{\xi}_n, \mathcal{T}_{-n}), \qquad n \geqslant 1,$$

where $\overline{\xi}_n := \frac{1}{n}\sum_{k=1}^n \xi_k$. By Lemma 2.6.37

$$\mathbb{E}(\xi_1|\mathscr{F}_{-n}) = \mathbb{E}(\xi_1|\sigma(\overline{\xi}_n, \mathcal{T}_{-n})) = \mathbb{E}(\xi_1|\sigma(\overline{\xi}_n)) = \overline{\xi}_n, \quad n \geqslant 1,$$

which shows that $(\overline{\xi}_n)_{n\geqslant 1}$ is a martingale with respect to $(\mathscr{F}_{-n})_{n\geqslant 1}$. By Theorem 3.3.5, the limit

$$\lim_{n\to\infty} \overline{\xi}_n = \lim_{n\to\infty} \mathbb{E}(\xi_1|\mathscr{F}_{-n}) =: \zeta$$

exists in $L^1(\Omega; X)$ and almost surely.

It remains to be shown that $\zeta = \mathbb{E}(\xi_1)$. For all $m \geqslant 1$,

$$\zeta = \lim_{n\to\infty} \frac{1}{n} \sum_{k=m+1}^{n} \xi_k$$

in $L^1(\Omega; X)$ and almost surely, and therefore ζ is strongly $\mathcal{T}_{-m}$-measurable.

For all $A \in \mathcal{T}_{-\infty}$ we have $\mu(A) \in \{0, 1\}$. To prove this let $\mathscr{G}_n := \sigma(\xi_1, \dots, \xi_n)$. By Theorem 3.3.2,

$$\lim_{n\to\infty} \mathbb{E}(\mathbf{1}_A|\mathscr{G}_n) = \mathbb{E}(\mathbf{1}_A|\mathscr{G}_\infty) = \mathbf{1}_A$$

in $L^1(\Omega; X)$ and almost surely. On the other hand, for each $n \geqslant 1$, $\mathbf{1}_A$ is $\mathcal{T}_{-n}$-measurable and therefore independent of $\mathscr{G}_n$, so that $\mathbb{E}(\mathbf{1}_A|\mathscr{G}_n) = \mathbb{E}(\mathbf{1}_A) = \mu(A)$ by Proposition 2.6.35. We conclude that $\mathbf{1}_A = \mu(A)$ almost surely, and the claim follows from this.

Since ζ is strongly measurable with respect to $\mathcal{T}_{-\infty}$ it equals a constant almost surely, and the constant is given by $\zeta = \mathbb{E}(\zeta) = \lim_{n\to\infty} \mathbb{E}(\overline{\xi}_n) = \mathbb{E}(\xi_1)$. This completes the proof. $\square$

3.3.c The Itô–Nisio theorem for martingales

Throughout this subsection X is a Banach space and $(S, \mathscr{A}, \mu)$ is a measure space. Unless otherwise stated, all martingales $(f_n)_{n\geqslant 0}$ will be taken with respect to a given σ-finite filtration $(\mathscr{F}_n)_{n\geqslant 0}$ which we fix once and for all. The aim is to prove that almost everywhere forward convergence of martingales $f_n \to f$ is equivalent to a number of *a priori* weaker conditions, the weakest among them being that $\langle f_n, x^*\rangle \to \langle f, x^*\rangle$ in measure for all $x^* \in X^*$.

We begin our analysis with a useful lemma.

Lemma 3.3.11. *Let $(f_n)_{n\geqslant 0}$ be a martingale in X satisfying $\int_S f^\star \,\mathrm{d}\mu < \infty$, where $f^\star = \sup_{n\geqslant 0} \|f_n\|$. For $n \geqslant 0$ define the vector measures $F_n : \mathscr{F}_\infty \to X$ by $F_n(A) := \int_A f_n \,\mathrm{d}\mu$. Then:*

(1) *for all $A \in \mathscr{F}_\infty$ the sequence $(F_n(A))_{n\geqslant 0}$ is convergent in X;*

(2) *the function* $F : \mathscr{F}_\infty \to X$ *defined by* $F(A) = \lim_{n\to\infty} F_n(A)$ *is an* X*-valued measure;*
(3) F *is of bounded variation and absolutely continuous with respect to* μ*.*

Proof. Define the finite measure $\nu : \mathscr{F}_\infty \to [0,\infty)$ by $\nu(A) = \int_A f^\star \,\mathrm{d}\mu$. Clearly, for all $n \geqslant 0$ and $A \in \mathscr{F}_\infty$, $\|F_n\|(A) \leqslant \nu(A)$.

To prove (1) choose an arbitrary $A \in \mathscr{F}_\infty$ and fix $\varepsilon > 0$. Since ν is a finite measure it follows from Lemma A.1.2 that there is a set $B \in \mathscr{G} := \bigcup_{n\geqslant 0} \mathscr{F}_n$ such that $\nu(A\Delta B) < \varepsilon$. Let $N \geqslant 1$ be an index such that $B \in \mathscr{F}_N$. By the martingale property, for all $n \geqslant N$ we have

$$F_n(B) = \int_B f_n \,\mathrm{d}\mu = \int_B f_N \,\mathrm{d}\mu = F_N(B). \tag{3.39}$$

For all $n, m \geqslant N$ this gives

$$\begin{aligned}\|F_n(A) - F_m(A)\| &\leqslant \|F_n(A) - F_n(B)\| + \|F_m(B) - F_m(A)\| \\ &\leqslant 2\nu(A\Delta B) < 2\varepsilon.\end{aligned}$$

Therefore $(F_n(A))_{n\geqslant 1}$ is a Cauchy sequence in X. It follows that the limit $F(A) := \lim_{n\to\infty} F_n(A)$ exists in X.

To prove (2), first note that $\|F(A)\| = \lim_{n\to\infty} \|F_n(A)\| \leqslant \nu(A)$ for all $A \in \mathscr{F}_\infty$. To prove that F is a vector measure let $A = \bigcup\limits_{i\geqslant 1} A_i$ with $A_1, A_2, \ldots \in \mathscr{F}_\infty$ disjoint. Then, since F is finitely additive,

$$\Big\|F(A) - \sum_{i=1}^n F(A_i)\Big\| = \Big\|F\Big(\bigcup_{i\geqslant n+1} A_i\Big)\Big\| \leqslant \nu\Big(\bigcup_{i\geqslant n+1} A_i\Big) = \sum_{i\geqslant n+1} \nu(A_i)$$

and the right-hand side tends to 0 as $n \to \infty$. This completes the proof of (2).

It is clear that F is absolutely continuous with respect to ν, and therefore also with respect to μ. If $S = \bigcup_{n=1}^N A_n$ is a measurable partition, then

$$\sum_{n=1}^N \|F(A_n)\| \leqslant \sum_{n=1}^N \nu(A_n) = \nu(S).$$

It follows that F is of bounded variation, with $\|F\|(S) \leqslant \nu(S)$. This proves (3). □

Lemma 3.3.12. *Suppose that* μ *is finite and let* $(f_n)_{n\geqslant 0}$ *be a bounded* L^1*-martingale in* X *with respect to* $(\mathscr{F}_n)_{n\geqslant 0}$*. Fix* $\lambda > 0$ *and consider the martingale* $(g_n)_{n\geqslant 0}$ *defined by* $g_n = f_{\tau\wedge n}$*,* $n \geqslant 0$*, where* $\tau = \inf\{n \geqslant 0 : \|f_n\| > \lambda\}$*. Then*

$$\int_S g^\star \,\mathrm{d}\mu \leqslant \lambda\mu(S) + \liminf_{n\to\infty} \|f_n\|_{L^1(S)}.$$

Proof. On the set $\{\tau = \infty\}$ one has $g^\star \leqslant \lambda$, and $g^\star = f_\tau$ on $\{\tau < \infty\}$. Therefore,

$$\begin{aligned}
\int_S g^\star \,\mathrm{d}\mu &\leqslant \lambda\mu(S) + \int_{\{\tau<\infty\}} \|f_\tau\|\,\mathrm{d}\mu \\
&= \lambda\mu(S) + \lim_{N\to\infty} \sum_{n=0}^{N} \int_{\{\tau=n\}} \|f_n\|\,\mathrm{d}\mu \\
&\leqslant \lambda\mu(S) + \liminf_{N\to\infty} \sum_{n=0}^{N} \int_{\{\tau=n\}} \|f_N\|\,\mathrm{d}\mu \\
&= \lambda\mu(S) + \liminf_{N\to\infty} \int_{\{\tau\leqslant N\}} \|f_N\|\,\mathrm{d}\mu \\
&\leqslant \lambda\mu(S) + \liminf_{n\to\infty} \|f_n\|_{L^1(S)},
\end{aligned}$$

where the second inequality used that $(\|f_n\|)_{n\geqslant 0}$ is a submartingale (see Example 3.1.5) and that $\{\tau = n\} \in \mathscr{F}_n$. $\square$

A sequence of μ-measurable functions $(f_n)_{n\geqslant 1}$ is said to *converge in measure* to f if for all $\varepsilon > 0$ and $r > 0$ and all sets $A \in \mathscr{A}$ of finite measure we have

$$\lim_{n\to\infty} \mu(A \cap \{|f_n - f| > r\}) = 0.$$

We refer to Section A.2 for more further details. A sequence of strongly μ-measurable X-valued functions $(f_n)_{n\geqslant 1}$ is said to *converge in measure* if $\|f_n - f\| \to 0$ in measure for $n \to \infty$.

Lemma 3.3.13. *Let $(f_n)_{n\geqslant 1}$ be a sequence of strongly $\mathscr{F}_n$-measurable X-valued functions. Let f be a strongly μ-measurable X-valued function. If for all $x^* \in X^*$ we have $\langle f_n, x^*\rangle \to \langle f, x^*\rangle$ in measure, then f has a strongly $\mathscr{F}_\infty$-measurable version.*

Proof. For each $x^* \in X^*$, using Lemma A.2.3 we find a subsequence (which depends on x^*) such that $\langle f_{n_k}, x^*\rangle \to \langle f, x^*\rangle$ almost surely. From this we find that $\langle f, x^*\rangle$ has an $\mathscr{F}_\infty$-measurable version. Therefore, by Proposition 1.1.16 (which can be applied since the standing assumptions imply that μ is σ-finite on $\mathscr{F}_\infty$) and the Pettis measurability theorem we find that f has a strongly $\mathscr{F}_\infty$-measurable version. $\square$

Theorem 3.3.14 (Itô–Nisio theorem for martingales). *Let $p \in [1, \infty]$, and suppose that $(f_n)_{n\geqslant 0}$ is a bounded L^p-martingale in X with respect to a σ-finite filtration $(\mathscr{F}_n)_{n\geqslant 0}$. Let f be a strongly μ-measurable X-valued function. The following assertions are equivalent:*

(1) $\lim_{n\to\infty} f_n = f$ *almost everywhere;*
(2) $\lim_{n\to\infty} f_n = f$ *in measure;*

(3) $\lim_{n\to\infty}\langle f_n, x^*\rangle = \langle f, x^*\rangle$ *almost everywhere;*
(4) $\lim_{n\to\infty}\langle f_n, x^*\rangle = \langle f, x^*\rangle$ *in measure.*

In this situation, $f \in L^p(S;X)$ *and* $\|f\|_p \leqslant \liminf_{n\to\infty} \|f_n\|_p$.

Proof. The implications (2)⇒(4) and (1)⇒(3) are obvious. The implications (1)⇒(2) and (3)⇒(4) follow from Lemma A.2.2.

For the proof of (4)⇒(1) we begin by observing that each f_n has a pointwise defined representative which is strongly $\mathscr{F}_n$-measurable. Thus, by Lemma 3.3.13, we may assume that f is strongly $\mathscr{F}_\infty$-measurable.

The remainder of the proof proceeds in three steps.

Step 1 – We first assume that $\mu(S) < \infty$ and that $(f_n)_{n\in\mathbb{Z}}$ is a bounded L^1-martingale satisfying $\|f^\star\|_1 < \infty$.

Fixing a $\mu|_{\mathscr{F}_\infty}$-simple function $g : S \to X^*$, by Lemma A.2.3 we find a subsequence $(f_{n_k})_{k\geqslant 1}$ (depending on the finite-dimensional range of g) such that $\lim_{k\to\infty}\langle f_{n_k}, g\rangle = \langle f, g\rangle$ almost everywhere. Then

$$\int_S |\langle f, g\rangle| \,\mathrm{d}\mu = \lim_{k\to\infty} \int_S |\langle f_{n_k}, g\rangle| \,\mathrm{d}\mu \leqslant \|f^\star\|_1 \|g\|_\infty,$$

the use of dominated convergence being justified by the integrability of the maximal function $f^\star$. Taking the supremum over all $\mu|_{\mathscr{F}_\infty}$-simple function $g \in L^\infty(S;X^*)$ of norm one and then letting $N \to \infty$, we infer that $f \in L^1(S;X)$ and $\|f\|_1 \leqslant \int_S f^\star \,\mathrm{d}\mu$.

Fix $x^* \in X^*$. Again by passing to a subsequence (which depends on the choice of x^*) we may assume that $\lim_{k\to\infty}\langle f_{n_k}, x^*\rangle = \langle f, x^*\rangle$ almost everywhere. On the other hand, as $f^\star \in L^1(S)$, the dominated convergence theorem can be applied to see that $\lim_{k\to\infty}\langle f_{n_k}, x^*\rangle = \langle f, x^*\rangle$ in $L^1(S)$.

Next, with the notations of Lemma 3.3.11, for all $m \geqslant 0$ and $A \in \mathscr{F}_m$ we obtain

$$\Big\langle \int_A f \,\mathrm{d}\mu, x^* \Big\rangle = \lim_{k\to\infty} \int_A \langle f_{n_k}, x^*\rangle \,\mathrm{d}\mu = \lim_{k\to\infty} \langle F_{n_k}(A), x^*\rangle = \langle F(A), x^*\rangle.$$

Since $x^* \in X^*$ was arbitrary it follows that

$$F(A) = \int_A f \,\mathrm{d}\mu. \tag{3.40}$$

On the other hand, by the definition of F and the fact that $A \in \mathscr{F}_m$,

$$F(A) = \lim_{n\to\infty} F_n(A) = \lim_{n\to\infty} \int_A f_n \,\mathrm{d}\mu = \int_A f_m \,\mathrm{d}\mu. \tag{3.41}$$

Combining (3.40) and (3.41) it follows that

$$\mathbb{E}(f|\mathscr{F}_m) = f_m.$$

The almost everywhere convergence $f_m \to f$ now follows from Theorem 3.3.2.

Step 2 – Next we consider the general case where $(f_n)_{n\geqslant 0}$ is a bounded L^1-martingale, still assuming that $\mu(S) < \infty$. Fix $\lambda > 0$ and let the martingale $(g_n)_{n\geqslant 0}$ be defined by $g_n = f_{\tau\wedge n}$, $n \geqslant 0$, where $\tau = \inf\{n \geqslant 0 : \|f_n\| > \lambda\}$. Clearly, $\lim_{n\to\infty}\langle g_n, x^*\rangle = \mathbf{1}_{\{\tau<\infty\}}\langle f_\tau, x^*\rangle + \mathbf{1}_{\{\tau=\infty\}}\langle f, x^*\rangle$ in measure. By Lemma 3.3.12 we have $\int_S g^\star \,\mathrm{d}\mu < \infty$ and therefore, by Step 1, we obtain that $g_n \to g := \mathbf{1}_{\{\tau<\infty\}} f_\tau + \mathbf{1}_{\{\tau=\infty\}} f$ almost everywhere. In particular, on the set $\{\tau = \infty\} = \{f^\star \leqslant \lambda\}$, $f_n = g_n$ converges almost everywhere to f. Considering a sequence $\lambda_m \to \infty$, it follows that f_n converges to f on the set $\{f^\star < \infty\}$ almost everywhere. By Doob's maximal inequality (applied to $\|f_n\|$),

$$\mu(f^\star = \infty) = \lim_{\lambda\to\infty} \mu(f^\star > \lambda) \leqslant \lim_{\lambda\to\infty} \lambda^{-1} \sup_{n\geqslant 0} \|f_n\|_1 = 0,$$

and we conclude that $f_n \to f$ almost everywhere.

Step 3 – We now remove the finiteness assumption on μ and consider the case where f is a bounded L^p-martingale with $p \in [1, \infty]$. Let $(S_i)_{i\geqslant 1}$ be an exhaustion of S by sets in $\mathscr{F}_0$ of finite μ-measure. For each $i \geqslant 1$, $(\mathbf{1}_{S_i} f_n)_{n\geqslant 1}$ defines a bounded L^1-martingale on the finite measure space $(S_i, \mathscr{F}|_{S_i}, \mu|_{S_i})$ in a natural way. By Step 2, $\mathbf{1}_{S_i} f_n \to \mathbf{1}_{S_i} f$ almost surely for each $i \geqslant 1$, and the result follows.

The final assertion follows from Fatou's lemma. $\square$

Remark 3.3.15. It is clear from Step 3 of the above proof that Theorem 3.3.14 extends to the case where $(f_n)_{n\geqslant 0}$ is a martingale such that $(\mathbf{1}_{S_0} f_n)_{n\geqslant 0}$ is L^1-bounded on every set $S_0 \in \mathscr{F}_0$ of finite measure.

3.3.d Martingale convergence and the RNP

In the next result we strengthen Theorem 3.3.2 in the case the Banach space X has the Radon–Nikodým property (RNP) introduced in Section 1.3.

Theorem 3.3.16 (Convergence of L^p-bounded martingales). *For a Banach space X the following are equivalent:*

(1) *X has the Radon–Nikodým property;*
(2) *for all (equivalently, some) $p \in [1, \infty)$, any L^p-bounded X-valued martingale $(f_n)_{n\geqslant 0}$ with respect to a σ-finite filtration converges almost everywhere in X;*
(3) *For all (equivalently, some) $p \in (1, \infty)$, every L^p-bounded X-valued martingale $(f_n)_{n\geqslant 0}$ with respect to a σ-finite filtration converges in $L^p(S; X)$;*
(4) *Every L^1-bounded X-valued martingale $(f_n)_{n\geqslant 0}$ with respect to a σ-finite filtration which is uniformly integrable converges in $L^1(S; X)$.*

The assertions (2)-(4) *are equivalent to their counterparts where only martingales on the unit interval* $[0, 1]$ *with the dyadic filtration are considered.*

For the definition of uniform integrability we refer to Section A.3.

Proof. We begin the proof of (1)⇒(2) for $p = 1$. As in the proof of Theorem 3.3.14, we first prove the result under the additional assumption that μ is finite and $f^\star \in L^1(S)$.

The measure $F : \mathscr{F}_\infty \to X$ defined by Lemma 3.3.11 has bounded variation and is absolutely continuous with respect to μ. Since X is assumed to have the RNP, there exists a function $f_\infty \in L^1(S, \mathscr{F}_\infty; X)$ such that

$$F(A) = \int_A f_\infty \,\mathrm{d}\mu, \quad A \in \mathscr{F}_\infty.$$

In particular, by (3.39), for all $B \in \mathscr{F}_n$ we have

$$\int_B f_\infty \,\mathrm{d}\mu = F(B) = F_n(B) = \int_B f_n \,\mathrm{d}\mu,$$

so that

$$\mathbb{E}(f_\infty | \mathscr{F}_n) = f_n, \quad n \geqslant 0.$$

Now the almost everywhere convergence of f_n to f_∞ follows from Theorem 3.3.2.

The case where μ is finite and $(f_n)_{n\geqslant 0}$ is a bounded L^1-martingale follows from Step 1 by repeating Step 2 of the proof of Theorem 3.3.14. The finiteness assumption on μ and the restriction to $p = 1$ can subsequently be removed as in Step 3 of the proof of Theorem 3.3.14.

Suppose next that (2) holds for some $p \in [1, \infty)$. As a preliminary observation we note that if $f_\infty = \lim_{n\to\infty} f_n$ is the almost everywhere limit of the L^p-bounded martingale $(f_n)_{n\geqslant 0}$, then by Fatou's lemma

$$\|f_\infty\|_{L^p(S;X)} \leqslant \liminf_{n\to\infty} \|f_n\|_{L^p(S;X)} \leqslant \sup_{n\geqslant 0} \|f_n\|_{L^p(S;X)} < \infty.$$

We will show next that if (2) holds for some $p \in (1, \infty)$, then (3) holds for the same $p \in (1, \infty)$. Indeed, we have $\|f_\infty - f_n\|^p \leqslant 2^p \|f^\star\|^p$, and by Doob's maximal inequality we have $\|f^\star\|_{L^p(S)} \leqslant p' \sup_{n\geqslant 0} \|f_n\|_{L^p(S;X)}$. Therefore (3) follows by dominated convergence.

If (2) holds for $p = 1$, then (4) holds. This follows from Lemma A.2.2 and Proposition A.3.5.

Suppose finally that either (2) or (3) holds, for some $p \in [1, \infty)$ or $p \in (1, \infty)$ respectively, or that (4) holds, in all three cases for martingales on the unit interval $[0, 1]$ with the dyadic filtration. In all these cases we will deduce that X has the RNP with respect to $[0, 1]$. By Theorem 1.3.26, this implies that X has the RNP.

Let $\mathscr{B} := \mathscr{B}([0, 1])$ be the Borel σ-algebra and let $F : \mathscr{B} \to X$ be a vector measure of bounded variation which is absolutely continuous with respect to the Lebesgue measure μ. By Lemma 1.3.6, $\|F\|$ is a finite measure on $\mathscr{B}$, which is easily seen to be absolutely continuous with respect to μ. Therefore, by the scalar Radon–Nikodým theorem, there is a function $g \in L^1(0,1)$ such that $\|F\|(B) = \int_B g \, \mathrm{d}\mu$ for all $B \in \mathscr{B}$.

Let $\mathscr{D}_n$ be the nth dyadic σ-algebra in $[0,1]$ and let $\mathscr{D}_n^*$ denote its set of atoms. Set

$$f_n := \sum_{D \in \mathscr{D}_n^*} \mathbf{1}_D \frac{F(D)}{\|F\|(D)},$$

where we use the convention $0/0 = 0$. Then for all $n \geqslant m \geqslant 0$ and for all $B \in \mathscr{D}_m$,

$$\begin{aligned} \int_B f_n \, \mathrm{d}\|F\| &= \sum_{D \in \mathscr{D}_n^*,\, D \subseteq B} \frac{F(D)}{\|F\|(D)} \int_B \mathbf{1}_D \, \mathrm{d}\|F\| \\ &= \sum_{D \in \mathscr{D}_n^*,\, D \subseteq B} \frac{F(D)}{\|F\|(D)} \|F\|(D) = F(B). \end{aligned} \tag{3.42}$$

In particular, $\int_B f_n \, \mathrm{d}\|F\| = \int_B f_m \, \mathrm{d}\|F\|$. This shows that $(f_n)_{n \geqslant 0}$ is a martingale on the measure space $([0,1], \mathscr{B}, \|F\|)$ with respect to the dyadic filtration $(\mathscr{D}_n)_{n \geqslant 0}$. Clearly, $\|f_n\| \leqslant 1$ almost everywhere. We claim that in any one of the cases (2)–(4), there exists a function $f_\infty \in L^1(0,1;X)$ such that $f_n \to f_\infty$ in $L^1(0,1,\|F\|;X)$. Indeed, if (2) holds, then $f_\infty = \lim_{n\to\infty} f_n$ exists $\|F\|$-almost surely and since $\|f_\infty - f_n\| \leqslant 2$, the claim follows from the dominated convergence theorem. If (3) holds, then $f_\infty = \lim_{n\to\infty} f_n$ in $L^p(0,1,\|F\|;X)$, and since $\|F\|$ is a finite measure the claim holds in this case as well. If (4) holds, then the claim holds since the functions $\|f_n\|$ are uniformly bounded on a finite measure space and therefore uniformly integrable (see Proposition A.3.4).

Let $G : \mathscr{B} \to X$ be defined by

$$G(B) = \int_B f_\infty \, \mathrm{d}\|F\| = \int_B g f_\infty \, \mathrm{d}\mu.$$

Thus G has density $g f_\infty$ with respect to μ. We complete the proof by showing that $F = G$. Since $\|f_\infty\| \leqslant 1$ we have $\|G(B)\| \leqslant \|F\|(B)$ for all $B \in \mathscr{B}$. If $D \in \mathscr{D}_m$, then (3.42) implies that for all $n \geqslant m$

$$\|F(D) - G(D)\| = \Big\| \int_D f_n - f_\infty \, \mathrm{d}\|F\| \Big\| \leqslant \|f_n - f_\infty\|_{L^1(0,1,\|F\|;X)}.$$

Since the latter tends to zero as n tends to infinity, we obtain $F(D) = G(D)$ for all $D \in \bigcup_{m \geqslant 1} \mathscr{D}_m$. By Lemma A.1.2 this implies $F(B) = G(B)$ for all $B \in \mathscr{B}$. □

3.4 Martingale decompositions

In various applications, it is useful to split a given general martingale into two or three parts, where each part takes care of specific singularities of the martingale but is free from some others. In this way, the analysis of the original martingale is divided into a conceptually simpler case study, where it is not necessary to address all difficulties at the same time. A general strategy to such decompositions goes via the stopping time technique, where one "stops at the first time n when something goes wrong" with the function f_n. Details on stopped martingales can be found in Section 3.1.c. We proceed to prove two classical decomposition results due to Gundy and Davis.

3.4.a Gundy decomposition

Gundy's decomposition of a martingale plays an important role in establishing weak type $(1,1)$ bounds for martingale transforms. It can be viewed as the probabilistic counterpart of the classical Calderón-Zygmund decomposition from harmonic analysis.

Theorem 3.4.1 (Gundy decomposition). *Let $f = (f_n)_{n\in\mathbb{Z}}$ be a martingale in $L^1(S;X)$. For every $\lambda > 0$ there exists a decomposition*

$$f = g + b + h,$$

where g (good), b (bad) and h (harmless) are martingales with respect to the same filtration, such that

$$g_{-\infty} = f_{-\infty}, \quad b_{-\infty} = h_{-\infty} = 0,$$

and the following estimates hold for all $n \in \mathbb{Z}$:

$$\begin{aligned} \|g_n\|_\infty \leqslant 2\lambda, \qquad & \|g_n\|_1 \leqslant 4\|f_n\|_1, \\ \lambda \cdot \mu(b_n^\star > 0) \leqslant 3\|f_n\|_1, & \\ \sum_{k\leqslant n} \|dh_k\|_1 \leqslant 4\|f_n\|_1. & \end{aligned}$$

Remark 3.4.2. From the first two bounds it is immediate that $g_n \in L^p(S;X)$ for all $1 \leqslant p \leqslant \infty$, with the estimate

$$\|g_n\|_p^p \leqslant \|g_n\|_\infty^{p-1}\|g_n\|_1 \leqslant (2\lambda)^{p-1} \cdot 4\|f_n\|_1.$$

Proof. By Corollary 3.3.7, the limit $f_{-\infty} = \lim_{n\to-\infty} f_n$ exists almost everywhere.

We first assume that $\|f_{-\infty}\| < \lambda$ almost everywhere, and lift this assumption only at the end of the proof. We start by motivating the construction to be given. To extract the bounded part g of f, it seems natural to consider the stopping time

$$\tau := \inf\{n : \|f_n\| > \lambda\}.$$

The assumption $\|f_{-\infty}\| < \lambda$ implies that $\lim_{n\to-\infty} \|f_n\| < \lambda$ almost everywhere, and therefore $\tau > -\infty$ almost everywhere. Thus $\|f_{n\wedge(\tau-1)}\| \leqslant \lambda$, so if we had some control of the jumps df_n, we could conclude that the stopped martingale $f_{n\wedge\tau}$ is not much bigger than λ either. However, such a control is generally unavailable. For this reason, we reserve for later use another stopping time σ, to be specified below, and make a further splitting of the martingale $f_{n\wedge\sigma\wedge\tau}$:

$$\begin{aligned} f_{n\wedge\sigma\wedge\tau} &= f_{n\wedge\sigma\wedge(\tau-1)} + \big(f_{n\wedge\sigma\wedge\tau} - f_{n\wedge\sigma\wedge(\tau-1)}\big) \\ &= f_{n\wedge\sigma\wedge(\tau-1)} + \mathbf{1}_{\{n\wedge\sigma\geqslant\tau\}} df_\tau =: \widetilde{g}_n + \widetilde{h}_n, \end{aligned}$$

which gives our first Ansatz for g and h. By definition of τ, it is immediate that $\|\widetilde{g}_n\|_\infty \leqslant \lambda$, but the problem is that, since $\tau - 1$ is not a stopping time, $\widetilde{g}$ is not a martingale, and we need to add a correction term. The second stopping time σ will then be chosen so that also this correction term remains under control. We now turn to the details.

The good part g

The difference sequence corresponding to our Ansatz $\widetilde{g}_n$ is

$$d\widetilde{g}_n = \mathbf{1}_{\{\sigma\wedge(\tau-1)\geqslant n\}} df_n = \mathbf{1}_{\{\sigma\geqslant n\}} \mathbf{1}_{\{\tau>n\}} df_n.$$

Let us write $\mathbb{E}_n$ for the conditional expectation with respect to $\mathscr{F}_n$. In order to ensure the martingale difference property $\mathbb{E}_{n-1}(dg_n) = 0$, we set

$$\begin{aligned} dg_n := d\widetilde{g}_n - \mathbb{E}_{n-1}(d\widetilde{g}_n) &= \mathbf{1}_{\{\sigma\geqslant n\}} \big(\mathbf{1}_{\{\tau>n\}} df_n - \mathbb{E}_{n-1}(\mathbf{1}_{\{\tau>n\}} df_n)\big) \\ &= \mathbf{1}_{\{\sigma\geqslant n\}} \big(\mathbf{1}_{\{\tau>n\}} df_n + \mathbb{E}_{n-1}(\mathbf{1}_{\{\tau=n\}} df_n)\big), \end{aligned}$$

where the last step follows from

$$\begin{aligned} 0 = \mathbf{1}_{\{\tau\geqslant n\}} \cdot 0 = \mathbf{1}_{\{\tau\geqslant n\}} \mathbb{E}_{n-1}(df_n) &= \mathbb{E}_{n-1}(\mathbf{1}_{\{\tau\geqslant n\}} df_n) \\ &= \mathbb{E}_{n-1}(\mathbf{1}_{\{\tau>n\}} df_n) + \mathbb{E}_{n-1}(\mathbf{1}_{\{\tau=n\}} df_n). \end{aligned}$$

Let also

$$g_{-\infty} := \widetilde{g}_{-\infty} = f_{-\infty}.$$

Summing up, this gives

$$\begin{aligned} g_n &= g_{-\infty} + \sum_{k\leqslant n} dg_k \\ &= f_{-\infty} + \sum_{k\leqslant n} \mathbf{1}_{\{\sigma\geqslant k\}} \big(\mathbf{1}_{\{\tau>k\}} df_k + \mathbb{E}_{k-1}(\mathbf{1}_{\{\tau=k\}} df_k)\big) \\ &= f_{-\infty} + \sum_{k\leqslant n} \mathbf{1}_{\{\sigma\wedge(\tau-1)\geqslant k\}} df_k + \sum_{k\leqslant n} \mathbf{1}_{\{\sigma\geqslant k\}} \mathbb{E}_{k-1}(\mathbf{1}_{\{\tau=k\}} df_k) \\ &= f_{n\wedge\sigma\wedge(\tau-1)} + \sum_{k\leqslant n\wedge\sigma} \mathbb{E}_{k-1}(\mathbf{1}_{\{\tau=k\}} df_k) \\ &=: \widetilde{g}_n + e_n. \end{aligned}$$

To ensure the boundedness of the error term e_n, we now fix the second stopping time σ as

$$\sigma := \inf \Big\{ n : \sum_{k \leqslant n+1} \mathbb{E}_{k-1}(\mathbf{1}_{\{\tau = k\}} \|df_k\|) > \lambda \Big\}.$$

It is then immediate that

$$\|g_n\|_\infty \leqslant \lambda + \lambda = 2\lambda,$$

and we can also estimate

$$\begin{aligned} \|\widetilde{g}_n\|_1 = \|f_{n\wedge\sigma\wedge(\tau-1)}\|_1 &\leqslant \|\mathbf{1}_{\{\tau>n\}} f_{n\wedge\sigma}\|_1 + \|\mathbf{1}_{\{\tau\leqslant n\}} f_{\sigma\wedge(\tau-1)}\|_1 \\ &\leqslant \|f_{n\wedge\sigma}\|_1 + \|\mathbf{1}_{\{\tau\leqslant n\}}\lambda\|_1 \\ &\leqslant \|f_n\|_1 + \mu(\tau \leqslant n)\lambda \leqslant 2\|f_n\|_1, \end{aligned}$$

using (3.1) for the penultimate step, and for the last step we note that by Doob's inequality (see Theorem 3.2.3),

$$\mu(\tau \leqslant n) = \mu(f_n^\star > \lambda) \leqslant \lambda^{-1}\|f_n\|_1.$$

In order to estimate the error term e_n, we claim that

$$\sum_{k\leqslant n} \|\mathbf{1}_{\{\tau=k\}} df_k\|_1 \leqslant 2\|f_n\|_1. \tag{3.43}$$

Indeed, we have pointwise

$$\|df_\tau\| \leqslant \|f_\tau\| + \|f_{\tau-1}\| \leqslant \|f_\tau\| + \lambda < 2\|f_\tau\|,$$

whereas the left-hand side of (3.43) equals

$$\|\mathbf{1}_{\{\tau\leqslant n\}} df_\tau\|_1 \leqslant 2\|\mathbf{1}_{\{\tau\leqslant n\}} f_\tau\|_1 \leqslant 2\|f_n\|_1,$$

where in the last step we applied (3.1).

Now (3.43) and the contractivity of the conditional expectation yield that

$$\begin{aligned} \|e_n\|_1 &\leqslant \int_S \sum_{k\leqslant n\wedge\sigma} \|\mathbb{E}_{k-1}(\mathbf{1}_{\{\tau=k\}} df_k)\| \,\mathrm{d}\mu \\ &\leqslant \int_S \sum_{k\leqslant n} \|\mathbb{E}_{k-1}(\mathbf{1}_{\{\tau=k\}} df_k)\| \,\mathrm{d}\mu \\ &= \sum_{k\leqslant n} \|\mathbb{E}_{k-1}(\mathbf{1}_{\{\tau=k\}} df_k)\|_1 \\ &\leqslant \sum_{k\leqslant n} \|\mathbf{1}_{\{\tau=k\}} df_k\|_1 \leqslant 2\|f_n\|_1, \end{aligned}$$

and therefore

$$\|g_n\|_1 \leqslant \|\widetilde{g}_n\|_1 + \|e_n\|_1 \leqslant 2\|f_n\|_1 + 2\|f_n\|_1 = 4\|f_n\|_1.$$

This completes the estimation of g_n.

The harmless part h

We have

$$f_{n\wedge\sigma\wedge\tau} = \widetilde{g}_n + \widetilde{h}_n = g_n + \widetilde{h}_n - e_n =: g_n + h_n,$$

where $f_{n\wedge\sigma\wedge\tau}$ and g_n are martingales, and hence h_n is one as well. Explicitly, it is given by

$$h_n = \widetilde{h}_n - e_n = \mathbf{1}_{\{\tau\leqslant n\wedge\sigma\}} df_\tau - \sum_{k\leqslant n\wedge\sigma} \mathbb{E}_{k-1}(\mathbf{1}_{\{\tau=k\}} df_k),$$

and its difference sequence by

$$\begin{aligned} dh_n &= \mathbf{1}_{\{\tau=n\leqslant\sigma\}} df_n - \mathbf{1}_{\{\sigma\geqslant n\}}\mathbb{E}_{n-1}(\mathbf{1}_{\{\tau=n\}} df_n) \\ &= \mathbf{1}_{\{\sigma\geqslant n\}}\big(\mathbf{1}_{\{\tau=n\}} df_n - \mathbb{E}_{n-1}(\mathbf{1}_{\{\tau=n\}} df_n)\big). \end{aligned}$$

Thus, by (3.43) and the contractivity of the conditional expectation, we obtain the required bound

$$\sum_{k\leqslant n} \|dh_k\|_1 \leqslant \sum_{k\leqslant n} 2\|\mathbf{1}_{\{\tau=k\}} df_k\|_1 \leqslant 4\|f_n\|_1.$$

The bad part b

The residual part of f_n is

$$b_n := f_n - (g_n + h_n) = f_n - f_{n\wedge\sigma\wedge\tau},$$

which is obviously a martingale. Moreover, we have

$$\{b_n \neq 0\} \subseteq \{n \neq n\wedge\sigma\wedge\tau\} = \{\sigma\wedge\tau < n\} = \{\sigma < n\} \cup \{\tau < n\},$$

and therefore

$$\mu(b_n^\star > 0) = \mu\Big(\bigcup_{k\leqslant n}\{b_k \neq 0\}\Big) \leqslant \mu(\sigma < n) + \mu(\tau < n),$$

where

$$\begin{aligned} \mu(\sigma < n) &= \mu\Big(\sum_{k\leqslant n} \mathbb{E}_{k-1}(\mathbf{1}_{\{\tau=k\}}\|df_k\|) > \lambda\Big) \\ &\leqslant \lambda^{-1}\Big\|\sum_{k\leqslant n} \mathbb{E}_{k-1}(\mathbf{1}_{\{\tau=k\}}\|df_k\|)\Big\|_1 \leqslant 2\lambda^{-1}\|f_n\|_1 \end{aligned}$$

by Doob's maximal inequality, the contractivity of the conditional expectation, and (3.43). For the other term again by Doob's maximal inequality,

$$\mu(\tau < n) = \mu(f_{n-1}^\star > \lambda) \leqslant \lambda^{-1}\|f_{n-1}\|_1 \leqslant \lambda^{-1}\|f_n\|_1.$$

Combining the estimates yields

$$\mu(b_n^\star > 0) \leqslant 3\lambda^{-1}\|f_n\|_1.$$

This completes the proof of the Gundy decomposition under the extra assumption that $f_{-\infty} < \lambda$ almost everywhere.

Completion of the proof

It remains to consider the case where $A := \{\|f_{-\infty}\| \geqslant \lambda\}$ has positive measure. Note that $A \in \mathscr{F}_{-\infty}$. We split f_n into two new martingales

$$f_n = f_n \mathbf{1}_{\complement A} + f_n \mathbf{1}_A =: f_n^0 + f_n^1.$$

Here $(f_n^0)_{n\in\mathbb{Z}}$ satisfies the assumptions of the case already treated, so we find a decomposition $f_n^0 = g_n^0 + h_n^0 + b_n^0$ with the asserted properties. The decomposition of f_n is then given by $g_n := g_n^0$, $h_n := h_n^0$ and $b_n := b_n^0 + f_n^1$. It only remains to check that b_n satisfies the required bound. To this end, we have

$$\lambda\mu(b_n^\star > 0) \leqslant \lambda\mu(b_n^{0,*} > 0) + \lambda\mu(f_n^{1,*} > 0) \leqslant 3\|f_n^0\|_1 + \lambda\mu(A),$$

where

$$\lambda\mu(A) = \lambda\mu(\|f_{-\infty}\| \geqslant \lambda) \leqslant \|f_{-\infty}\mathbf{1}_A\|_1 \leqslant \|f_n\mathbf{1}_A\|_1,$$

and hence

$$3\|f_n^0\|_1 + \lambda\mu(A) \leqslant 3\|f_n\mathbf{1}_{\complement A}\|_1 + \|f_n\mathbf{1}_A\|_1 \leqslant 3\|f_n\|_1.$$

The Gundy decomposition is now proved in this case as well. □

3.4.b Davis decomposition

The Davis decomposition of f splits a martingale $f = (f_n)_{n\in\mathbb{Z}}$ into two parts, one having increments that are pointwise dominated by a bounded predictable sequence and the other having pointwise absolute summability of its martingale differences. As in the previous section we write $\mathbb{E}_n$ for $\mathbb{E}(\cdot|\mathscr{F}_n)$ for $n \in \mathbb{Z}$ to simplify the notation.

Theorem 3.4.3 (Davis decomposition). *Given a martingale $(f_n)_{n\in\mathbb{Z}}$ with $f_n^\star \in L^p(S)$ for some $p \in [1,\infty)$, there is a decomposition $f_n = g_n + h_n$ into martingales adapted to the same filtration, such that the good part starts at $g_{-\infty} = 0$ and its differences have a predictable majorant:*

$$\|dg_n\| \leqslant v_n \in L^0(\mathscr{F}_{n-1}),$$

the harmless part satisfies $h_{-\infty} = f_{-\infty}$, and its differences are pointwise absolutely summable:

$$\|h_n^\star\|_p \leqslant \Big\| \|h_{-\infty}\| + \sum_{k\leqslant n} \|dh_k\| \Big\|_p \leqslant 9p\|f_n^\star\|_p;$$

moreover, we have the norm estimates

$$\|g_n^\star\|_p \leqslant 10p\|f_n^\star\|_p \leqslant 10pp'\|f_n\|_p,$$
$$\|v_n^\star\|_p \leqslant 8\|f_{n-1}^\star\|_p \leqslant 8p'\|f_{n-1}\|_p.$$

If $p > 1$, then the condition that $f_n^\star \in L^p(S)$ is ensured by $f_n \in L^p(S;X)$, thanks to Doob's inequality.

Proof. We split

$$df_n = df_n \mathbf{1}_{\{\|df_n\| \leqslant 2(df)_{n-1}^\star\}} + df_n \mathbf{1}_{\{\|df_n\| > 2(df)_{n-1}^\star\}} =: y_n + z_n.$$

On $\{\|df_n\| > 2(df)_{n-1}^\star\}$ we have

$$\|df_n\| = 2\|df_n\| - \|df_n\| < 2\|df_n\| - 2(df)_{n-1}^\star \leqslant 2(df)_n^\star - 2(df)_{n-1}^\star,$$

and hence

$$\sum_{k \leqslant n} \|z_k\| \leqslant 2 \sum_{k \leqslant n} ((df)_k^\star - (df)_{k-1}^\star) \leqslant 2(df)_n^\star \leqslant 4 f_n^\star.$$

By dual Doob's inequality (see Proposition 3.2.8), we also have

$$\Big\| \sum_{k \leqslant n} \mathbb{E}_{k-1} \|z_k\| \Big\|_p \leqslant p \Big\| \sum_{k \leqslant n} \|z_k\| \Big\|_p \leqslant 4p \|f_n^\star\|_p.$$

Clearly $(I - \mathbb{E}_{k-1}) z_k$ is a martingale difference, and we may define

$$h_n := h_{-\infty} + \sum_{k \leqslant n} dh_k := f_{-\infty} + \sum_{k \leqslant n} (I - \mathbb{E}_{k-1}) z_k,$$

where the series converges absolutely almost everywhere and satisfies

$$\begin{aligned} \Big\| \|h_{-\infty}\| + \sum_{k \leqslant n} \|dh_k\| \Big\|_p &\leqslant \|f_{-\infty}\|_p + 4\|f_n^\star\|_p + 4p\|f_n^\star\|_p \\ &\leqslant 9p\|f_n^\star\|_p \leqslant 9pp'\|f_n\|_p. \end{aligned}$$

We define $g_n := f_n - h_n$, so in particular $g_{-\infty} = 0$ and

$$\|g_n^\star\|_p \leqslant \|f_n^\star\|_p + \|h_n^\star\|_p \leqslant 10p\|f_n^\star\|_p \leqslant 10pp'\|f_n\|_p.$$

Moreover, we have

$$dg_n = df_n - dh_n = (I - \mathbb{E}_{n-1}) df_n - (I - \mathbb{E}_{n-1}) z_n = (I - \mathbb{E}_{n-1}) y_n,$$

and thus

$$\begin{aligned} \|dg_n\| \leqslant (1 + \mathbb{E}_{n-1})\|y_n\| &\leqslant (1 + \mathbb{E}_{n-1})(2(df)_{n-1}^\star) \\ &= 4(df)_{n-1}^\star \leqslant 8 f_{n-1}^\star =: v_n, \end{aligned}$$

where of course $\|v_n^\star\|_p = 8\|f_{n-1}^\star\|_p \leqslant 8p'\|f_{n-1}\|_p$. □

3.5 Martingale transforms

We now introduce an important class of transformations with far-reaching applications.

3.5.a Basic properties

Throughout this subsection, X and Y are Banach spaces and $(S, \mathscr{A}, \mu)$ is a measure space equipped with a σ-finite filtration $(\mathscr{F}_n)_{n\in\mathbb{Z}}$, which without any loss of generality may be assumed to be *generating* in the sense that $\mathscr{A} = \sigma(\bigcup_{n\in\mathbb{Z}} \mathscr{F}_n)$. To simplify notation, we abbreviate $\mathbb{E}_n f := \mathbb{E}(f|\mathscr{F}_n)$ for $n \in \mathbb{Z}$. For $f \in L^p(S;X)$, $1 \leqslant p < \infty$, we also denote

$$f_{-\infty} := \lim_{n\to-\infty} \mathbb{E}_n f,$$

which exists almost everywhere by Corollary 3.3.7. We have $f_{-\infty} = \mathbb{E}(f|\mathscr{F}_{-\infty})$ if μ is σ-finite on $\mathscr{F}_{-\infty} := \bigcap_{n\in\mathbb{Z}} \mathscr{F}_{-\infty}$, but we prefer the notation $f_{-\infty}$ to emphasise the fact that it may not be a proper conditional expectation in general.

A sequence of X-valued functions $(w_n)_{n\in\{-\infty\}\cup\mathbb{Z}}$ is called *predictable* if each w_n is strongly $\mathscr{F}_{n-1}$-measurable, where of course $-\infty - 1 = -\infty$. A sequence of $\mathscr{L}(X,Y)$-valued functions $(v_n)_{n\in\{-\infty\}\cup\mathbb{Z}}$ is called *strongly predictable* if for each $x \in X$, the mapping $s \mapsto v_n(s)x$ is strongly $\mathscr{F}_{n-1}$-measurable.

We fix a strongly predictable sequence $(v_n)_{n\in\{-\infty\}\cup\mathbb{Z}}$ with the additional property

$$\|v_n\|_{L^\infty_{\rm so}(S;\mathscr{L}(X,Y))} := \sup_{\|x\|\leqslant 1} \|v_n x\|_{L^\infty(S;X)} < \infty, \quad \forall n \in \{-\infty\} \cup \mathbb{Z}. \tag{3.44}$$

Definition 3.5.1 (Martingale transform as an operator on test functions). *Let $1 \leqslant p < \infty$ and suppose that $f \in L^p(S;X)$ has a finitely non-zero difference sequence. The* martingale transform *of f by the sequence v is defined as*

$$T_v f := v_{-\infty} f_{-\infty} + \sum_{n\in\mathbb{Z}} v_n df_n, \tag{3.45}$$

where, as always, $df_n := \mathbb{E}_n f - \mathbb{E}_{n-1} f$.

In Proposition 3.5.9 below, we shall see how to extend (3.45) to general functions $f \in L^p(S;X)$, $p \in [1,\infty)$, under a boundedness condition of the transform on its initial test function domain. For $p \in (1,\infty)$, this is relatively routine, since the space of $f \in L^p(S;X)$ with a finite difference sequence is dense in $L^p(S;X)$ in this range by Theorem 3.3.8, but the case of $p = 1$ requires some care, since the density may fail in this case.

When the sequence $(v_n)_{n\in\{-\infty\}\cup\mathbb{Z}}$ is fixed we will often write T instead of T_v. It is of interest to determine conditions under which the densely defined operator T may be extended boundedly to all of $L^p(S;X)$, or to some other function spaces. We denote

$$\|T\|_{L^p(S;X)\to L^p(S;Y)}$$
$$:= \sup\Big\{\frac{\|Tf\|_{L^p(S;Y)}}{\|f\|_{L^p(S;X)}} : \ f \not\equiv 0,\ f = f_{-\infty} + \sum_{n\in\mathbb{Z}} df_n \text{ finitely non-zero}\Big\},$$

and we say that T is *bounded* from $L^p(S;X)$ into $L^p(S;Y)$ if this extended real number is finite. The weak-type norms $\|T\|_{L^p(S;X)\to L^{p,\infty}(S;Y)}$ and the corresponding boundedness notions are defined similarly, just replacing $\|Tf\|_{L^p(S;Y)}$ by $\|Tf\|_{L^{p,\infty}(S;Y)}$.

The *maximally truncated martingale transform* is the sub-linear operator

$$T^\star f := (Tf)^\star := \sup_{n\in\mathbb{Z}} \Big\| v_{-\infty} f_{-\infty} + \sum_{k\leqslant n} v_k df_k \Big\|_Y$$

which, as suggested by the notation, is just the Doob maximal function of the martingale $g_n = v_{-\infty}f_{-\infty} + \sum_{k\leqslant n} v_k df_k$. The latter is a martingale by the following lemma:

Lemma 3.5.2. *Let $\mathscr{G}$ be a σ-finite sub-σ-algebra of $\mathscr{F}$. Let $g : S \to \mathscr{L}(X,Y)$ be such that for all $x \in X$, gx is strongly $\mathscr{G}$-measurable and*

$$\|g\|_{L^\infty_{\rm so}(S;\mathscr{L}(X,Y))} := \sup_{\|x\|\leqslant 1} \|gx\|_{L^\infty(S;Y)} < \infty.$$

Then for all $f : S \to X$ which are σ-integrable over $\mathscr{G}$, we have $\mathbb{E}(gf|\mathscr{G}) = g\mathbb{E}(f|\mathscr{G})$.

Note that gf is σ-integrable over $\mathscr{G}$. Indeed, if is a μ-simple function then $\|gf\|_{L^1(S;Y)} \leqslant C\|f\|_{L^1(S;X)}$ and this extends to $f \in L^1(S;X)$ by approximation. Now for general f one can apply the previous observation to $\mathbf{1}_A f \in L^1(S;X)$ with $A \in \mathscr{G}$. For scalar-valued functions g the lemma is an immediate consequence of Proposition 2.6.31.

Proof. First let $f := \mathbf{1}_A \otimes x$, where $x \in X$ and $A \in \mathscr{A}$ has finite measure. By Proposition 2.6.31,

$$\mathbb{E}(gf|\mathscr{G}) = \mathbb{E}(g\mathbf{1}_A \otimes x|\mathscr{G}) = g\mathbb{E}(\mathbf{1}_A|\mathscr{G}) \otimes x = g\mathbb{E}(f|\mathscr{G}).$$

By linearity this extends to all μ-simple functions $f : S \to X$ and by density to all $f \in L^1(S;X)$. The general case then follows via (2.28). ☐

Lemma 3.5.3 (Adjoint of a martingale transform). *Suppose that also the adjoint sequence v^* is strongly predictable with $\|v_n^*\|_{L^\infty_{\rm so}(S;\mathscr{L}(Y^*,X^*))} < \infty$. Then*

$$\langle T_v f, g\rangle = \langle f, T_{v^*} g\rangle$$

for all $f \in L^p(S;X)$ and $g \in L^{p'}(S;Y^*)$ with finitely non-zero difference sequences, $1 \leqslant p \leqslant \infty$.

Here, of course, $v_n^*(s) := (v_n(s))^*$. Since $L^p(S;X)$ and $L^{p'}(S;Y^*)$ are norming for $L^{p'}(S;X^*)$ and $L^p(S;Y)$, and since functions with a finitely non-zero difference sequence are dense, it follows that

$$\|T_v\|_{\mathscr{L}(L^p(S;X),L^p(S;Y))} = \|T_{v^*}\|_{\mathscr{L}(L^{p'}(S;Y^*),L^{p'}(S;X^*))}$$

for all such v. We refer to T_{v^*} as the *formal adjoint* of T_v and also use the alternative notation $T^* := T_v^* := T_{v^*}$, particularly when v is not indicated in the notation.

Proof. This is a direct computation

$$\begin{aligned}\langle T_v f, g\rangle &= \Big\langle v_{-\infty} f_{-\infty} + \sum_{n\in\mathbb{Z}} v_n df_n, g_{-\infty} + \sum_{m\in\mathbb{Z}} dg_m \Big\rangle \\ &= \langle v_{-\infty} f_{-\infty}, g_{-\infty}\rangle + \sum_{n\in\mathbb{Z}} \langle v_n df_n, dg_n\rangle \\ &= \langle f_{-\infty}, v_{-\infty}^* g_{-\infty}\rangle + \sum_{n\in\mathbb{Z}} \langle df_n, v_n^* dg_n\rangle \\ &= \Big\langle f_{-\infty} + \sum_{n\in\mathbb{Z}} df_n, v_{-\infty}^* g_{-\infty} + \sum_{m\in\mathbb{Z}} v_m^* dg_m \Big\rangle = \langle f, T_{v^*} g\rangle,\end{aligned}$$

where the cancellation of the cross terms in the second and second-to-last steps follows, say for $n < m$ and $dh_n = v_n df_n$, from Proposition 2.6.31 and

$$\begin{aligned}\langle dh_n, dg_m\rangle = \langle dh_n, (\mathbb{E}_m - \mathbb{E}_{m-1})g\rangle &= \langle (\mathbb{E}_m - \mathbb{E}_{m-1}) dh_n, g\rangle \\ &= \langle dh_n - dh_n, g\rangle = 0\end{aligned}$$

and other similar identities. □

The main result of this section is an extrapolation theorem which shows that various norm bounds for a martingale transform are equivalent. For this statement, it is convenient to introduce the *martingale Hardy spaces* $H^p(S;X)$, which are defined as

$$H^p(S;X) := \{f \in L^p(S;X) : f^\star \in L^p(S)\}, \qquad \|f\|_{H^p} := \|f^\star\|_{L^p},$$

where $f^\star = \sup_{n\in\mathbb{Z}} \|\mathbb{E}(f|\mathscr{F}_n)\|$ is the Doob maximal function of the martingale generated by f, relative to the given filtration $(\mathscr{F}_n)_{n\in\mathbb{Z}}$. The space $H^p(S;X)$ certainly depends on the chosen filtration, although we do not indicate it explicitly in order not to overburden the notation.

Since the filtration is assumed to generate $\mathscr{A}$, we have $\|f\| \leqslant f^\star$ pointwise almost everywhere, and therefore

$$\|f\|_{L^p} \leqslant \|f\|_{H^p}.$$

Moreover, for $p \in (1,\infty]$, the Hardy space $H^p(S;X)$ coincides with $L^p(S;X)$, thanks to Doob's inequality, but it is equipped with a different norm, as

$$\|f\|_{H^p} \leqslant p'\|f\|_{L^p}.$$

For $p = 1$, however, $H^1(S;X)$ is a strict subspace of $L^1(S;X)$ and plays an important role as a substitute space for some estimates that fail in $L^1(S;X)$. The theory of (martingale) Hardy spaces is rich and interesting in its own right, but we shall not develop it in any detail here, aside from its appearance in the following theorem.

Theorem 3.5.4 (Extrapolation of martingale transform inequalities). *Let T be a fixed martingale transform associated with a strongly predictable $\mathscr{L}(X,Y)$-valued sequence v, and let $T^\star$ be the corresponding maximally truncated martingale transform. The following assertions are equivalent:*

(1) *T is bounded from $H^p(S;X)$ to $L^{p,\infty}(S;Y)$ for some $p \in [1,\infty)$;*
(2) *$T^\star$ is bounded from $L^p(S;X)$ to $L^p(S)$ for all $p \in (1,\infty)$;*
(3) *$T^\star$ is bounded from $H^1(S;X)$ to $L^1(S)$;*
(4) *$T^\star$ is bounded from $L^1(S;X)$ to $L^{1,\infty}(S)$.*

Moreover, we have the quantitative estimates

$$\|T^\star\|_{H^p\to L^p} \leqslant 700\cdot\Big(1+\frac{p}{q}\Big)\cdot p\cdot\|T\|_{H^q\to L^{q,\infty}}, \quad p,q\in[1,\infty), \tag{3.46}$$

$$\|T^\star\|_{L^1\to L^{1,\infty}} \leqslant 2000\cdot\|T\|_{L^q\to L^{q,\infty}}, \quad q\in[1,\infty), \tag{3.47}$$

and, if also the adjoint sequence v^* is strongly predictable, then

$$\|T\|_{L^p\to L^p} \leqslant 1400\cdot pp'\cdot\|T\|_{L^q\to L^q}, \quad p,q\in(1,\infty). \tag{3.48}$$

If the underlying filtration is a Paley–Walsh filtration, then (3.48) may be improved to

$$\|T\|_{L^p\to L^p} \leqslant 100\cdot\Big(\frac{p}{q}+\frac{p'}{q'}\Big)\|T\|_{L^q\to L^q}, \quad p,q\in(1,\infty). \tag{3.49}$$

Remark 3.5.5. The estimates (3.48) and (3.49) have the correct order in the sense that, for many martingale transforms, the L^p-norm grows at the rate $pp' = p+p'$ as $p\to 1$ or $p\to\infty$. The bound (3.49) captures this behaviour even more precisely, if p is close to q. Moreover, if we start from the estimate for $\|T\|_{L^q\to L^q}$ for, say, a large q, the bound for $\|T\|_{L^p\to L^p}$ for large p has the order $p/q\cdot\|T\|_{L^q\to L^q} \eqsim p$ if $\|T\|_{L^q\to L^q} \eqsim q$. Note that if $q = 2$, then the multiplicative constant becomes $50(p+p')$.

The three implications (2)⇒(1), (3)⇒(1), and (4)⇒(1) are trivial. The main effort will consist of showing that (1)⇒((2) and (3)), and that (1)⇒(4).

Before dwelling into the lengthy details of the proof of Theorem 3.5.4, we present some simple examples that illustrate the use of the theorem. In the first example we need the orthogonality of martingale differences with values in Hilbert spaces. As this will be used several times, we record it as:

Proposition 3.5.6 (Orthogonality of martingale differences). *Let X be a Hilbert space. If f is in $L^2(S;X)$, then all of its martingale differences are orthogonal. In particular,*

$$\|f\|_{L^2(S;X)} = \Big(\|f_{-\infty}\|^2_{L^2(S;X)} + \sum_{n\in\mathbb{Z}} \|df_n\|^2_{L^2(S;X)}\Big)^{1/2}.$$

Proof. The orthogonality follows for $m < n$ from $\mathbb{E}(df_n|\mathscr{F}_m) = 0$ and

$$\int_S (df_n|df_m)\,\mathrm{d}\mu = \int_S \mathbb{E}((df_n|df_m)|\mathscr{F}_m)\,\mathrm{d}\mu = \int_S (\mathbb{E}(df_n|\mathscr{F}_m)|df_m)\,\mathrm{d}\mu = 0,$$

using Proposition 2.6.31 in the middle step. The same holds if one replaces df_m by $f_{-\infty}$. □

Example 3.5.7. Let X and Y be Hilbert spaces and let $(v_n)_{n\in\mathbb{Z}\cup\{-\infty\}}$ in $\mathscr{L}(X,Y)$ be strongly predictable and uniformly bounded. Then the corresponding martingale transform $T = T_v$ is L^2-bounded with

$$\|T\|_{\mathscr{L}(L^2(S;X),L^2(S;Y))} \leqslant \sup_{n\in\{-\infty\}\cup\mathbb{Z}} \|v_n\|_\infty.$$

Indeed, this is immediate from Proposition 3.5.6 and the pointwise bounds $\|v_{-\infty}f_{-\infty}\| \leqslant \|v_{-\infty}\|_\infty\|f_{-\infty}\|$ and $\|v_n df_n\| \leqslant \|v_n\|_\infty\|df_n\|$. By Theorem 3.5.4 the L^2-boundedness of T also implies its L^p-boundedness for every $p\in(1,\infty)$. In Corollary 4.5.15 we will obtain a sharp bound for its L^p-norm.

Example 3.5.8. Let $X = \ell^q$ with $q\in(1,\infty]$ and $Y = \ell^q(\ell^\infty(\mathbb{Z}))$ and let $p\in(1,\infty)$. For each $n\in\mathbb{Z}$, define $v_n\in\mathscr{L}(\ell^q, Y)$ by

$$v_n(x) := \sum_{k\geqslant n} x\otimes e_k \quad\text{and}\quad v_{-\infty}(x) := \sum_{k\in\mathbb{Z}} x\otimes e_k, \qquad x\in\ell^q,$$

with $e_k\in\ell^\infty(\mathbb{Z})$ the standard unit vectors. For all $f = (f^{(j)})_{j\in\mathbb{Z}}\in L^p(S;\ell^q)$ with finitely non-zero difference sequence relative to some given filtration on S we have, pointwise on S,

$$\begin{aligned}\|Tf\|_Y &= \Big\|\sum_{k\in\mathbb{Z}} f_{-\infty}\otimes e_k + \sum_{k\in\mathbb{Z}}\sum_{n\leqslant k} df_n\otimes e_k\Big\|_Y \\ &= \Big\|\sum_{k\in\mathbb{Z}} f_k\otimes e_k\Big\|_Y = \|(f^{(j)\star})_{j\in\mathbb{Z}}\|_{\ell^q},\end{aligned}$$

where $f_k = (f_k^{(j)})_{j\in\mathbb{Z}} := \mathbb{E}(f|\mathscr{F}_k)$ and $f_n^{(j)\star} := \sup_{k\leqslant n}|((f_k^{(j)}(s))|$. From this identity it is clear that the associated martingale transform T is bounded from $L^p(S;\ell^q)$ into $L^p(S;Y)$ of norm $\leqslant C$ if and only if $\|(f^{(j)\star})_{j\in\mathbb{Z}}\|_{L^p(S;\ell^q)} \leqslant C\|(f^{(j)})_{j\in\mathbb{Z}}\|_{L^p(S;\ell^q)}$ for all $(f^{(j)})_{j\in\mathbb{Z}}\in L^p(S;\ell^q)$. As we have seen in Theorem 3.2.7, this bound holds for $q\in(1,\infty]$ and $p\in(1,\infty)$. The case $p = q\in(1,\infty)$

is a simple consequence of the scalar case of Doob's maximal inequality (see Theorem 3.2.2). The boundedness in the case $p \neq q$ with $p \in (1,\infty)$ and $q \in (1,\infty]$ follows from Theorem 3.5.4.

This provides an alternative proof of the ℓ^q-version of Doob's inequality (Theorem 3.2.7) which, as we have seen, is a precursor to the Fefferman–Stein theorem (Theorem 3.2.28).

Theorem 3.5.4 also provides a weak L^1-estimate. By the same method of proof one can include a weight (as was done in Theorem 3.2.3).

The next result provides a representation for the extension of the martingale transform T to a bounded operator on all of $L^p(S;X)$.

Proposition 3.5.9 (Martingale transform as an operator on $L^p(S;X)$). *Suppose that a martingale transform T satisfies the equivalent conditions of Theorem 3.5.4. Let $p \in [1,\infty)$ and $f \in L^p(S;X)$. Then the series*

$$v_{-\infty} f_{-\infty} + \sum_{n\in\mathbb{Z}} v_n df_n \tag{3.50}$$

converges pointwise almost everywhere.

If $p \in (1,\infty)$, the convergence also takes place in $L^p(S;Y)$, and the limit coincides with $\widetilde{T}_p f$, where $\widetilde{T}_p : L^p(S;X) \to L^p(S;Y)$ is the unique extension of T, initially defined on functions with a finite martingale difference sequence, to a bounded linear operator.

Proof. For any $p \in [1,\infty)$, $f \in L^p(S;X)$, and integers $m < n$, let us denote

$$T_{m,n} f := \sum_{k=m+1}^{n} v_k df_k = T(f_n - f_m),$$

where the identity follows from the fact that $f_n - f_m$ is a function with a finite difference sequence, on which T is initially defined.

If $p \in (1,\infty)$, then

$$\|T_{m,n} f\|_p = \|T(f_n - f_m)\|_p \leqslant \|T\|_{L^p \to L^p} \|f_n - f_m\|_p,$$

which tends to 0 as $m, n \to \infty$ or $m, n \to -\infty$ by Theorem 3.3.8. Thus $T_{m,n} f$ has a limit in $L^p(S;Y)$ as $n \to \infty$ and $m \to -\infty$. Since $f_{-\infty} + (f_n - f_m) \to f$ in $L^p(S;X)$ as $n \to \infty$ and $m \to -\infty$, it follows that

$$\widetilde{T}_p f = \lim_{\substack{n\to+\infty \\ m\to-\infty}} T[f_{-\infty} + (f_n - f_m)] = v_{-\infty} f_{-\infty} + \lim_{\substack{n\to+\infty \\ m\to-\infty}} T_{m,n} f,$$

so that $\widetilde{T}_p f$ is given by the $L^p(S;Y)$ limit of the series (3.50).

Next let $p \in [1,\infty)$. To prove pointwise convergence of $T_{m,n} f$, we first observe that

$$
\begin{aligned}
&\Big\| \sup_{|m|,|n|\leqslant N} \|T_{m,n}f\|_Y \Big\|_{L^{p,\infty}(S)} \\
&= \Big\| \sup_{|m|,|n|\leqslant N} \|T(f_n - f_{-N}) - T(f_m - f_{-N})\|_Y \Big\|_{L^{p,\infty}(S)} \\
&\leqslant 2\|T^\star(f_N - f_{-N})\|_{L^{p,\infty}(S)} \leqslant 4\|T^\star\|_{L^p(S;X)\to L^{p,\infty}(S)}\|f\|_{L^p(S;X)},
\end{aligned}
$$

and then, by Fatou's lemma,

$$
\Big\| \sup_{m,n\in\mathbb{Z}} \|T_{m,n}f\|_Y \Big\|_{L^{p,\infty}(S)} \leqslant 4\|T^\star\|_{L^p(S;X)\to L^{p,\infty}(S)}\|f\|_{L^p(S;X)}. \tag{3.51}
$$

Hence, by Proposition 2.3.6 and Remark 2.3.7, to show the pointwise convergence of $T_{m,n}f$ for all $f \in L^p(S;X)$, it suffices to show this for all f in a dense subspace. The mentioned convergence is trivial if f has a finite difference sequence, and such functions are dense in $L^p(S;X)$ for $p \in (1,\infty)$ by Theorem 3.3.8, so the proof in this case is complete.

Concerning $p = 1$, the finite difference sequences may fail to be dense, so we need a little additional step: Note that we have already settled the case $p \in (1,\infty)$, and the space $L^1(S;X) \cap L^p(S;X)$ is dense in $L^1(S;X)$. Thus, the pointwise convergence of $T_{m,n}f$ for all $f \in L^1(S;X) \cap L^p(S;X)$, together with the bound (3.51) for $p = 1$ and Proposition 2.3.6 with Remark 2.3.7 completes the proof. □

Necessity of bounded transforming sequences

Before turning to the main implications of Theorem 3.5.4, there is a technical issue that we would like to clean out of the way. Namely, in some intermediate estimates of Tf, not only the operator norm of T but also the size of the transforming sequence v makes an appearance, and we would like to absorb everything into bounds involving the norm of T only. To this effect, we prove the following:

Proposition 3.5.10. *Given a martingale transform T, it is possible to redefine its transforming sequence v in a way that does not change the operator T, so that the following holds:*

If, for some $p \in [1,\infty)$ and all martingales f with a finitely non-zero difference sequence, we have

$$
\|Tf\|_{L^{p,\infty}} \leqslant N\|f\|_{H^p}, \quad \textit{respectively} \quad \|Tf\|_{L^p} \leqslant N\|f\|_{H^p},
$$

then the redefined sequence v satisfies

$$
\|v_n\|_{L^\infty_{\mathrm{so}}} \leqslant 2^{1/p}N, \quad \textit{respectively} \quad \|v_n\|_{L^\infty_{\mathrm{so}}} \leqslant N,
$$

for all $n \in \mathbb{Z} \cup \{-\infty\}$. The same result holds with H^p replaced by L^p.

Here notations are as in (3.44).

In the subsequent analysis, we will always assume that we are using the redefined sequence v from the beginning. To some extent, Proposition 3.5.10 is a mere curiosity, as almost all concrete examples of martingale transforms that we ever encounter have norm at least 1 for some trivial reasons, as well as $\|v_n\|_{L^\infty_{\mathrm{so}}} \leqslant 1$. With this in mind, the proposition may be easily taken for granted, and the rest of this subsection skipped by an impatient reader, but we nevertheless include the proof for completeness.

In order to explain the redefinition of v_n, we need the following notions: A set $A \in \mathscr{A}$ will be called *non-trivial* if both A and its complement have positive measure. If $\mathscr{F} \subseteq \mathscr{G} \subseteq \mathscr{A}$ are sub-σ-algebras, a set $F \in \mathscr{F}$ is said to *split* in $\mathscr{G}$ if there exist disjoint sets $G, G' \in \mathscr{G} \setminus \overline{\mathscr{F}}$ such that $F = G \cup G'$. Here $\overline{\mathscr{F}}$ is defined as the set of all $A \in \mathscr{A}$ such that $\mu(A \Delta F) = 0$ for some $F \in \mathscr{F}$. In this situation, each of the of sets F, G and G' is non-trivial.

Let us return now to the martingale transforms

$$Tf := v_{-\infty} f_{-\infty} + \sum_{n \in \mathbb{Z}} v_n df_n,$$

with notations and assumptions as in (3.45).

Lemma 3.5.11 (Redefinition of v_n). *Without affecting the operator T, we may set*

- $v_n = 0$ *on every $A \in \mathscr{F}_{n-1}$ which does not split in $\mathscr{F}_n$;*
- $v_{-\infty} = 0$ *on the purely infinite part of $\mathscr{F}_{-\infty}$ (cf. Proposition A.1.4).*

Proof. Any martingale difference df_n must vanish on such a set, so that the value of $v_n df_n$ is unaffected by the value of v_n on A. Similarly, the function $f_{-\infty}$ vanishes on the purely infinite part of $\mathscr{F}_{-\infty}$. $\square$

It is for the v so re-defined that we prove the estimates of Proposition 3.5.10. The proof of the proposition relies on:

Lemma 3.5.12 (Abundance of martingale differences). *Let $\mathscr{F} \subseteq \mathscr{G} \subseteq \mathscr{A}$ be σ-finite sub-σ-algebras and suppose that the set $F \in \mathscr{F}$ splits in $\mathscr{G}$. Then for every $\delta > 0$ there exists a non-zero real-valued measurable function $d \in L^\infty(S, \mathscr{F})$, supported in F, such that*

$$\mathbb{E}(d|\mathscr{F}) = 0, \quad \|d\|_{L^p} \leqslant (2^{1/p} + \delta)\|d\|_{L^{p,\infty}} \quad \forall 1 \leqslant p < \infty.$$

Proof. Let $G, G' \in \mathscr{G} \setminus \overline{\mathscr{F}}$ be disjoint sets such that $F = G \cup G'$. Both G and G' are non-trivial.

If $\mu(F) = \infty$, by the σ-finiteness of $\mathscr{F}$ we can find an exhaustion $F = \bigcup_{i \geqslant 1} F_i$ with sets $F_1 \subseteq F_2 \subseteq \ldots$ in $\mathscr{F}$ of finite measure. Pick $i_0 \in \mathbb{N}$ so large that $F_{i_0} \cap G$ and $F_{i_0} \cap G'$ are both non-trivial. There exists an index $i \geqslant i_0$ such that $F_i \cap G \in \mathscr{G} \setminus \overline{\mathscr{F}}$. Indeed, if not, this would lead to the contradiction

$G = \bigcup_{i \geqslant i_0}(F_i \cap G) \in \overline{\mathscr{F}}$. As a consequence we also have $F_i \cap G' \in \mathscr{G} \setminus \overline{\mathscr{F}}$. This argument shows that there is no loss of generality in assuming that $\mu(F) < \infty$.

We claim that

$$\mu(\{\mathbb{E}_{\mathscr{F}} \mathbf{1}_G \in (0,1)\}) > 0. \tag{3.52}$$

Since $g := \mathbb{E}_{\mathscr{F}} \mathbf{1}_G$ takes its values in $[0,1]$, we have $g = \mathbf{1}_{\{g=1\}} + \mathbf{1}_{\{g \in (0,1)\}} g$, and the claim is equivalent to

$$\mathbb{E}_{\mathscr{F}} \mathbf{1}_G \neq \mathbf{1}_{G_1}, \qquad G_1 := \{\mathbb{E}_{\mathcal{F}} \mathbf{1}_G = 1\} \in \mathscr{F}.$$

Assume to the contrary that the two functions are equal almost everywhere. Then, in fact,

$$\mathbb{E}_{\mathscr{F}} \mathbf{1}_G = \mathbf{1}_{G_1} = \mathbf{1}_{G_1}^2 = \mathbf{1}_{G_1} \mathbb{E}_{\mathscr{F}} \mathbf{1}_G = \mathbb{E}_{\mathscr{F}}(\mathbf{1}_{G_1 \cap G}),$$

and integrating over the underlying space leads to $\mu(G) = \mu(G_1) = \mu(G_1 \cap G)$. This easily implies that $\mu(G \Delta G_1) = 0$, which contradicts the fact that $G \notin \overline{\mathscr{F}}$.

Thus we have verified (3.52), and by symmetry also $\mu(\{\mathbb{E}_{\mathscr{F}} \mathbf{1}_{G'} \in (0,1)\}) > 0$. Since $\mathbb{E}_{\mathscr{F}} \mathbf{1}_G + \mathbb{E}_{\mathscr{F}} \mathbf{1}_{G'} = \mathbb{E}_{\mathscr{F}} \mathbf{1}_F = \mathbf{1}_F$, after permuting the roles of G and G' if necessary, we may assume that $\mu(\{\mathbb{E}_{\mathscr{F}} \mathbf{1}_G \in [\frac{1}{2}, 1)\}) > 0$. Dividing the interval $[\frac{1}{2}, 1)$ into finitely many sub-intervals of small length, by the pigeonhole principle we deduce that $\mu(\{\mathbb{E}_{\mathscr{F}} \mathbf{1}_G \in [a,b)\}) > 0$ for some $\frac{1}{2} \leqslant a < b \leqslant 1$, where $b - a$ can be chosen smaller than any preassigned $\varepsilon > 0$.

We then set $D := \{\mathbb{E}_{\mathscr{F}} \mathbf{1}_G \in [a,b)\} \in \mathscr{F}$, and

$$d := \mathbf{1}_D(\mathbf{1}_G - \mathbb{E}_{\mathscr{F}} \mathbf{1}_G) = \mathbf{1}_{D \cap G}(1 - \mathbb{E}_{\mathscr{F}} \mathbf{1}_G) - \mathbf{1}_{D \setminus G} \mathbb{E}_{\mathscr{F}} \mathbf{1}_G,$$

so that $\mathbb{E}_{\mathscr{F}} d = \mathbf{1}_D \mathbb{E}_{\mathscr{F}}(\mathbf{1}_G - \mathbb{E}_{\mathscr{F}} \mathbf{1}_G) = 0$, and d is indeed a martingale difference. The vanishing integral also implies that

$$\mu(D \cap G)(1-a) \leqslant \int_{D \cap G} (1 - \mathbb{E}_{\mathscr{F}} \mathbf{1}_G) \, \mathrm{d}\mu = \int_{D \setminus G} \mathbb{E}_{\mathscr{F}} \mathbf{1}_G \leqslant \mu(D \setminus G) b.$$

Moreover,

$$\|d\|_{L^p}^p \leqslant \mu(D \cap G)(1-a)^p + \mu(D \setminus G) b^p, \quad \|d\|_{L^{p,\infty}}^p \geqslant \mu(D \setminus G) a^p,$$

and thus

$$\begin{aligned} \Big(\frac{\|d\|_{L^p}}{\|d\|_{L^{p,\infty}}}\Big)^p &\leqslant \frac{\mu(D \cap G)}{\mu(D \setminus G)} \Big(\frac{1-a}{a}\Big)^p + \Big(\frac{b}{a}\Big)^p \\ &\leqslant \frac{b}{1-a} \Big(\frac{1-a}{a}\Big)^p + \Big(\frac{b}{a}\Big)^p \leqslant \Big(\frac{1-a}{a}\Big)^{p-1} + 1 + \eta \leqslant 2 + \eta, \end{aligned}$$

where the second-to-last bound holds provided that $b < a + \varepsilon$ for a sufficiently small $\varepsilon = \varepsilon(\eta)$, and the last one follows by maximising $(1-a)/a$ over $a \in [\frac{1}{2}, 1]$. The statement of the lemma follows by taking the pth root and observing that $(2+\eta)^{1/p} < 2^{1/p} + \delta$ if η is small enough. $\square$

Now we are ready for:

Proof of Proposition 3.5.10. Fix $n \in \mathbb{Z}$ and a vector $x \in X$ of norm one (we may of course assume that $\dim(X) \geqslant 1$). It is enough to show that $\|v_n x\|_\infty \leqslant (2^{\frac{1}{p}} + \delta)N$ for an arbitrary $\delta > 0$. In order to do this let $\lambda > 0$ be arbitrary. Consider the set $A := \{\|v_n x\| > \lambda\}$ which belongs to $\mathscr{F}_{n-1}$. It suffices to show that $\lambda \leqslant (2^{\frac{1}{p}} + \delta)N$ whenever $\mu(A) > 0$. If $\mu(A) > 0$ then A splits in $\mathscr{F}_n$. Indeed, otherwise by our convention (see Lemma 3.5.11) $v_n = 0$ on A and this would imply $A = \varnothing$. By Lemma 3.5.12, there is a non-trivial (real-valued) function df_n in $L^\infty(S;\mathscr{F}_n)$, supported on A, such that $\mathbb{E}(df_n|\mathscr{F}_{n-1}) = 0$ and $\|df_n\|_{L^p} \leqslant (2^{1/p} + \delta)\|df_n\|_{L^{p,\infty}}$. But then

$$\begin{aligned} \lambda\|df_n\|_{L^{p,\infty}} &\leqslant \|v_n x\, df_n\|_{L^{p,\infty}} = \|T(x\, df_n)\|_{L^{p,\infty}} \\ &\leqslant N\|x\, df_n\|_{H^p} = N\|x\, df_n\|_{L^p} \leqslant (2^{1/p} + \delta)N\|df_n\|_{L^{p,\infty}}. \end{aligned} \tag{3.53}$$

Since df_n is non-trivial, we may divide it out to conclude that $\lambda \leqslant (2^{\frac{1}{p}} + \delta)N$, and the result follows.

For $n = -\infty$, the argument is similar but easier: If $\{\|v_{-\infty}x\| > \lambda\} \in \mathscr{F}_{-\infty}$ is a non-zero set, take some non-zero subset $A \in \mathscr{F}_{-\infty}$ of finite measure, and consider the function $f_{-\infty} := \mathbf{1}_A$. Then

$$\begin{aligned} \lambda\|f_{-\infty}\|_{L^{p,\infty}} &\leqslant \|v_{-\infty}x\, f_{-\infty}\|_{L^{p,\infty}} = \|T(x\, f_{-\infty})\|_{L^{p,\infty}} \\ &\leqslant N\|x\, f_{-\infty}\|_{L^p} = N\mu(A)^{1/p} = N\|f_{-\infty}\|_{L^{p,\infty}}. \end{aligned}$$

As before, we find that $\lambda \leqslant N$ and hence $\|v_{-\infty}x\|_\infty \leqslant N$.

To prove the final assertion we replace $L^{p,\infty}$ by L^p in the assumptions and in all the computations. It is then clear that the step in (3.53) that introduced the factor $2^{1/p} + \delta$ becomes redundant. □

3.5.b Extrapolation of L^p-inequalities

This subsection is concerned with the proof of the implications (1) ⇒((2) and (3)) of Theorem 3.5.4, that is, the fact that a single weak-type L^q-bound already implies the strong L^p-bounds for all $p \in (1,\infty)$ as well as the strong end-point estimate on H^1. Quantitatively, we establish the bounds (3.46), (3.48) and (3.49). We proceed in several steps, which also reveal some additional information and relations between the various bounds that have not been stated in Theorem 3.5.4. We start with a technical consequence of the bound (1) of Theorem 3.5.4, which has an additional free parameter with respect to which we may optimise in the subsequent estimates.

We recall from Definition 3.1.13 the notation ${}^k f_n = f_n - f_{n\wedge k}$.

Lemma 3.5.13. *Suppose that, for some $q \in [1,\infty)$, a martingale transform T is bounded from $H^q(S;X)$ to $L^{q,\infty}(S;Y)$ with norm K. Then, for all $\alpha > 0$ it satisfies*

$$\frac{\mu(F \cap \{\|T({}^k f)\| > \alpha \|{}^k f\|_\infty\})}{\mu(F)} \leqslant \Big(\frac{K}{\alpha}\Big)^q \quad \forall F \in \mathscr{F}_k,\ k \in \mathbb{Z} \cup \{-\infty\},$$

where f is a martingale with finitely non-zero difference sequence.

Of course the bound is only interesting for $\alpha > K$ so that $K/\alpha < 1$.

Proof. Observe that

$$\mathbf{1}_F T({}^k f) = \mathbf{1}_F \sum_{n>k} v_n df_n = \sum_{n>k} v_n d(\mathbf{1}_F f)_n = T({}^k(\mathbf{1}_F f)) = T(\mathbf{1}_F {}^k f)$$

so that

$$\begin{aligned} \mu\big(F \cap \{\|T({}^k f)\| > \alpha \|{}^k f\|_\infty\}\big) &\leqslant \mu\big(\|T(\mathbf{1}_F {}^k f)\| > \alpha \|{}^k f\|_\infty\big) \\ &\leqslant \Big(\frac{1}{\alpha \|{}^k f\|_\infty} \|T(\mathbf{1}_F {}^k f)\|_{L^{q,\infty}}\Big)^q \\ &\leqslant \Big(\frac{K}{\alpha \|{}^k f\|_\infty} \|\mathbf{1}_F {}^k f\|_{H^q}\Big)^q. \end{aligned} \tag{3.54}$$

We note that $\mathbb{E}_n(\mathbf{1}_F {}^k f) = \mathbf{1}_F \mathbb{E}_n({}^k f)$ for $n > k$ and

$$\mathbb{E}_n(\mathbf{1}_F {}^k f) = \mathbb{E}_n \mathbb{E}_k(\mathbf{1}_F {}^k f) = \mathbb{E}_n(\mathbf{1}_F \mathbb{E}_k {}^k f) = \mathbb{E}_n(\mathbf{1}_F \cdot 0) = 0 \quad \text{for } n \leqslant k,$$

so that

$$\|\mathbf{1}_F {}^k f\|_{H^q} = \|\mathbf{1}_F ({}^k f)^\star\|_q \leqslant \mu(F)^{1/q} \|({}^k f)^\star\|_\infty = \mu(F)^{1/q} \|{}^k f\|_\infty.$$

Thus we may conclude the computation (3.54) with

$$\Big(\frac{K}{\alpha \|{}^k f\|_\infty} \|\mathbf{1}_F {}^k f\|_{H^q}\Big)^q \leqslant \Big(\frac{K}{\alpha}\Big)^q \mu(F),$$

and this completes the proof. □

From the technical estimate, we can deduce a distributional inequality, which roughly says that the set where, at the same time, $T^\star f := (Tf)^\star$ is very large and f is small (this being the main obstacle to the possibility of bounding the former by the latter), is only a small fraction of the set where $T^\star f$ is quite large. While we have the object that we want to estimate, namely $T^\star f$, on both sides of the inequality, the small fraction eventually allows us to absorb the right-hand side to the left in an integrated version of the inequality below. Distributional estimates of this form are commonly known as *good λ inequalities.*

Lemma 3.5.14 (Good-λ inequality for martingale transforms). *Suppose that a martingale transform T satisfies*

$$\frac{\mu(F \cap \{\|T({}^k f)\| > \alpha \|{}^k f\|_\infty\})}{\mu(F)} \leqslant \eta \qquad \forall F \in \mathscr{F}_k,\ k \in \mathbb{Z} \cup \{-\infty\}, \tag{3.55}$$

for some $\alpha, \eta > 0$*. Then, for any* $\theta > 1$ *and* $\delta \leqslant (\theta - 1)/(3\alpha + \|v\|_{L^\infty_{\mathrm{so}}})$*, where* v *is the predictable sequence defining* T*, we have the estimate*

$$\mu(T^\star f > \theta\lambda, f^\star \vee w^\star \leqslant \delta\lambda) \leqslant \eta \cdot \mu(T^\star f > \lambda) \qquad \forall \lambda > 0, \tag{3.56}$$

whenever f *is a martingale with a finitely non-zero difference sequence and a strongly predictable majorant:* $\|df_n\| \leqslant w_n \in L^0(\mathscr{F}_{n-1})$*, and* $w^\star := \sup_n w_n$*.*

Proof. Fix $\theta > 1$, $\delta > 0$ as above and $\lambda > 0$, denote $g := Tf$, and define the stopping times

$$\begin{aligned}
\rho &:= \inf\{n \in \mathbb{Z} : \|g_n\| > \lambda\}, \\
\nu &:= \inf\{n \in \mathbb{Z} : \|g_n\| > \theta\lambda\}, \\
\sigma &:= \inf\{n \in \mathbb{Z} : f_n^\star \vee w_{n+1}^\star > \delta\lambda\}.
\end{aligned}$$

Rewritten in terms of these stopping times, the claim (3.56) reads as:

$$\mu(\nu < \infty, \sigma = \infty) \leqslant \eta \cdot \mu(\rho < \infty). \tag{3.57}$$

Note that $\rho \leqslant \nu$ since $\theta > 1$, so in particular $\{\nu < \infty, \sigma = \infty\} \subseteq \{\rho < \infty\}$; the point of (3.57) is to show that the former is a strictly smaller subset of the latter.

We now investigate the martingale ${}^\rho f^{\nu\wedge\sigma}$ and its transform ${}^\rho g^{\nu\wedge\sigma} = T({}^\rho f^{\nu\wedge\sigma})$. Pointwise, we have

$$\|{}^\rho f^{\nu\wedge\sigma}\| \leqslant \mathbf{1}_{\{\sigma<\infty\}}(2f_{\sigma-1}^\star + \|df_\sigma\|) + \mathbf{1}_{\{\sigma=\infty\}} 2f^\star \leqslant 2\delta\lambda + \delta\lambda = 3\delta\lambda,$$

where we used that $\|df_\sigma\| \leqslant w_\sigma \leqslant \delta\lambda$. Here and below, we interpret $df_{-\infty} = 0$. On the set $\{\nu < \infty, \sigma = \infty\}$ (where also $\rho \leqslant \nu < \infty$), we also find that

$$\|{}^\rho g^{\nu\wedge\sigma}\| = \|{}^\rho g^\nu\| = \|g_\nu - g_\rho\| = \|g_\nu - g_{\rho-1} - dg_\rho\| > \theta\lambda - \lambda - \delta\lambda\|v\|_{L^\infty_{\mathrm{so}}},$$

since

$$\|dg_\rho\| = \|v_\rho df_\rho\| \leqslant \|v\|_{L^\infty_{\mathrm{so}}} w_\rho \leqslant \|v\|_{L^\infty_{\mathrm{so}}} \delta\lambda$$

for $\rho < \infty = \sigma$. Altogether, we conclude that on $\{\nu < \infty, \sigma = \infty\} \subseteq \{\rho < \infty\}$,

$$\|T({}^\rho f^{\nu\wedge\sigma})\| > (\theta - 1 - \delta\|v\|_{L^\infty_{\mathrm{so}}})\lambda \geqslant 3\alpha\delta\lambda \geqslant \alpha\|{}^\rho f^{\nu\wedge\sigma}\|_\infty.$$

Thus we can apply the assumed condition (3.55) (with $F = \{\rho = k\} \in \mathscr{F}_k$ and ${}^k f^{\nu\wedge\sigma}$ in place of ${}^k f$) to obtain

$$\begin{aligned}
\mu(\nu < \infty, \sigma = \infty) &= \sum_{k\in\mathbb{Z}\cup\{-\infty\}} \mu(\nu < \infty, \sigma = \infty, \rho = k) \\
&\leqslant \sum_{k\in\mathbb{Z}\cup\{-\infty\}} \mu(\{\rho = k\} \cap \{\|T({}^k f^{\nu\wedge\sigma})\| > \alpha\|{}^k f^{\nu\wedge\sigma}\|_\infty\}) \\
&\leqslant \sum_{k\in\mathbb{Z}\cup\{-\infty\}} \eta \cdot \mu(\rho = k) \leqslant \eta \cdot \mu(\rho < \infty).
\end{aligned}$$

$\square$

From the good-λ inequality we can deduce a preliminary L^p-bound, which still contains the predictable majorant w:

Lemma 3.5.15. *Suppose that a martingale transform T is bounded from H^q to $L^{q,\infty}$ for some $q \in [1,\infty)$. Then it also satisfies*

$$\|T^\star f\|_p \leqslant 37(1+\tfrac{p}{q})\|T\|_{H^q\to L^{q,\infty}}\|f^\star \vee w^\star\|_p, \tag{3.58}$$

whenever f is a martingale with a finitely non-zero difference sequence and a predictable majorant w.

Proof. Writing K for the norm of T from H^q to $L^{q,\infty}$ and fixing $\theta > 1$, we have from Lemma 3.5.13 that (3.55) holds with $\eta := (K/\alpha)^q = 2^{-1}\theta^{-p}$, provided that we choose $\alpha = 2^{1/q}K\theta^{p/q}$.

For f as in the assumptions, and δ as in Lemma 3.5.14 (whose assumption (3.55) we just checked), we then have the good-λ inequality (3.56). Integrating with respect to λ, we deduce that

$$\begin{aligned}
\|T^\star f\|_p^p &= \theta^p \int_0^\infty p\lambda^{p-1}\mu(T^\star f > \theta\lambda)\,\mathrm{d}\lambda \\
&\leqslant \theta^p \int_0^\infty p\lambda^{p-1}\mu(T^\star f > \theta\lambda, f^\star \vee w^\star \leqslant \delta\lambda)\,\mathrm{d}\lambda \\
&\qquad + \theta^p \int_0^\infty p\lambda^{p-1}\mu(f^\star \vee w^\star > \delta\lambda)\,\mathrm{d}\lambda \\
&\leqslant \eta\theta^p \int_0^\infty p\lambda^{p-1}\mu(T^\star f > \lambda)\,\mathrm{d}\lambda + \Big(\frac{\theta}{\delta}\Big)^p \|f^\star \vee w^\star\|_p^p \\
&= \eta\theta^p \|T^\star f\|_p^p + \Big(\frac{\theta}{\delta}\Big)^p \|f^\star \vee w^\star\|_p^p.
\end{aligned}$$

The claim of the Lemma is vacuous unless $\|f^\star\|_{L^p} < \infty$, and in this case the finitely non-zero property of the difference sequence ensures that also $\|T^\star f\|_{L^p} < \infty$. This allows us to absorb the first terms of the above computations on the left hand side and conclude that

$$\begin{aligned}
\frac{\|T^\star f\|_p}{\|f^\star \vee w^\star\|_p} &\leqslant \frac{\theta/\delta}{(1-\eta\theta^p)^{1/p}} = 2^{1/p}\frac{\theta}{\delta} && \text{by the choice } \eta = 2^{-1}\theta^{-p} \\
&= 2^{1/p}\frac{\theta(3\alpha + \|v\|_{L^\infty_{\mathrm{so}}})}{\theta - 1} && \text{taking the maximal admissible } \delta \\
&\leqslant 2^{\frac{1}{p}+\frac{1}{q}}\frac{\theta(3K\theta^{p/q} + K)}{\theta - 1} && \text{by the choice of } \alpha \text{ and Prop. 3.5.10.}
\end{aligned}$$

By minimising the leading term $\theta^{1+p/q}/(\theta-1)$ as a function of θ, we are led to choose $\theta = 1 + q/p$. Then $\theta^{p/q} \leqslant e$ and $\theta/(\theta-1) = 1 + p/q$, so that

$$\frac{\|T^\star f\|_p}{\|f^\star \vee w^\star\|_p} \leqslant 2^{\frac{1}{p}+\frac{1}{q}}(1+\tfrac{p}{q})(3e+1)K \tag{3.59}$$

and this yields the required bound. □

Now we are ready to complete the proof of several claims of Theorem 3.5.4:

Proof of Theorem 3.5.4, the implications (1)⇒(2), (1)⇒(3), *and* (3.46). Let again $K := \|T\|_{H^q \to L^{q,\infty}}$. Given a general f, we apply Lemma 3.5.15 to the good part of its Davis decomposition $f = g + h$ (Theorem 3.4.3; note that g starts at 0 and has finitely non-zero difference sequences if f has these properties), to find that

$$\begin{aligned}\|T^\star g\|_p &\leqslant 37K(1+\tfrac{p}{q})\|g^\star \vee w^\star\|_p \\ &\leqslant 37K(1+\tfrac{p}{q})(10p\|f^\star\|_p + 8\|f^\star\|_p) \leqslant 666K(1+\tfrac{p}{q})p\|f\|_{H^p}.\end{aligned}$$

For the harmless part, we have, from Theorem 3.4.3 and Proposition 3.5.10,

$$\|(Th)^\star\|_p \leqslant \|v\|_{L^\infty_{\mathrm{so}}} \Big\| \|h_{-\infty}\| + \sum_j \|dh_j\| \Big\|_p \leqslant 2^{1/q}K \cdot 9p\|f^\star\|_p \leqslant 18Kp\|f\|_{H^p},$$

so that altogether

$$\begin{aligned}\|T^\star f\|_p \leqslant \|T^\star g\|_p + \|T^\star h\|_p &\leqslant 684K(1+\tfrac{p}{q})p\|f\|_{H^p} \\ &\leqslant 684K(1+\tfrac{p}{q})pp'\|f\|_p.\end{aligned}$$

Note that this contains both (2) (for $p \in (1,\infty)$, when the last estimate is meaningful) and (3) (for $p = 1$, when we should stop at the penultimate estimate involving the H^1 norm). □

Proof of (3.48). Using the trivial bounds

$$\|T\|_{L^p \to L^p} \leqslant \|T^\star\|_{L^p \to L^p}, \quad \|T\|_{H^q \to L^{q,\infty}} \leqslant \|T\|_{L^q \to L^q},$$

the established bound (3.46), written for both T and its formal adjoint $T^* : L^{q'}(S;Y^*) \to L^{q'}(S;X^*)$, and recalling that the norms of adjoint operators are equal, we have

$$\begin{aligned}\|T\|_{L^p(S;X) \to L^p(S;Y)} &\leqslant 700\|T\|_{L^q(S;X) \to L^q(S;Y)}(1+\tfrac{p}{q})pp', \\ \|T\|_{L^p(S;X) \to L^p(S;Y)} &\leqslant 700\|T^*\|_{L^{q'}(S;Y^*) \to L^{q'}(S;X^*)}(1+\tfrac{p'}{q'})pp'.\end{aligned}$$

Combining the alternative bounds, we finally deduce that

$$\begin{aligned}\|T\|_{L^p(S;X) \to L^p(S;Y)} &\leqslant 700\|T\|_{L^q(S;X) \to L^q(S;Y)} \min\Big\{1+\frac{p}{q}, 1+\frac{p'}{q'}\Big\}pp' \\ &\leqslant 1400\|T\|_{L^q(S;X) \to L^q(S;Y)}pp',\end{aligned}$$

observing that $\min(p/q, p'/q') \leqslant 1$. □

Proof of (3.49). If the martingale transform is relative to the Paley–Walsh filtration, instead of invoking the Davis decomposition we may apply (3.58) directly to f, observing that by Proposition 3.1.10 the sequence $\|df\|$ is predictable in this case (so that we can take the weights $w_n = \|df_n\|$). Moreover, to obtain a better constant we apply the slightly better estimate (3.59) with the result that

$$\begin{aligned}
\|T^\star f\|_p &\leqslant 10K2^{\frac{1}{p}+\frac{1}{q}}(1+\tfrac{p}{q})\|f^\star \vee (df)^\star\|_p \\
&\leqslant 20K2^{\frac{1}{p}+\frac{1}{q}}(1+\tfrac{p}{q})\|f^\star\|_p \\
&\leqslant 20K2^{\frac{1}{p}+\frac{1}{q}}(1+\tfrac{p}{q})p'\|f\|_p \\
&= 20KC(p,q)pp'\|f\|_p,
\end{aligned}$$

where $C(p,q) = 2^{\frac{1}{p}+\frac{1}{q}}\big(\frac{1}{p}+\frac{1}{q}\big)$. Repeating the above argument with adjoints in this case, we deduce that

$$\begin{aligned}
\|T\|_{L^p(S;X)\to L^p(S;Y)} &\leqslant 20\|T\|_{L^q(S;X)\to L^q(S;Y)} \min\{C(p,q), C(p',q')\}pp' \\
&\leqslant 80\|T\|_{L^q(S;X)\to L^q(S;Y)}\Big(\frac{p}{q}+\frac{p'}{q'}\Big),
\end{aligned}$$

where the verification of the elementary inequality

$$\min\{C(p,q), C(p',q')\} \leqslant 4\Big(\frac{1}{pq'}+\frac{1}{p'q}\Big)$$

is left to the reader. □

3.5.c End-point estimates in L^1

Here we deal with the remaining implication (1)⇒(4) of Theorem 3.5.4, and establish the quantitative bound (3.47). Interestingly, this will require the use of the other fundamental martingale decomposition, due to Gundy. Note that (1) of Theorem 3.5.4 has essentially two cases: either $T : L^q \to L^{q,\infty}$ for some $q \in (1,\infty)$ (since $H^q = L^q$ in this case) or $T : H^1 \to L^{1,\infty}$. But in the latter case we can apply the part (1)⇒(2) of the theorem already proved, so that we can always assume that we have $T : L^q \to L^{q,\infty}$ for some $q \in (1,\infty)$ to begin with. Indeed, invoking the full strength of (2) of Theorem 3.5.4, we could assume more, but this will not be needed.

We begin by considering the smaller operator T in place of $T^\star$:

Proposition 3.5.16. *For any $q \in (1,\infty)$,*

$$\|T\|_{L^1\to L^{1,\infty}} \leqslant 36\,\|T\|_{L^q\to L^{q,\infty}}.$$

Proof. Let us fix $q \in (1,\infty)$ and denote $K := \|T\|_{L^q\to L^{q,\infty}}$. Given $f \in L^1(S;X)$ and $\lambda > 0$, we split $f = g + b + h$ according to the Gundy decomposition (Theorem 3.4.1 with $f = \mathbb{E}(f|\mathscr{F}_n)$) at level $\alpha\lambda$ in place of λ, where $\alpha > 0$ is yet to be chosen. Then

$$\mu(\|Tf\| > \lambda) \leqslant \mu(\|Tg\| > \tfrac{1}{2}\lambda) + \mu(\|Tb\| > 0) + \mu(\|Th\| > \tfrac{1}{2}\lambda).$$

We estimate the three terms separately. To the g term we apply the assumed $L^q \to L^{q,\infty}$ boundedness to find that

$$\begin{aligned}\mu(\|Tg\| > \lambda/2) &\leqslant (2/\lambda)^q \|Tg\|_{L^{q,\infty}}^q \leqslant (2K/\lambda)^q \|g\|_q^q \\ &\leqslant (2K/\lambda)^q \|g\|_\infty^{q-1} \|g\|_1 \leqslant (2K/\lambda)^q (2\alpha\lambda)^{q-1} 4\|f\|_1 \\ &= 2\frac{(4\alpha K)^q}{\alpha\lambda}\|f\|_1.\end{aligned}$$

For the b term we observe that

$$\|Tb\| \leqslant \|v_{-\infty} b_{-\infty}\| + \sum_k \|v_k db_k\| = 0 \qquad \text{on } \{b^\star = 0\},$$

and therefore

$$\mu(\|Tb\| > 0) \leqslant \mu(b^\star > 0) \leqslant \frac{3}{\alpha\lambda}\|f\|_1.$$

Finally, for the h term we find, via Proposition 3.5.10,

$$\begin{aligned}\mu(\|Th\| > \lambda/2) &\leqslant \frac{2}{\lambda}\|Th\|_1 \leqslant \frac{2}{\lambda}\sum_k \|v_k dh_k\|_1 \\ &= \frac{2}{\lambda}\|v\|_{L^\infty_{\rm so}} \sum_k \|dh_k\|_1 \leqslant \frac{2}{\lambda} \cdot 2K \cdot 4\|f\|_1 = \frac{16}{\lambda}K\|f\|_1.\end{aligned}$$

We now choose $\alpha = (4K)^{-1}$ to conclude that

$$\mu(\|Tf\| > \lambda) \leqslant \frac{8K}{\lambda}\|f\|_1 + \frac{12K}{\lambda}\|f\|_1 + \frac{16K}{\lambda}\|f\|_1 = \frac{36K}{\lambda}\|f\|_1.$$

□

We next aim to dominate $T^\star$ in terms of T. For $p > 1$, we have already established all the bounds for both operators claimed in Theorem 3.5.4, but had we only the boundedness of T on one of these L^p, a corresponding estimate for $T^\star$ would immediately follow from the factorisation $T^\star = M \circ T$ and the boundedness of Doob's maximal operator $M : f \mapsto \sup_n \mathbb{E}_n\|f\|$ on these spaces. However, this reasoning breaks down in estimates related to L^1, since M maps neither L^1 nor $L^{1,\infty}$ to itself. To compensate for this fact, we make use of the variant

$$M_\delta f := (M(\|f\|^\delta))^{1/\delta} = \sup_n (\mathbb{E}_n\|f\|^\delta)^{1/\delta},$$

with $0 < \delta < p$, which puts the good mapping properties of M on L^p and $L^{p,\infty}$, $p \in (1,\infty)$, at our disposal in a larger range of $p \in (\delta, \infty)$:

Lemma 3.5.17. *For $0 < \delta < p < \infty$, we have*

$$\|M_\delta f\|_{L^p} \leqslant \Big(\frac{p}{p-\delta}\Big)^{1/\delta}\|f\|_{L^p}, \quad \|M_\delta f\|_{L^{p,\infty}} \leqslant \Big(\frac{p}{p-\delta}\Big)^{1/\delta}\|f\|_{L^{p,\infty}}.$$

In particular, these bounds are valid for $p = 1$ as soon as $\delta \in (0, 1)$.

Proof. From Doob's maximal inequality (see Theorem 3.2.3) we obtain

$$\|M_\delta f\|_{L^p} = \big\|M(\|f\|^\delta)\big\|_{L^{p/\delta}}^{1/\delta} \leqslant \big((p/\delta)'\big\|\|f\|^\delta\big\|_{L^{p/\delta}}\big)^{1/\delta} = \Big(\frac{p}{p-\delta}\Big)^{1/\delta}\|f\|_{L^p}.$$

Exactly the same steps are valid with L^p replaced by $L^{p,\infty}$ throughout. □

We define yet another maximal function by

$$M_\delta^\# g := \sup_n (\mathbb{E}_n\|_n g\|^\delta)^{1/\delta},$$

where

$${}_n g := g - g_n = g - \mathbb{E}_n g.$$

(This function should not be confused with the martingale ${}^n g$ defined in Definition 3.1.13). We can then estimate the Doob maximal function by the sum of the two new maximal functions:

Lemma 3.5.18. *We have*

$$g^\star \leqslant 2^{1/\delta-1}\big(M_\delta g + M_\delta^\# g\big), \qquad \delta \in (0, 1].$$

Proof. From the defining identity $g_n = g - {}_n g$, we obtain $\|g_n\|^\delta \leqslant \|g\|^\delta + \|_n g\|^\delta$, and then

$$\|g_n\|^\delta = \mathbb{E}_n\|g_n\|^\delta \leqslant \mathbb{E}_n\|g\|^\delta + \mathbb{E}_n\|_n g\|^\delta.$$

The claim follows by raising to the power $1/\delta$, using the convexity of $t \mapsto t^{1/\delta}$, and taking the supremum over all n. □

When applied to $g = Tf$, we will show that this leads to the following pointwise estimate, which is a martingale analogue of a well-known inequality of Cotlar for singular integrals in harmonic analysis:

Proposition 3.5.19 (Cotlar's inequality for martingale transforms). *Let $T : L^1 \to L^{1,\infty}$ be a bounded martingale transform. Then*

$$T^\star f \leqslant 2^{1/\delta-1} M_\delta(Tf) + \Big(\frac{2}{1-\delta}\Big)^{1/\delta}\|T\|_{L^1 \to L^{1,\infty}} Mf.$$

The proof of Proposition 3.5.19 will be streamlined by isolating a couple of simple lemmas:

Lemma 3.5.20. *For $\phi \in L^{1,\infty}(S)$, $F \in \mathscr{A}$ and $\delta \in (0, 1)$,*

$$\int_F \phi^\delta \, \mathrm{d}\mu \leqslant \frac{1}{1-\delta}\|\phi\|_{L^{1,\infty}}^\delta \mu(F)^{1-\delta}.$$

Proof. We assume that $\mu(F) \in (0, \infty)$, for otherwise the claim is trivial. Then

$$\begin{aligned}\int_F \phi^\delta \, \mathrm{d}\mu &= \int_0^\infty \delta t^{\delta-1} \mu(F \cap \{\phi > t\}) \, \mathrm{d}t \\ &\leqslant \int_0^A \delta t^{\delta-1} \mu(F) \, \mathrm{d}t + \int_A^\infty \delta t^{\delta-2} \|\phi\|_{L^{1,\infty}} \, \mathrm{d}t \\ &= A^\delta \mu(F) + \frac{\delta}{1-\delta} A^{\delta-1} \|\phi\|_{L^{1,\infty}},\end{aligned}$$

and the choice $A = \|\phi\|_{L^{1,\infty}} / \mu(F)$ gives the claimed bound. □

Lemma 3.5.21. *Let ϕ and ψ be non-negative $\mathscr{F}$-measurable functions such that*

$$\fint_F \phi^\delta \, \mathrm{d}\mu \leqslant \Big(\fint_F \psi \, \mathrm{d}\mu \Big)^\delta \qquad \forall F \in \mathscr{F} \text{ with } 0 < \mu(F) < \infty.$$

Then $\phi \leqslant \psi$ almost everywhere.

Proof. Assume for contradiction that $\mu(\phi > \psi) > 0$. Then $\mu(\phi \geqslant b > a \geqslant \psi) > 0$ for some $a, b \in \mathbb{Q}_+$. Now $\{\phi \geqslant b > a \geqslant \psi\} \in \mathscr{F}$. Let F be a subset of this set of finite positive measure. With this F, the assumption gives us $b^\delta \leqslant a^\delta$, a contradiction. □

Proof of Proposition 3.5.19. Let $g = Tf$ and $K = \|T\|_{L^1 \to L^{1,\infty}}$. After an application of Lemma 3.5.18, it remains to be shown that

$$M_\delta^\# g := \sup_n (\mathbb{E}_n \|{}_n g\|^\delta)^{1/\delta} \leqslant \frac{2K}{(1-\delta)^{1/\delta}} Mf. \tag{3.60}$$

We first estimate $\mathbb{E}_n \|{}_n g\|^\delta$ for a fixed n. For $F \in \mathscr{F}_n$, we have

$$\mathbf{1}_F \cdot {}_n g = \mathbf{1}_F \sum_{k>n} v_k df_k = \sum_{k>n} v_k d_k(\mathbf{1}_F f) = T({}_n(\mathbf{1}_F f)),$$

and thus by Lemma 3.5.20

$$\int_F \|{}_n g\|^\delta \, \mathrm{d}\mu = \int_F \|T({}_n(\mathbf{1}_F f))\|^\delta \, \mathrm{d}\mu \leqslant \frac{1}{1-\delta} \|T({}_n(\mathbf{1}_F f))\|_{L^{1,\infty}}^\delta \mu(F)^{1-\delta},$$

where

$$\|T({}_n(\mathbf{1}_F f))\|_{L^{1,\infty}} \leqslant K \|{}_n(\mathbf{1}_F f)\|_{L^1} \leqslant 2K \|\mathbf{1}_F f\|_{L^1}.$$

Hence

$$\begin{aligned}\fint_F \mathbb{E}_n \|{}_n g\|^\delta \, \mathrm{d}\mu &= \fint_F \|{}_n g\|^\delta \, \mathrm{d}\mu \\ &\leqslant \frac{(2K)^\delta}{1-\delta} \Big(\fint_F \|f\| \, \mathrm{d}\mu \Big)^\delta = \frac{(2K)^\delta}{1-\delta} \Big(\fint_F \mathbb{E}_n \|f\| \, \mathrm{d}\mu \Big)^\delta.\end{aligned}$$

An application of Lemma 3.5.21 shows that

$$(\mathbb{E}_n\|_n g\|^\delta)^{1/\delta} \leqslant \frac{2K}{(1-\delta)^{1/\delta}}\mathbb{E}_n\|f\|.$$

Taking the supremum over n completes the proof (3.60), and thus of Proposition 3.5.19. □

The proof of the remaining implication (1)⇒(4) of Theorem 3.5.4, as well as the remaining quantitative bound (3.47), is now completed by the following:

Proposition 3.5.22.

$$\|T^\star\|_{L^1\to L^{1,\infty}} \leqslant 48\|T\|_{L^1\to L^{1,\infty}} \leqslant 1728\|T\|_{L^q\to L^{q,\infty}} \qquad \forall q\in(1,\infty).$$

Proof. Using the estimate $\|a+b\|_{L^{1,\infty}} \leqslant 2\|a\|_{L^{1,\infty}} + 2\|b\|_{L^{1,\infty}}$ for $a,b\in L^{1,\infty}$ and Cotlar's estimate (see Proposition 3.5.19) we obtain

$$\begin{aligned}
\|T^\star f\|_{L^{1,\infty}} &\leqslant 2^{1/\delta}\|M_\delta(Tf)\|_{L^{1,\infty}} + 2\Big(\frac{2}{1-\delta}\Big)^{1/\delta}\|T\|_{L^1\to L^{1,\infty}}\|Mf\|_{L^{1,\infty}} \\
&\leqslant \Big(\frac{2}{1-\delta}\Big)^{1/\delta}\|Tf\|_{L^{1,\infty}} + 2\Big(\frac{2}{1-\delta}\Big)^{1/\delta}\|T\|_{L^1\to L^{1,\infty}}\|f\|_{L^1} \\
&\qquad \text{by Lemma 3.5.17 and Doob's inequality} \\
&\leqslant \frac{3\cdot 2^{1/\delta}}{(1-\delta)^{1/\delta}}\|T\|_{L^1\to L^{1,\infty}}\|f\|_{L^1},
\end{aligned}$$

Now the first estimate follows with the choice $\delta=\frac{1}{2}$. The second is then immediate from Proposition 3.5.16. □

3.5.d Martingale type and cotype

The next definition introduces two notions that have rich connections to questions of uniform convexity and smoothness in the geometry of Banach spaces. We offer the present limited treatment of these notions mainly as an illustration of the application of operator-valued martingale transforms to some questions where the presence of such transforms is perhaps not evident at first sight.

Definition 3.5.23. *Let $1\leqslant p\leqslant 2\leqslant q\leqslant\infty$. A Banach space X is said to have:*

(i) martingale type p, *if*

$$\|f_N\|_{L^p(S;X)} \leqslant C\Big(\|f_0\|^p_{L^p(S;X)} + \sum_{n=1}^N \|df_n\|^p_{L^p(S;X)}\Big)^{1/p}$$

for all X-valued L^p-martingales $(f_n)_{n=0}^N$.

(ii) martingale cotype q, *if*

$$\Big(\|f_0\|_{L^q(S;X)}^q + \sum_{n=1}^N \|df_n\|_{L^q(S;X)}^q\Big)^{1/q} \leqslant C\|f_N\|_{L^q(S;X)}$$

(with an obvious modification if $q=\infty$) for all X-valued L^q-martingales $(f_n)_{n=0}^N$.

The best constants will be denoted by $\tau_{p,X}^{\mathrm{mart}}$ and $c_{q,X}^{\mathrm{mart}}$, respectively.

Remark 3.5.24. Here, and in other instances below, whenever we quantify over "all martingales", it is understood that the quantification concerns martingales of any length, over any measure space, with any σ-finite filtration.

By an approximation argument as in Lemma 3.6.2 below, an equivalent definition would be obtained if one only considers simple martingales on probability spaces, and one even may assume the σ-algebras $(\mathscr{F}_k)_{k=0}^n$ to be finite. This fact will not be needed, however.

Every Banach space has martingale type 1 and martingale cotype ∞; the latter follows by writing $df_n = (f_0+\sum_{k=1}^n df_k) - (f_0+\sum_{k=1}^{n-1} df_k)$. Proposition 3.5.6 shows that every Hilbert space has martingale type 2 and cotype 2, with constants $\tau_{2,X}^{\mathrm{mart}} = c_{2,X}^{\mathrm{mart}} = 1$.

Roughly speaking, the plan is to show that the definitions of martingale type p and martingale cotype q can be equivalently formulated in terms of L^r-martingales for any given $r\in(1,\infty)$. This will be deduced from the extrapolation theorem for martingale transforms (Theorem 3.5.4). We begin with the case of martingale cotype, for which the relevant martingale transforms are somewhat more readily identified.

Proposition 3.5.25. *Let X be a Banach space and $q\in[2,\infty)$. The following assertions are equivalent:*

(1) *X has martingale cotype q;*
(2) *for some (equivalently, for all) $r\in(1,\infty)$, we have*

$$\Big\|\Big(\|f_0\|_X^q + \sum_{n=1}^N \|df_n\|_X^q\Big)^{1/q}\Big\|_{L^r(S)} \leqslant C\|f_N\|_{L^r(S;X)} \tag{3.61}$$

for all X-valued L^r-martingales $(f_n)_{n=0}^N$.

For every $r\in(1,\infty)$, the best constant $C = C_{r,q,X}$ in this inequality satisfies $C \eqsim_{q,r} c_{q,X}^{\mathrm{mart}}$, with implied constants depending only on q and r, but not on the space X.

Proof. Observe that the defining condition of martingale cotype q is the special case $q=r$ of (3.61); thus the content of the Proposition is that if this inequality holds for one $r \in (1,\infty)$, then it holds for all $r \in (1,\infty)$. This follows from Theorem 3.5.4 as soon as we identify the left-hand side as the norm of an (operator-valued) martingale transform acting on f, and this is easy:

Let e_n, $n = 0, 1, \dots, N$, be the canonical unit vectors in ℓ^q_{N+1}. Then

$$\Big(\|f_0\|_X^q + \sum_{n=1}^N \|df_n\|_X^q\Big)^{1/q} = \Big\|e_0 \otimes f_0 + \sum_{n=1}^N e_n \otimes df_n\Big\|_{\ell^q_{N+1}(X)}.$$

Thus (3.61) asserts the boundedness from $L^r(S;X)$ to $L^r(S;\ell^q_{N+1}(X))$ of the martingale transform $T = T_v$, with the operator-valued transforming sequence $(v_n)_{n=0}^N$, where $v_n \in \mathscr{L}(X, \ell^q_{N+1}(X))$ is given by $v_n(x) = e_n \otimes x$. Now the Proposition is an immediate corollary to Theorem 3.5.4. □

Before turning to the analogous statement for martingale type, we prove the boundedness of a certain projection.

Lemma 3.5.26. *Let X be a Banach space, and fix $p \in (1,\infty)$ and $r \in (1,\infty)$. On the space $L^r(S;\ell^p(X))$, consider the operator $\mathbb{D}_n := \mathbb{E}(\cdot|\mathscr{F}_n) - \mathbb{E}(\cdot|\mathscr{F}_{n-1})$ for $n \geqslant 1$. Then*

$$PF := P(F_n)_{n\geqslant 0} := (\mathbb{E}_0 F_0, \mathbb{D}_1 F_1, \mathbb{D}_2 F_2, \dots)$$

is a bounded projection on $L^r(S;\ell^p(X))$, whose norm is bounded above by a constant depending only on p and r.

Proof. Writing $\mathbb{E}_n = \mathbb{E}(\cdot|\mathscr{F}_n)$ we have

$$PF = P(F_n)_{n\geqslant 0} = (\mathbb{E}_0 F_0, \mathbb{E}_1 F_1, \dots) - (0, \mathbb{E}_0 F_1, \mathbb{E}_1 F_2, \dots).$$

Thus the $L^p(S;\ell^p(X))$-boundedness of P will follow from the bound

$$\|(\mathbb{E}_{n_j} F_n)_{n\geqslant 0}\|_{L^r(S;\ell^p(X))} \leqslant C_{r,p} \|(F_n)_{n\geqslant 0}\|_{L^r(S;\ell^p(X))}. \tag{3.62}$$

But this bound follows from the vector-valued Doob inequality (Theorem 3.2.7). Alternatively, one can prove the bound by an extrapolation argument as follows. For $r = p$, observe that (3.62) is immediate from Fubini's theorem and the L^p-contractivity of conditional expectations. Now, the left-hand side of (3.62) can be written as an operator-valued martingale transform of the right-hand side, and hence the extrapolation result of Theorem 3.5.4 can be used. □

Now we are in a position to prove the analogue of Proposition 3.5.25 for martingale type:

Proposition 3.5.27. *Let X be a Banach space and $p \in (1,2]$. The following assertions are equivalent:*

(1) *X has martingale type p;*

(2) *for some (equivalently, for all) $r \in (1,\infty)$, we have*

$$\|f_N\|_{L^r(S;X)} \leqslant C\Big\|\Big(\|f_0\|_X^p + \sum_{n=1}^N \|df_n\|_X^p\Big)^{1/p}\Big\|_{L^r(S)} \tag{3.63}$$

for all X-valued L^r-martingales $(f_n)_{n=0}^N$.

For every $r \in (1,\infty)$, the best constant $C = C_{r,p,X}$ in this inequality satisfies $C \eqsim_{r,p} \tau_{p,X}^{\rm mart}$, with implied constants depending only on r and p, but not on the space X.

Proof. The martingale type p estimate, as defined earlier, is the special case for $r = p$ of the estimate (3.63), which can be reformulated as

$$\|f_n\|_{L^r(S;X)} \leqslant C\Big\|e_0 \otimes f_0 + \sum_{n=1}^N e_n \otimes df_n\Big\|_{L^r(S;\ell^p_{N+1}(X))},$$

where e_n, $n = 0,1,\dots,N$, are the canonical unit vectors in ℓ^p_{N+1}. Writing e_n^*, $n = 0,1,\dots,N$, for the corresponding vectors in $(\ell^p_{N+1})^* \eqsim \ell^{p'}_{N+1}$, and abbreviating $\mathbb{E}_0 := \mathbb{E}(\cdot|\mathscr{F}_0)$, $\mathbb{D}_n := \mathbb{E}(\cdot|\mathscr{F}_n) - \mathbb{E}(\cdot|\mathscr{F}_{n-1})$, we have

$$f_n = f_0 + \sum_{n=1}^N df_n = \langle \mathbb{E}_0(F), e_0^*\rangle + \sum_{n=1}^N \langle \mathbb{D}_k(F), e_n^*\rangle,$$

where $F : S \to \ell^p_{N+1}(X)$ is defined by

$$F = e_0 \otimes f_0 + \sum_{n=1}^N e_n \otimes df_n.$$

Therefore, defining $v_j : \ell^p_{N+1}(X) \to X$ by $v_j(\sum_{n=0}^N e_n \otimes x_n) = x_j$, (3.63) translates into the statement that

$$\|T_v F\|_{L^r(S;X)} \leqslant C\|F\|_{L^r(S;\ell^p_{N+1}(X))} \tag{3.64}$$

for all $F \in L^r(S;\ell^p_{N+1}(X))$ of the form $F = e_0 \otimes f_0 + \sum_{n=1}^N e_n \otimes df_n$.

By Theorem 3.5.4, if (3.64) holds for *all* $F \in L^r(S;\ell^p_{N+1}(X))$ for some fixed value of $r \in (1,\infty)$, then it holds for all $r \in (1,\infty)$. To complete the proof, we need to show that (3.64) holds for all $F \in L^r(S;\ell^p_{N+1}(X))$ once it holds for all F of the special form $F = e_0 \otimes f_0 + \sum_{n=1}^N e_n \otimes df_n$. Since $T_v F = T_v \circ P(F)$, where

$$PF = P(F_k)_{n=0}^N = (\mathbb{E}_0 F_0, \mathbb{D}_1 F_1, \dots, \mathbb{D}_N F_N)$$

where PF is of the mentioned special form, the equivalence of (3.64) for general and special F follows from the $L^r(S;\ell^p_{N+1}(X))$-boundedness of P which was proved in Lemma 3.5.26. $\square$

As an immediate consequence of the preceding propositions we obtain the following innocent-looking result:

Corollary 3.5.28. *Let X be a Banach space.*

(1) *If X has martingale cotype q for some $q \in [2, \infty]$, then X has martingale cotype r for all $r \in [q, \infty]$.*
(2) *If X has martingale type p for some $p \in [1, 2]$, then X has martingale type r for all $r \in [1, p]$.*

Proposition 3.5.29. *Let $p \in (1, 2]$. A Banach space X has martingale type p if and only if X^* has martingale cotype p', in which case we have $\tau_{p,X}^{\rm mart} \leqslant c_{p',X^*}^{\rm mart} \leqslant 2\tau_{p,X}^{\rm mart}$.*

Proof. We shall use the notation of Lemma 3.5.26.

'Only if': By duality, we can pick a sequence (not necessarily a martingale) $(g_n)_{n=0}^N$ of norm $1 + \varepsilon$ in $\ell_{N+1}^p(L^p(S; X))$ such that

$$\begin{aligned}\Big(\|f_0\|_{L^{p'}(S;X^*)}^q + \sum_{n=1}^N \|df_n\|_{L^{p'}(S;X^*)}^q\Big)^{1/q} &= \int_S [\langle f_0, g_0\rangle + \sum_{n=1}^N \langle df_n, g_n\rangle] \,\mathrm{d}\mu \\ &= \int_S \Big\langle f_0 + \sum_{n=1}^N df_n, \sum_{n=0}^N \mathbb{D}_n g_n \Big\rangle \,\mathrm{d}\mu \\ &\leqslant \|f_N\|_{L^{p'}(S;X^*)} \Big\| \sum_{n=0}^N \mathbb{D}_n g_n \Big\|_{L^p(S;X)}.\end{aligned}$$

Then we apply the defining inequality of martingale type p and use that $\|\mathbb{D}_n g_n\|_{L^p(S;X)} \leqslant 2\|g_n\|_{L^p(S;X)}$, so that

$$\Big\| \sum_{n=0}^N \mathbb{D}_n g_n \Big\|_{L^p(S;X)} \leqslant 2\tau_{p,X}^{\rm mart} \|(g_n)_{n=0}^N\|_{\ell_{N+1}^p(L^p(S;X))} \leqslant 2\tau_{p,X}^{\rm mart}(1 + \varepsilon).$$

'If': By duality, there is a function $h \in L^{p'}(S; X^*)$ of norm $1+\varepsilon$ such that

$$\begin{aligned}\|f_N\|_{L^p(S;X)} &= \int_S \langle f_n, h\rangle \,\mathrm{d}\mu \\ &= \int_S [\langle f_0, \mathbb{D}_0 h\rangle + \sum_{n=1}^N \langle df_n, \mathbb{D}_n h\rangle] \,\mathrm{d}\mu \\ &\leqslant \Big(\|f_0\|_{L^p(S;X)}^p + \sum_{n=1}^N \|df_n\|_{L^p(S;X)}^p\Big)^{1/p} \Big(\sum_{n=0}^N \|\mathbb{D}_n h\|_{L^{p'}(S;X^*)}^{p'}\Big)^{1/p'}.\end{aligned}$$

We can now apply the martingale cotype p' inequality to get

$$\Big(\sum_{n=0}^{N} \|\mathbb{D}_n h\|_{L^{p'}(S;X^*)}^{p'}\Big)^{1/p'} \leqslant c_{p',X^*}^{\mathrm{mart}} \|h\|_{L^{p'}(S;X^*)} \leqslant c_{p',X^*}^{\mathrm{mart}}(1+\varepsilon).$$

□

By a slight modification of this argument, one could establish a duality directly between the equivalent formulations (3.61) and (3.63), with $q = p'$ and r' in place of r in one of the estimates. In this way, only one of Proposition 3.5.25 and 3.5.27 would have to be proved directly, as the other one could be achieved by duality.

We do not give an extensive treatment of martingale type and cotype as this is only slightly related to our main objectives. We do summarise some easy facts. As in the case of the Fourier transform (see Proposition 2.4.17) one can show that martingale cotype interpolates. By a duality argument one obtains the same result for martingale type. In Proposition 4.3.5 we will show that non-trivial martingale type or cotype implies reflexivity.

Proposition 3.5.30. *Let X be a Banach space, let $p \in [1,2]$ and $q \in [2,\infty)$, and let $r \in (1,\infty)$. Let $(T,\mathscr{B},\nu)$ be a measure space.*

(1) *If X has martingale type p, then $L^r(T;X)$ has martingale type $p \wedge r$.*
(2) *If X has martingale cotype q, then $L^r(T;X)$ has martingale cotype $q \vee r$*

As $\mathbb{K}$ has martingale type 2 and martingale cotype 2, it follows that L^p-spaces with $p \in (1,\infty)$ have martingale type $p \wedge 2$ and martingale cotype $p \vee 2$.

Proof. We will prove the result for martingale type, the case of martingale cotype being similar. If $r < p$, note that X has martingale type r as well (with a possibly larger constant), and we may replace p by r and thereby assume that $1 < p \leqslant r < \infty$. As explained below Definition 3.5.23 by Lemma 3.6.2 we may assume each of the $\mathscr{F}_n$ for $n = 0,\dots,N$ consists of finitely many sets. Now let $f : S \to L^r(T;X)$ be a martingale with respect to $(\mathscr{F}_n)_{n=0}^N$. Then as $\mathscr{F}_N$ is finite, one can check that $s \mapsto f_N(s)(t)$ is a martingale for almost all $t \in T$. Therefore, applying Proposition 3.5.27 it follows that

$$\begin{aligned}
\|f\|_{L^r(S;L^r(T;X))} &= \Big(\int_T \|f\|_{L^r(S;X)}^r \,\mathrm{d}\nu\Big)^{1/r} \\
&\leqslant C\Big(\int_T \int_S \Big(\|f_0\|^p + \sum_{n=1}^{N} \|df_n\|^p\Big)^{r/p} \,\mathrm{d}\mu\,\mathrm{d}\nu\Big)^{1/r} \\
&\leqslant C\Big(\int_S \Big(\|f_0\|_{L^r(T;X)}^p + \sum_{n=1}^{N} \|df_n\|_{L^r(T;X)}^p\Big)^{r/p} \,\mathrm{d}\mu\,\mathrm{d}\nu\Big)^{1/r}.
\end{aligned}$$

where we applied the triangle inequality in $L^{r/p}(T)$ in the last step. □

3.6 Approximate models for martingales

While the definition of martingales is rather general, for the purposes of proving some L^p-norm estimates they can often be replaced, without loss of generality, by much more well-structured objects. The goal of this section is to explore some reductions of this type on a reasonably general level so that they are readily applicable to a variety of situations that we shall encounter in the subsequent chapters. Some first applications are already considered later in the section at hand.

3.6.a Universality of Paley–Walsh martingales

Recall from Definition 3.1.8 that a filtration on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is called a *Paley–Walsh filtration* if each set belonging to $\mathscr{F}_n$ is a union of atoms of size 2^{-n}. A *Paley–Walsh martingale* is a martingale with respect to a Paley–Walsh filtration.

Theorem 3.6.1. *Let $(f_k)_{k=0}^n$ be an X-valued L^p-martingale on a divisible measure space $(S, \mathscr{A}, \mu)$, relative to a σ-finite filtration. For any $\varepsilon > 0$, there exists another X-valued L^p-martingale $(g_j)_{j=0}^N$ and an increasing function*

$$\phi : \{0, 1, \dots, n\} \to \{0, 1, \dots, N\}$$

with $\phi(n) = N$ such that:

(i) $\|f_k - g_{\phi(k)}\|_p < \varepsilon$ *for all* $k = 0, 1, \dots, n$,
(ii) $(g_j)_{j=0}^N$ *is supported on a set* $E \in \mathscr{A}$ *of finite measure,*
(iii) $(g_j)_{j=0}^N$ *is a Paley–Walsh martingale in* $L^p(E, \mu_E; X)$, *where*

$$\mu_E := \mu(E)^{-1} \mu|_E$$

is the restriction of μ on E, normalised to a probability measure.

If $f_0 = 0$, we may take $g_0 = 0$.

The proof of Theorem 3.6.1 consists of several subsequent reductions, and we start with a reduction to simple martingales on a finite measure space:

Lemma 3.6.2. *Let $(f_k)_{k=0}^n$ be an $L^p(S; X)$-martingale defined on a measure space $(S, \mathscr{A}, \mu)$, relative to a σ-finite filtration $(\mathscr{F}_k)_{k=0}^n$. For any $\varepsilon > 0$, there is another sequence $(g_k)_{k=0}^n$ of X-valued functions such that:*

(i) *each g_k is a simple function supported on a set $E \in \mathscr{F}_0$ of finite μ-measure,*
(ii) *the restrictions of the g_k to E form a martingale with respect to a filtration in E,*
(iii) $\|f_k - g_k\|_p < \varepsilon$ *for all* $k = 0, 1, \dots, n$.

If $f_0 = 0$, we may take $g_0 = 0$.

Proof. By σ-finiteness and dominated convergence, we can find an $E \in \mathscr{F}_0$ of finite measure such that $\|\mathbf{1}_{\complement E} f_n\|_p < \delta$. Since $\mathbb{E}(\mathbf{1}_E f_j|\mathscr{F}_k) = \mathbf{1}_E \mathbb{E}(f_j|\mathscr{F}_k)$ for all $k, j = 0, 1, \dots, n$, it is immediate that the martingale property of f is inherited by $(\mathbf{1}_E f_k)_{k=0}^n$. Replacing f_k by these approximations in the first step, we henceforth assume that these functions are supported on the finite set E, and make all the remaining constructions inside this set only.

Fix an arbitrary $\delta > 0$. By density, we can find simple functions $s_k \in L^p(E, \mathscr{F}_k; X)$ such that:

$$\|s_0 - f_0\|_p < \delta, \qquad \|s_k - df_k\|_p < \delta \quad \forall k = 1, \dots, n,$$

where we choose $s_0 = 0$ if $f_0 = 0$. We then define

$$\mathscr{G}_0 := \sigma(E, s_0) \subseteq \mathscr{F}_0, \qquad \mathscr{G}_k := \sigma(\mathscr{G}_{k-1}, s_k) \subseteq \mathscr{F}_k \quad \forall k = 1, \dots, n,$$

so that $(\mathscr{G}_k)_{k=0}^n$ is a filtration of finite σ-algebras of E to which $(s_k)_{k=0}^n$ is adapted. Next, we set

$$g_0 := s_0 \quad (= 0 \text{ if } f_0 = 0), \qquad dg_k := s_k - \mathbb{E}(s_k|\mathscr{G}_{k-1}) \quad \forall k = 1, \dots, n,$$

so that this is a martingale difference sequence relative to $(\mathscr{G}_k)_{k=0}^n$ on E.

We have $\|g_0 - f_0\|_p < \delta$ and, using

$$\mathbb{E}(df_k|\mathscr{G}_{k-1}) = \mathbb{E}(\mathbb{E}(df_k|\mathscr{F}_{k-1})|\mathscr{G}_{k-1}) = \mathbb{E}(0|\mathscr{G}_{k-1}) = 0,$$

also the estimate

$$\|dg_k - df_k\|_p = \|s_k - df_k - \mathbb{E}(s_k - df_k|\mathscr{G}_{k-1})\|_p \leqslant 2\|s_k - df_k\|_p < 2\delta.$$

Summing up the difference sequence, we get $\|g_k - f_k\|_p < (1+2k)\delta \leqslant (1+2n)\delta$, which proves the lemma, since $\delta > 0$ can be picked arbitrarily. $\square$

With Lemma 3.6.2 at hand, we have reduced the proof of Theorem 3.6.1 to the case of martingales supported on a set E of finite μ-measure. For the rest of the argument, we will be working on the set E only, and we switch to using the rescaled probability measure $\mu_E = \mu(E)^{-1}\mu$. Since E is fixed, any approximations that we can make arbitrarily good in the $L^p(E, \mu_E; X)$ norm, will also be arbitrarily good with respect to the original $L^p(S, \mu; X)$ norm. We will also write Ω instead of E, and denote the rescaled probability measure again by μ, for simplicity of notation, as there will be no further reference to the original ambient space $(S, \mathscr{A}, \mu)$.

In the following lemma, we make use of the divisibility assumption via two measure-theoretic results proved in the Appendix (Corollary A.1.9 and Proposition A.1.10).

Lemma 3.6.3. *If $(f_k)_{k=0}^n$ is a simple martingale on a divisible probability space $(\Omega, \mathscr{A}, \mu)$, then it is also a martingale with respect to a filtration $(\mathscr{F}_k)_{k=0}^n$ of divisible σ-algebras.*

Proof. Let $\mathscr{E}_k := \sigma(f_0, \ldots, f_k)$, $k \geqslant 0$, be the minimal filtration for which $(f_k)_{k=0}^n$ is a martingale; each of these σ-algebras is finite since each f_k is simple. The divisibility of $(\Omega, \mathscr{A}, \mu)$ guarantees the existence of a random variable u on Ω, uniformly distributed in $[0, 1)$ and independent of $\mathscr{E}_n$ (Proposition A.1.10). The σ-algebra $\mathscr{U} := \sigma(u) = \{u^{-1}(B) : B \in \mathscr{B}([0, 1))\}$ is clearly divisible by the divisibility of the Borel σ-algebra $\mathscr{B}([0, 1))$ and the fact that $\mu(u^{-1}(B)) = \mu(u \in B) = |B|$ by the uniform distribution of u. Thus also $\mathscr{F}_k := \sigma(\mathscr{E}_k, \mathscr{U})$, having a divisible sub-σ-algebra $\mathscr{U}$, is itself divisible (see Corollary A.1.9).

From Proposition 2.6.36, it follows that

$$\mathbb{E}(f_n|\mathscr{F}_k) = \mathbb{E}(f_n|\sigma(\mathscr{E}_k, \mathscr{U})) = \mathbb{E}(f_n|\mathscr{E}_k) = f_k,$$

so that f_k is also a martingale with respect to the filtration $(\mathscr{F}_k)_{k=0}^n$ of divisible σ-algebras, as required. $\square$

We say that a σ-algebra $\mathscr{F}$ of a probability space Ω is *dyadic* if it is generated by a partition $\mathscr{F}^*$ of Ω into 2^m atoms of equal measure 2^{-m}. A *dyadic filtration* is a filtration each of whose constituting σ-algebras is dyadic. We do not insist in this definition that the number of atoms of the nth σ-algebra be 2^n; if we do, we get a Paley–Walsh filtration.

Lemma 3.6.4. *Let $(\Omega, \mathscr{F}, \mu)$ be a divisible probability space and $\mathscr{G} \subseteq \mathscr{F}$ a dyadic sub-σ-algebra. Let $p \in [1, \infty)$ and let $f \in L^p(\Omega, \mathscr{F}, \mu; X)$ be a simple function. For any $\varepsilon > 0$ there is another dyadic sub-σ-algebra $\mathscr{H} \subseteq \mathscr{F}$ such that $\mathscr{H} \supseteq \mathscr{G}$ and a simple $s \in L^p(\Omega, \mathscr{H}, \mu; X)$ such that $\|s - f\|_p < \varepsilon$.*

Proof. It is enough to consider $f = \mathbf{1}_F$ with $F \in \mathscr{F}$ and find, for any given $\delta > 0$, $s = \mathbf{1}_H$ with $H \subseteq F$ and $H \in \mathscr{H}$ so that $\mu(F \setminus H) < \delta$. Let m be an integer such that $2^{-m} < \delta$. For each atom $A \in \mathscr{G}^*$, we use divisibility to find a subset $A_F \subseteq A \cap F$ having measure of the form $\mu(A_F) = k_A 2^{-m} \mu(A)$ for an integer k_A such that $\mu(A_F) > \mu(A \cap F) - 2^{-m}\mu(A)$. Setting $H := \bigcup_{A \in \mathscr{G}^*} A_F$, we have $H \subseteq F$ and

$$\mu(F \setminus H) = \sum_{A \in \mathscr{G}^*} \mu(A \cap F \setminus A_F) \leqslant \sum_{A \in \mathscr{G}^*} 2^{-m}\mu(A) = 2^{-m} < \delta.$$

Now we partition each A_F and $A \setminus A_F$ into subsets of measure $2^{-m}\mu(A) = 2^{-m-k}$ for some k independent of A, since $\mathscr{G}$ is dyadic. This produces a partition of Ω, which generates a dyadic σ-algebra $\mathscr{H} \supseteq \mathscr{G}$ for which $H \in \mathscr{H}$. This completes the proof. $\square$

Lemma 3.6.5. *If $(f_k)_{k=0}^n$ is a simple martingale with respect to a divisible filtration $(\mathscr{F}_k)_{k=0}^n$ of a probability space $(\Omega, \mathscr{A}, \mu)$, then for any $p \in [1, \infty)$ and $\varepsilon > 0$ there exists another simple martingale $(h_k)_{k=0}^n$ with respect to a dyadic filtration $(\mathscr{H}_k)_{k=0}^n$ of $(\Omega, \mathscr{A}, \mu)$ such that $\|f_k - h_k\|_p < \varepsilon$. If $f_0 = 0$, we may take $h_0 = 0$.*

Proof. The proof depends on an inductive application of Lemma 3.6.4. In the basic step, we apply Lemma 3.6.4 to $\mathscr{G} = \{\varnothing, \Omega\}$ (the trivial σ-algebra), $\mathscr{F} = \mathscr{F}_0$ and $f = f_0$ to produce a dyadic σ-algebra $\mathscr{H}_0 \subseteq \mathscr{F}_0$ and a simple $s_0 \in L^p(\mathscr{H}_0, \mu; X)$ such that $\|s_0 - f_0\|_p < \delta$. Of course may choose $s_0 = 0$ if $f_0 = 0$.

In the inductive step, we apply Lemma 3.6.4 to $\mathscr{G} = \mathscr{H}_{k-1}$ (the dyadic σ-algebra constructed in the previous step), $\mathscr{F} = \mathscr{F}_k$ and $f = df_k$ to produce a dyadic σ-algebra $\mathscr{H}_k \subseteq \mathscr{F}_k$ with $\mathscr{H}_k \supseteq \mathscr{H}_{k-1}$, and a simple $s_k \in L^p(\Omega, \mathscr{H}_k, \mu; X)$ such that $\|s_k - df_k\|_p < \delta$.

We next define

$$h_0 := s_0, \qquad dh_k := s_k - \mathbb{E}(s_k|\mathscr{H}_{k-1}) \quad \forall k = 1, \dots, n.$$

This is a martingale difference sequence relative to $(\mathscr{H}_k)_{k=0}^n$ by construction. Moreover, we have $\|h_0 - f_0\|_p < \delta$ and, using

$$\mathbb{E}(df_k|\mathscr{H}_{k-1}) = \mathbb{E}(\mathbb{E}(df_k|\mathscr{F}_{k-1})|\mathscr{H}_{k-1}) = \mathbb{E}(0|\mathscr{H}_{k-1}) = 0,$$

also the estimate

$$\|dh_k - df_k\|_p = \|s_k - df_k - \mathbb{E}(s_k - df_k|\mathscr{H}_{k-1})\|_p \leqslant 2\|s_k - df_k\|_p < 2\delta.$$

Summing the difference sequence, we get $\|h_k - f_k\|_p < (1+2k)\delta \leqslant (1+2n)\delta$, which can be taken as small as we like. $\square$

Lemma 3.6.6. *If $(h_k)_{k=0}^n$ is a martingale with respect to a dyadic filtration $(\mathscr{H}_k)_{k=0}^n$ of a probability space $(\Omega, \mathscr{A}, \mu)$, then there is a Paley–Walsh martingale $(g_j)_{j=0}^N$ such that $h_k = g_{\phi(k)}$ for some increasing function $\phi : \{0, 1, \dots, n\} \to \{0, 1, \dots, N = \phi(n)\}$. If h_0 is a constant, we have $\phi(0) = 0$.*

Proof. By definition of a dyadic filtration, each $\mathscr{H}_k$ is generated by a partition $\mathscr{H}_k^*$ of $2^{\phi(k)}$ atoms of equal measure $2^{-\phi(k)}$, for some increasing function $\phi : \{0, 1, \dots, n\} \to \{0, 1, \dots, N := \phi(n)\}$. We define $\mathscr{G}_{\phi(k)} := \mathscr{H}_k$ (as well as $\mathscr{G}_0 := \{\varnothing, \Omega\}$, if this is not already implied by the previous definition). The remaining instances of $\mathscr{G}_j$ are easily defined by backwards induction, combining pairs of two atoms of $\mathscr{G}_j$ (chosen to be contained in the same atom of $\mathscr{G}_{\phi(k-1)}$ if $\phi(k-1) < j \leqslant \phi(k)$) into atoms of $\mathscr{G}_{j-1}$ in each step. It is easy to see that in this way we produce a Paley–Walsh filtration $(\mathscr{G}_j)_{j=0}^N$ with $\mathscr{G}_{\phi(k)} = \mathscr{H}_k$, so defining $g_j := \mathbb{E}(f_n|\mathscr{G}_j)$ we obtain a Paley–Walsh martingale with $g_{\phi(k)} = f_k$. $\square$

Proof of Theorem 3.6.1. This amounts to a concatenation of the previous lemmas. First, we approximate the original martingale by a simple martingale on a finite measure space (Lemma 3.6.2). Such a simple martingale may also be viewed as a martingale with respect to a divisible filtration (Lemma 3.6.3). This allows us to approximate it by a dyadic martingale (Lemma 3.6.5). And the dyadic martingale is a sampling of a Paley–Walsh martingale, in the sense that $h_k = g_{\phi(k)}$, by the last Lemma 3.6.6. $\square$

Application to Doob's inequality

We illustrate the typical way of applying Theorem 3.6.1, in terms of a simple corollary. A related application will appear in Subsection 3.6.b below, and another one in the next chapter.

Corollary 3.6.7. *Consider Doob's maximal inequality* $\|f^\star\|_p \leqslant c_p \|f\|_p$ *for*

(a) *all* L^p*-martingales on any* σ*-finite measure space,*
(b) *all Paley–Walsh martingales on any probability space,*
(c) *all Paley–Walsh martingales on one fixed divisible probability space.*

Then $c_p = p'$ *is the optimal constant in this inequality in all cases.*

Recall that we showed by example the optimality of $c_p = p'$ in case (a) (Subsection 3.2.a). However, this example was based on martingales with respect to a particular, non-Paley–Walsh filtration. Without the knowledge of Theorem 3.6.1, one could incorrectly suspect that Paley–Walsh martingales might possibly behave better under the maximal operator than general martingales. The proof below shows the sharpness of the constant in case (b) without exhibiting an explicit example of a near-critical martingale of type (b).

Proof. The equality of the best constant in (b) and (c) is immediate, since the L^p-norm of any function of a real Rademacher sequence $(r_j)_{j=1}^\infty$ is invariant under the replacement of $(r_j)_{j=1}^\infty$ by another real Rademacher sequence, possibly living on a different probability space. So the main concern is identifying c_p(b) with c_p(a), which we already know to be p'.

Since $f^\star = \sup_{k\in\mathbb{Z}} \|f_k\|_X$ is the increasing limit of $\sup_{|k|\leqslant n} \|f_k\|_X$, by monotone convergence and re-indexing it is clear that both (a) and (b) are equivalent to their finite versions

$$\Big\| \sup_{0\leqslant k\leqslant n} \|f_k\|_X \Big\|_p \leqslant c_p \|f_n\|_p \tag{3.65}$$

for the respective classes of martingales. It suffices to prove that (3.65) for Paley–Walsh martingales implies the same bound for all L^p-martingales $(f_k)_{k=0}^n$.

To this end, consider a martingale $(f_k)_{k=0}^n$ in some $L^p(S;X)$. Replacing S by the divisible space $S\times[0,1]$ and defining f_k on S by the equidistributed $(s,t)\mapsto f_k(s)$ on $S\times[0,1]$ if necessary, we assume without loss of generality that S is divisible. We then apply Theorem 3.6.1 to produce a $(g_j)_{j=0}^N$ living on a set $E\subseteq S$ of finite measure with the properties listed in the theorem. Then

$$\begin{aligned}
&\Big\| \sup_{0\leqslant k\leqslant n} \|f_k\|_X \Big\|_{L^p(S,\mu)} \\
&\qquad \leqslant \Big\| \sup_{0\leqslant k\leqslant n} \|g_{\phi(k)}\|_X \Big\|_{L^p(S,\mu)} + \sum_{0\leqslant k\leqslant n} \Big\| f_k - g_{\phi(k)} \Big\|_{L^p(S,\mu;X)}
\end{aligned}$$

$$\leqslant \mu(E)^{1/p} \Big\| \sup_{0 \leqslant j \leqslant N} \|g_j\|_X \Big\|_{L^p(E,\mu_E)} + (n+1)\varepsilon,$$

where, using that (3.65) holds for the Paley–Walsh martingale $(g_j)_{j=0}^N$,

$$\begin{aligned} \Big\| \sup_{0 \leqslant j \leqslant N} \|g_j\|_X \Big\|_{L^p(S,\mu_E)} &\leqslant c_p \|g_N\|_{L^p(S,\mu_E;X)} = c_p \mu(E)^{-1/p} \|g_{\phi(n)}\|_{L^p(S,\mu;X)} \\ &\leqslant c_p \mu(E)^{-1/p} \big(\|f_n\|_{L^p(S,\mu;X)} + \varepsilon \big), \end{aligned}$$

so that

$$\begin{aligned} \Big\| \sup_{0 \leqslant k \leqslant n} \|f_k\|_X \Big\|_{L^p(S,\mu)} &\leqslant \mu(E)^{1/p} \frac{c_p}{\mu(E)^{1/p}} \big(\|f_n\|_{L^p(S,\mu;X)} + \varepsilon \big) + (n+1)\varepsilon \\ &= c_p \|f_n\|_{L^p(S,\mu;X)} + (c_p + n + 1)\varepsilon. \end{aligned}$$

Letting $\varepsilon \to 0$, this completes the proof. □

3.6.b The Rademacher maximal function

For vector-valued functions, it is sometimes useful to have a variant of Doob's maximal function which takes into account not only the absolute size, but also to some extend the "directions" of the different f_k.

Let $(r_k)_{k \in \mathbb{Z}}$ be a real Rademacher sequence (see Definition 3.2.9).

Definition 3.6.8. *The* Rademacher maximal function *of a martingale* $f = (f_k)_{k \in \mathbb{Z}}$ *in* $L^0(S;X)$ *is defined by*

$$M_{\mathrm{Rad}} f := \sup \Big\{ \Big\| \sum_{k \in \mathbb{Z}} r_k \lambda_k f_k \Big\|_{L^2(\Omega;X)} : \lambda \in \overline{B}_{\ell^2(\mathbb{Z})} \text{ finitely non-zero} \Big\}.$$

Here $\overline{B}_{\ell^2(\mathbb{Z})}$ denotes the closed unit ball of $\ell^2(\mathbb{Z})$.

Remark 3.6.9. The special choice $\lambda_k = \delta_{kj}$ for $j \in \mathbb{Z}$ shows that

$$M_{\mathrm{Rad}} f \geqslant \sup_{j \in \mathbb{Z}} \|f_j\|_X = f^\star,$$

with $f^\star$ the Doob maximal function. When $X = H$ is a Hilbert space, then by orthogonality

$$\Big\| \sum_{k \in \mathbb{Z}} r_k \lambda_k f_k \Big\|^2_{L^2(\Omega;H)} = \sum_{k \in \mathbb{Z}} \|\lambda_k f_k\|_H^2 \leqslant \sum_{k \in \mathbb{Z}} |\lambda_k|^2 \sup_{j \in \mathbb{Z}} \|f_j\|_H^2,$$

and therefore

$$M_{\mathrm{Rad}} f = \sup_{j \in \mathbb{Z}} \|f_j\|_H.$$

Thus the Rademacher maximal function M_{Rad} is a Banach space-theoretic generalisation of the Doob maximal function.

As it turns out, the analogue of Doob's inequality is not automatically true for the Rademacher maximal function in general Banach spaces. This leads to the following condition that a space may or may not satisfy:

Definition 3.6.10. *A Banach space X has the* RMF property *if*

$$\|M_{\mathrm{Rad}} f\|_{L^p(S)} \leqslant C_{p,X} \|f\|_{L^p(S;X)} \tag{3.66}$$

for all $p \in (1,\infty)$ and all L^p-martingales $(f_k)_{k\in\mathbb{Z}}$ with values in X.

Here, as always, the quantification extends over all measure spaces $(S, \mathscr{A}, \mu)$ with a σ-finite filtration relative to which the martingale can be defined. To check that a Banach space has the RMF property, it of course suffices to check the defining inequality for finitely non-zero L^p-martingales.

Proving some basic properties of the Rademacher maximal function will provide us with an opportunity to illustrate the application of the main theorems of both this and the previous section, namely Theorems 3.5.4 and 3.6.1. As an application of Theorem 3.5.4, we first deal with the role of p in Definition 3.6.10.

Proposition 3.6.11. *A Banach space X has the RMF property if and only if the defining inequality (3.66) holds for some $p \in (1,\infty)$. For any two values of $p \in (1,\infty)$, the ratio of the best constants $C_{p,X}$ in (3.66) is bounded from above and below by constants that only depend on p but not on X.*

Proof. This again depends on identifying the left-hand side of (3.66) as the norm of a martingale transform of f. It is convenient to first observe that by relabelling the indices, (3.66) is equivalent to

$$\|M_{\mathrm{Rad}} (f_k)_{k=0}^n\|_{L^p(S)} \leqslant C \|f_n\|_{L^p(S;X)}$$

for all finite martingales $(f_k)_{k=0}^n$ with values in X. Let e_k^*, $k = 0, 1, \ldots, n$, be the canonical unit vectors in $(\ell^2_{n+1})^*$. Then we see that

$$\begin{aligned} M_{\mathrm{Rad}} (f_k)_{k=0}^n &= \Big\| \sum_{k=0}^n r_k e_k^* \otimes f_k \Big\|_{\mathscr{L}(\ell^2_{n+1}, L^2(\Omega;X))} \\ &= \Big\| \Big(\sum_{k=0}^n r_k e_k^* \Big) \otimes f_0 + \sum_{j=1}^n \Big(\sum_{k=j}^n r_k e_k^* \Big) \otimes df_j \Big\|_{\mathscr{L}(\ell^2_{n+1}, L^2(\Omega;X))}. \end{aligned}$$

Thus (3.66) can be reinterpreted as the boundedness from $L^p(S;X)$ to $L^p(S; \mathscr{L}(\ell^2_{n+1}, L^2(\Omega;X)))$ of the martingale transform $T = T_v$, with the operator-valued transforming sequence

$$v_j \in \mathscr{L}(X, \mathscr{L}(\ell^2_{n+1}, L^2(\Omega;X))), \quad v_j(x) = \sum_{k=j}^n r_k e_k^* \otimes x.$$

Note that each v_j is constant as a function on S with values in the operator space $\mathscr{L}(X, \mathscr{L}(\ell^2_{n+1}, L^2(\Omega; X)))$; the Rademacher functions r_k appearing in the definition of v_j are functions on Ω, not on S. Hence the sequence $(v_j)_{j=0}^n$ is trivially predictable.

Now the p-independence of (3.66) follows directly from Theorem 3.5.4. $\square$

A variant of the argument of Corollary 3.6.7 also gives an analogous result for the Rademacher maximal function:

Corollary 3.6.12. *Let X be a Banach space and $p \in (1, \infty)$. Then the following conditions are equivalent:*

(1) *X has the RMF property;*
(2) *the inequality*

$$\Big\| M_{\rm Rad}(f_k)_{k=0}^n \Big\|_{L^p(S)} \leqslant C \|f_n\|_{L^p(S;X)} \tag{3.67}$$

holds for all finite Paley–Walsh martingales $(f_k)_{k=0}^n$ on any probability space $(S, \mathscr{A}, \mu)$;
(3) *the inequality*

$$\Big\| \sum_{k=0}^n \sum_{I \in \mathscr{D}_k[0,1)} r_I e_I^* \otimes \mathbf{1}_I \langle f \rangle_I \Big\|_{L^p(0,1;\mathscr{L}(\ell^2(\mathscr{D}[0,1)), L^2(\Omega;X))} \leqslant C \|f\|_{L^p(0,1;X)} \tag{3.68}$$

holds for all $f \in L^p(0,1;X)$ and $n = 0, 1, 2, \ldots$, where

$$\mathscr{D}_k[0,1) := \{2^{-k}[j-1, j) : \ j = 1, \ldots, 2^k\},$$

and $\langle f \rangle_I$ denotes the average of f over I.

Moreover, the best constants in (3.67) and (3.68) are equal to the best constant in (3.66).

Proof. Recall that the RMF property is (after a monotone convergence argument detailed in the proof of Proposition 3.6.11) simply the statement that (3.67) holds for all X-valued L^p-martingales on any σ-finite measure space S (either for one or for all $p \in (1, \infty)$, by Proposition 3.6.11).

(1)$\Leftrightarrow$(2): Given an X-valued L^p-martingale $f = (f_k)_{k=0}^n$, where S is divisible without loss of generality, let $g = (g_j)_{j=0}^N$ be an approximating Paley–Walsh martingale with support in a set $E \subseteq S$ of finite measure as provided by Theorem 3.6.1. Then we have the pointwise bound

$$\begin{aligned} M_{\rm Rad}(f_k)_{k=0}^n &= \Big\| \sum_{k=0}^n r_k e_k^* \otimes f_k \Big\|_{\mathscr{L}(\ell^2_{n+1}; L^2(\Omega;X))} \\ &\leqslant M_{\rm Rad}(g_{\phi(k)})_{k=0}^n + \sum_{k=0}^n \|f_k - g_{\phi(k)}\|_X, \end{aligned}$$

noting that

$$\|r_k e_k^* \otimes x\|_{\mathscr{L}(\ell^2_{n+1};L^2(\Omega;X))} = \|x\|_X.$$

After this point, the proof essentially repeats that of Corollary 3.6.7. Taking the L^p-norms, it follows that

$$\|M_{\mathrm{Rad}}(f_k)_{k=0}^n\|_{L^p(S,\mu)} \leqslant \mu(E)^{1/p}\|M_{\mathrm{Rad}}(g_j)_{j=0}^N\|_{L^p(E,\mu_E)} + (n+1)\varepsilon$$

and

$$\|M_{\mathrm{Rad}}(g_j)_{j=0}^N\|_{L^p(E,\mu_E)} \leqslant C\|g_N\|_{L^p(E,\mu_E;X)} \leqslant C\mu(E)^{-1/p}\big(\|f_n\|_{L^p(S,\mu;X)}+\varepsilon\big),$$

so that

$$\|M_{\mathrm{Rad}}(f_k)_{k=0}^n\|_{L^p(S,\mu)} \leqslant C\|f_n\|_{L^p(S,\mu;X)} + (C+n+1)\varepsilon.$$

Letting $\varepsilon \to 0$ completes the proof.

(2)⇔(3): Since the functions

$$r_k(t) := \mathrm{sgn}(\sin(2^{k+1}\pi t)), \quad t \in [0,1),\ k \geqslant 1,$$

realise a real Rademacher sequence, (3.67) can be equivalently stated on the unit interval. The direct translation of (3.67) in this setting yields the inequality

$$\Big\|\sum_{k=0}^n r_k e_k^* \otimes \sum_{I\in\mathscr{D}_k[0,1)} \mathbf{1}_I\langle f\rangle_I\Big\|_{L^p(0,1;\mathscr{L}(\ell^2_{n+1},L^2(\Omega;X))} \leqslant C\|f\|_{L^p(0,1;X)},$$

which seemingly differs from (3.68), in that the factor $r_k e_k^*$, which multiplies the average $\mathbf{1}_I\langle f\rangle_I$ on a dyadic interval I, is indexed by the level $k = -\log_2 \ell(I)$ of the interval I, instead of the interval I itself. However, looking at the integrand of the L^p-norm on the left of (3.68) at any fixed point $t \in [0,1)$, we find a sum

$$\sum_{k=0}^n \sum_{I\in\mathscr{D}_k[0,1)} r_I e_I^* \otimes \mathbf{1}_I(t)\langle f\rangle_I = \sum_{k=0}^n \sum_{\substack{I\in\mathscr{D}_k[0,1)\\ I\ni t}} r_I e_I^* \otimes \langle f\rangle_I,$$

which contains exactly one interval I of each given length $\ell(I)$ such that $2^{-n} \leqslant \ell(I) \leqslant 1$. Then it is clear that the $\mathscr{L}(\ell^2(\mathscr{D}[0,1)), L^2(\Omega;X))$-norm of this quantity at each t, and therefore also its $L^p(0,1;\mathscr{L}(\ell^2(\mathscr{D}[0,1)), L^2(\Omega;X)))$ norm, is invariant under replacing $r_I e_I^*$ by $r_k e_k^*$ with $k := -\log_2(\ell(I))$, and $\ell^2(\mathscr{D}[0,1))$ by ℓ^2_{n+1}. This shows the equivalence (2)⇔(3). □

Examples of spaces with and without the RMF property

We conclude our discussion of the Rademacher maximal function with several examples of its behaviour in concrete spaces.

Example 3.6.13 (Type 2 *implies RMF).* A Banach space has *type 2* if there exists a constant $\tau_{2,X}$ such that for all finite sequences $(x_n)_{n=1}^N$ in X we have

$$\mathbb{E}\Big\|\sum_{n=1}^N r_n x_n\Big\|^2 \leqslant \tau_{2,X}^2 \sum_{n=1}^N \|x_n\|^2.$$

Trivially, every martingale type 2 space (as defined in Definition 3.5.23) has type 2. In particular, Hilbert spaces and L^p-spaces, $2 \leqslant p < \infty$, have type 2. The notion of type and the related dual notion of cotype will be taken up in more depth in the Volume II.

If X has type 2 and $(f_k)_{k=0}^n$ is a martingale with values in X, then $M_{\mathrm{Rad}} f \leqslant \tau_{2,X} f^\star$, where $f^\star$ is the Doob maximal function. Indeed, we have the pointwise estimate

$$\begin{aligned} M_{\mathrm{Rad}} f &= \sup_{\|\lambda\|_{\ell^2}\leqslant 1} \Big(\mathbb{E}\Big\|\sum_{k=0}^n r_k \lambda_k f_k\Big\|_X^2\Big)^{1/2} \\ &\leqslant \tau_{2,X} \sup_{\|\lambda\|_{\ell^2}\leqslant 1} \Big(\sum_{k=0}^n |\lambda_k|^2 \|f_k\|_X^2\Big)^{1/2} = \tau_{2,X} \sup_{0\leqslant k\leqslant n} \|f_k\|_X = \tau_{2,X} f^\star. \end{aligned}$$

That X has the RMF property then follows from Doob's maximal inequality.

This example shows, among other things, that L^p-spaces with $2 \leqslant p < \infty$ enjoy the RMF property. The next example shows that ℓ^p has the RMF property for all $1 < p < \infty$. Using Rubio de Francia's extension of the Fefferman–Stein theorem, the same method of proof shows that in fact every UMD Banach function space (in particular, every L^p-space with $1 < p < \infty$) has the RMF property. We refer to the Notes for a discussion.

Example 3.6.14 (ℓ^p has the RMF property for $1 < p < \infty$*).* As an application of the Fefferman–Stein theorem (Theorem 3.2.28) we will show that ℓ^p enjoys the RMF property for all $1 < p < \infty$.

Fix $1 < p < \infty$. For a locally integrable function $u : [0, 1) \to \ell^p$, the (dyadic) *lattice maximal function* $M_{\mathrm{lat}} u$ is defined by

$$M_{\mathrm{lat}} u(t) := \sup_{I\ni t} |\langle u\rangle_I|, \quad t \in [0, 1),$$

where the supremum extends over all dyadic intervals in $[0, 1)$ and $|\cdot|$ is the coordinate-wise modulus in ℓ^p, i.e., $|(u_k)_{k\in\mathbb{Z}}| := (|u_k|)_{k\in\mathbb{Z}}$. By the Fefferman–Stein theorem, we know that if $u \in L^2(0, 1; \ell^p)$, then $M_{\mathrm{lat}} u \in L^2(0, 1; \ell^p)$ and $\|M_{\mathrm{lat}} u\|_{L^2(0,1;\ell^p)} \lesssim \|u\|_{L^2(0,1;\ell^p)}$.

We also consider the version of the Rademacher maximal function

$$M_{\mathrm{Rad}}u(t) := \sup_{\lambda \in \overline{B}_{\ell^2(\mathscr{D}[0,1))}} \Big(\mathbb{E}\Big\| \sum_{\substack{I\in\mathscr{D}[0,1)\\ I\ni t}} \varepsilon_I \lambda_I \langle f\rangle_I \Big\|_{\ell^p}^2\Big)^{1/2}$$

that was implicit in Corollary 3.6.12. By the Kahane–Khintchine inequality (Theorem 3.2.23), both in ℓ^p and coordinate-wise in $\mathbb{K}$ after Fubini's theorem), we have

$$\Big(\mathbb{E}\Big\| \sum_{\substack{I\in\mathscr{D}[0,1)\\ I\ni t}} r_I \lambda_I \langle u\rangle_I \Big\|_{\ell^p}^2\Big)^{1/2} \leqslant \kappa_{2,p}\kappa_{p,2} \Big\|\Big(\sum_{\substack{I\in\mathscr{D}[0,1)\\ I\ni t}} |\lambda_I|^2 |\langle u\rangle_I|^2\Big)^{1/2}\Big\|_{\ell^p},$$

and clearly

$$\Big\|\Big(\sum_{\substack{I\in\mathscr{D}[0,1)\\ I\ni t}} |\lambda_I|^2 |\langle u\rangle_I|^2\Big)^{1/2}\Big\|_{\ell^p} \leqslant \Big(\sum_{\substack{I\in\mathscr{D}[0,1)\\ I\ni t}} |\lambda_I|^2\Big)^{1/2} \Big\| \sup_{\substack{I\in\mathscr{D}[0,1)\\ I\ni t}} |\langle u\rangle_I| \Big\|_{\ell^p}.$$

Taking the supremum over $\lambda \in \overline{B}_{\ell^2(\mathscr{D}[0,1))}$, we get the pointwise bound

$$M_{\mathrm{Rad}}u(t) \leqslant \kappa_{2,p}\kappa_{p,2}\|M_{\mathrm{lat}}u(t)\|_{\ell^p}.$$

Taking $L^2(0,1)$-norms, this gives

$$\|M_{\mathrm{Rad}}u\|_{L^2(0,1)} \lesssim \|M_{\mathrm{lat}}u\|_{L^2(0,1;\ell^p)} \lesssim \|u\|_{L^2(0,1;\ell^p)}$$

by the Fefferman–Stein theorem in the last step.

To conclude we show that ℓ^1 does not have the RMF property.

Example 3.6.15 (Failure of RMF for ℓ^1). Fix $n \geqslant 1$ and set

$$f_n := \sum_{j=1}^{2^n} \mathbf{1}_{[(j-1)2^{-n}, j2^{-n})} \otimes e_j,$$

where e_j is the jth unit vector of ℓ^1. Then $\|f_n\|_{L^p(0,1;\ell^1)} = 1$ for all $p \in [1,\infty]$. For $k = 0,\dots,n$ let $\mathscr{F}_k$ be the σ-algebra generated by the sets $[(j-1)2^{-k}, j2^{-k})$, $j = 1,\dots,2^k$, and let $f_k := \mathbb{E}(f_n|\mathscr{F}_k)$. For $t \in [0,2^{-n})$ we have

$$f_{n-j}(t) = \frac{1}{2^j}\sum_{k=1}^{2^j} e_k, \qquad j = 0,1,\dots,n.$$

For other $t \in [0,1)$ we have similar expressions, but with a shift in the index of the e_k's.

Now let $n = 2m$, and consider the sequence λ given by $\lambda_{n-2i} = m^{-1/2}$, $i = 1,\dots,m$, and $\lambda_j = 0$ otherwise. Then clearly $\|\lambda\|_{\ell^2} = 1$, and hence for $0 < t < 2^{-n}$, we have

$$
\begin{aligned}
M_{\mathrm{Rad}} f_n(t) &\geqslant \mathbb{E}\Big\| \sum_{i=1}^m \varepsilon_i \lambda_{n-2i} f_{n-2i}(t) \Big\|_{\ell^1} \\
&= m^{-1/2} \mathbb{E}\Big\| \sum_{i=1}^m \varepsilon_i \frac{1}{2^{2i}} \sum_{k=1}^{2^{2i}} e_k \Big\|_{\ell^1} \\
&\geqslant m^{-1/2} \Big(\mathbb{E}\Big\| \sum_{i=1}^m \varepsilon_i \frac{1}{2^{2i}} \sum_{k=2^{2(i-1)}+1}^{2^{2i}} e_k \Big\|_{\ell^1} - \sum_{i=1}^m \frac{1}{2^{2i}} 2^{2(i-1)} \Big) \\
&= m^{-1/2} \Big(\sum_{i=1}^m \frac{2^{2i} - 2^{2(i-1)}}{2^{2i}} - \sum_{i=1}^m 2^{2(i-1)-2i} \Big) \\
&= m^{-1/2} \sum_{i=1}^m (1 - 2 \cdot 2^{-2}) = m^{-1/2} \sum_{i=1}^m \frac{1}{2} = \frac{\sqrt{m}}{2} = \frac{\sqrt{n}}{2\sqrt{2}}.
\end{aligned}
$$

A similar computation (the only difference being in the labels of the e_k) holds for all $t \in [0,1)$, and taking the $L^p(0,1)$ norm shows that

$$
\|M_{\mathrm{Rad}} f_n\|_{L^p(0,1)} \geqslant \frac{\sqrt{n}}{2\sqrt{2}} = \frac{\sqrt{n}}{2\sqrt{2}} \|f_n\|_{L^p(0,1;\ell^1)}.
$$

Since the same construction can be repeated with arbitrarily large n, we see that no L^p bound can hold for M_{Rad} in ℓ^1.

This example can be strengthened to show that the RMF property implies non-trivial type. We refer to the Notes for more details.

3.6.c Approximate models for martingale transforms

In this subsection, we provide an approximation procedure somewhat parallel to that of Theorem 3.6.1 but with the added element that we simultaneously approximate a given martingale and a transforming predictable sequence.

As in Theorem 3.6.1, we deal finite martingales $(f_k)_{k=0}^n$ with respect to a σ-finite filtration $(\mathscr{F}_k)_{k=0}^n$ of a measure space $(S, \mathscr{A}, \mu)$. In this context, a predictable sequence $(v_k)_{k=0}^n$ consists of functions such that v_k is $\mathscr{F}_{k-1}$-measurable for $k \geqslant 1$, and $\mathscr{F}_0$-measurable for $k = 0$. Note that this agrees with our conventions in the case of filtrations indexed by $\{-\infty\} \cup \mathbb{Z}$ if we re-index $-\infty$ by 0.

For the transform of $f = (f_k)_{k=0}^n$ by $v = (v_k)_{k=0}^n$, we here adopt the notation

$$
(v \star f)_k := v_0 f_0 + \sum_{j=1}^k v_j df_j, \qquad k = 0, \dots, n.
$$

This emphasises the bilinear dependence of the transform on both v and f, which is natural in the present context, where both quantities will be allowed to vary; this is in contrast with, say, Theorem 3.5.4, which dealt with a fixed transforming sequence throughout the argument.

The main result of this subsection is the following result of arguably technical nature; the full reward for its understanding will only be provided in the following chapter. Note that we here aim at a slightly different class of model objects than the Paley–Walsh martingales used in Theorem 3.6.1, the desired property being that of *incrementality* defined in the statement of the proposition.

Proposition 3.6.16. *Let $(f_k)_{k=0}^n$ be a martingale in $L^p(S;X)$, and $(v_k)_{k=0}^n$ a scalar-valued bounded predictable sequence with respect to the same σ-finite filtration $(\mathscr{F}_k)_{k=0}^n$ on a measure space $(S,\mathscr{A},\mu)$. For any $\varepsilon > 0$, there is another martingale $(g_j)_{j=0}^N$ and a predictable sequence $(w_j)_{j=0}^N$, as well as an increasing function $\phi : \{0,1,\dots,n\} \to \{0,1,\dots,\phi(n)=N\}$, such that:*

(i) *$(g_j)_{j=0}^N$ and $(w_j)_{j=0}^N$ are supported on a set $E \in \mathscr{F}_0$ of finite μ-measure,*
(ii) *their restrictions to E form a martingale and a predictable sequence with respect to a filtration $(\mathscr{G}_j)_{j=0}^N$ of finite σ-algebras of E,*
(iii) *the filtration is* incremental *in the sense that the number of generating atoms satisfies*

$$\#\mathscr{G}_j^* = \#\mathscr{G}_{j-1}^* + 1 \qquad \forall j = 1,\dots,N,$$

(iv) *we have the bound*

$$\sup_{0\leqslant j\leqslant N} \|w_j\|_\infty \leqslant \sup_{0\leqslant k\leqslant n} \|v_k\|_\infty,$$

(v) *$w_j = \lambda_j \mathbf{1}_E$ takes a constant value $\lambda_j \in \mathbb{K}$ on E, for all $j = 1,\dots,N$,*
(vi) *$\|f_k - g_{\phi(k)}\|_p < \varepsilon$ and $\|(v \star f)_k - (w \star g)_{\phi(k)}\|_p < \varepsilon$ for all $k = 0,\dots,n$.*

If v_0 is constant, we may take $\mathscr{G}_0 = \{\varnothing, E\}$. If $f_0 = 0$, we may take $g_0 = 0$.

Note that the constancy of the new sequence w_j asserted in (v) is more restrictive than its predictability asserted in (ii); it is part of the proposition that even this stronger property can be arranged.

We begin with a reduction to simple martingales and predictable sequences on a finite measure space, in parallel with Lemma 3.6.2.

Lemma 3.6.17. *Let $(f_k)_{k=0}^n$ be a martingale and $(v_k)_{k=0}^n$ be a scalar-valued bounded predictable sequence, both with respect to the same σ-finite filtration $(\mathscr{F}_k)_{k=0}^n$ on a measure space $(S,\mathscr{A},\mu)$. For any $\varepsilon > 0$, there is another martingale $(g_k)_{k=0}^n$ and a predictable sequence $(w_k)_{k=0}^n$ such that:*

(i) *$(g_k)_{k=0}^n$ and $(w_k)_{k=0}^n$ are supported on a set $E \in \mathscr{F}_0$ of finite μ-measure,*
(ii) *their restrictions to E form a martingale and a predictable sequence with respect to a filtration of finite σ-algebras of E,*
(iii) *$\|w_k\|_\infty \leqslant \|v_k\|_\infty$ for all $k = 0,1,\dots,n$,*
(iv) *$\|f_k - g_k\|_p < \varepsilon$ and $\|(v \star f)_k - (w \star g)_k\|_p < \varepsilon$ for all $k = 0,1,\dots,n$.*

If $f_0 = 0$, we may take $g_0 = 0$.

Proof. The proof is an elaborate version of that of Lemma 3.6.2.

Let f and v be a martingale and a predictable sequence with respect to the σ-finite filtration $(\mathscr{F}_k)_{k=0}^n$. Fix $\delta > 0$. By σ-finiteness and dominated convergence, we can find an $E \in \mathscr{F}_0$ of finite measure such that $\|\mathbf{1}_{\complement E} f_n\|_p + \|\mathbf{1}_{\complement E}(v \star f)_n\|_p < \delta$. Since $\mathbb{E}(\mathbf{1}_E \phi|\mathscr{F}_k) = \mathbf{1}_E \mathbb{E}(\phi|\mathscr{F}_k)$ for all $\phi \in \{f_j, v_j\}$ and all $j, k = 0, 1, \ldots, n$, it is immediate that the martingale and predictable properties of f and v are inherited by $(\mathbf{1}_E f_k)_{k=0}^n$ and $(\mathbf{1}_E v_k)_{k=0}^n$. Replacing f and v by these approximations in the first step, we shall henceforth assume that these functions are supported on the finite set E, and make all the remaining constructions inside this set only.

By density in both spaces, we can find simple functions $s_k \in L^p(E, \mathscr{F}_k; X)$ and $w_k \in L^\infty(E, \mathscr{F}_{(k-1)_+})$, such that:

- $\|s_0 - f_0\|_p < \delta$ (where we choose $s_0 = 0$ if $f_0 = 0$),
- $\|s_k - df_k\|_p < \delta$ for each $k = 1, \ldots, n$,
- $\|w_0 - v_0\|_\infty < \delta$ and $\|w_0\|_\infty \leqslant \|v_0\|_\infty$,
- $\|w_k - v_k\|_\infty < \delta$ and $\|w_k\|_\infty \leqslant \|v_k\|_\infty$ for each $k = 1, \ldots, n$.

We then define

- $\mathscr{G}_0 := \sigma(E, s_0, w_0, w_1) \subseteq \mathscr{F}_0$
- $\mathscr{G}_k := \sigma(\mathscr{G}_{k-1}, s_k, w_{(k+1)\wedge n}) \subseteq \mathscr{F}_k$ for each $k = 1, \ldots, n$,

so that $(\mathscr{G}_k)_{k=0}^n$ is a filtration of finite σ-algebras of E for which $(w_k)_{k=0}^n$ is predictable and $(s_k)_{k=0}^n$ is adapted.

Let $g_0 := s_0$ $(= 0$ if $f_0 = 0)$, and for $k \geqslant 1$, we let $dg_k := s_k - \mathbb{E}(s_k|\mathscr{G}_{k-1})$ be a martingale difference sequence relative to $(\mathscr{G}_k)_{k=0}^n$ on E. We have $\|g_0 - f_0\|_p < \delta$ and, for $k \geqslant 1$,

$$\mathbb{E}(df_k|\mathscr{G}_{k-1}) = \mathbb{E}(\mathbb{E}(df_k|\mathscr{F}_{k-1})|\mathscr{G}_{k-1}) = \mathbb{E}(0|\mathscr{G}_{k-1}) = 0,$$

which implies that

$$\|dg_k - df_k\|_p = \|s_k - df_k - \mathbb{E}(s_k - df_k|\mathscr{G}_{k-1})\|_p \leqslant 2\|s_k - df_k\|_p < 2\delta,$$

and then also

$$\begin{aligned}\|v_k df_k - w_k dg_k\|_p &\leqslant \|v_k - w_k\|_\infty \|df_k\|_p + \|w_k\|_\infty \|df_k - dg_k\|_p \\ &\leqslant \delta\|df_k\|_p + \|v_k\|_\infty \cdot 2\delta = (\|df_k\|_p + 2\|v_k\|_\infty)\delta,\end{aligned}$$

and similarly $\|v_0 f_0 - w_0 g_0\|_p \leqslant (\|f_0\|_p + 2\|v_0\|_\infty)\delta$.

Summing up the difference sequences, it is immediate that the condition (iv) is satisfied provided that the δ is chosen sufficiently small. □

After the previous reduction to finite σ-algebras, incrementality is reached simply by constructing finitely many intermediate σ-algebras between each $\mathscr{F}_k$ and $\mathscr{F}_{k+1}$. The details are given in the following lemma.

Lemma 3.6.18. *Let $(f_k)_{k=0}^n$ be a martingale and $(v_k)_{k=0}^n$ be predictable, both with respect to the same filtration $(\mathscr{F}_k)_{k=0}^n$ of finite σ-algebras. Then there are sequences $(g_j)_{j=0}^N$ and $(w_j)_{j=0}^N$ such that:*

(i) $(g_j)_{j=0}^N$ *is a martingale with respect to an incremental filtration* $(\mathscr{G}_j)_{j=0}^N$,
(ii) w_0 *is* $\mathscr{G}_0$*-measurable and* $w_j = \lambda_j$ *is a constant for all* $j = 1, \ldots, N$,
(iii) $f_k = g_{\phi(k)}$ *and* $(v \star f)_k = (w \star g)_{\phi(k)}$ *for all* $k = 0, 1, \ldots, N$, *where* $\phi : \{0, 1, \ldots, n\} \to \{0, 1, \ldots, \phi(n) = N\}$ *is increasing, and*
(iv) *the range of values of* $(w_j)_{j=0}^N$ *is a subset of the values of* $(v_k)_{k=0}^n$.

If v_0 *is a constant, we may in addition arrange that* $\mathscr{G}_0$ *is trivial. If* $f_0 = 0$, *we may take* $g_0 = 0$.

Proof. Let $\phi(k) := \#\mathscr{F}_k^* - \#\mathscr{F}_0^*$ be the increase of the number of atoms from the initial σ-algebra to the kth one, and denote $\mathscr{G}_{\phi(k)} := \mathscr{F}_k$ for each $k = 0, 1, \ldots, n$. For the intermediate values $j = \phi(k) + 1, \ldots, \phi(k+1) - 1$, we define recursively $\mathscr{G}_j := \sigma(\mathscr{G}_{j-1}, A)$, picking any atom $A \in \mathscr{G}_{\phi(k+1)}^* \setminus \mathscr{G}_{j-1}^*$. It is routine to check that this produces an incremental filtration $(\mathscr{G}_j)_{j=0}^N$, with $N := \phi(n)$.

We define $g_j := \mathbb{E}(f_n|\mathscr{G}_j)$ for all $j = 0, 1, \ldots, N$, which gives a martingale such that $g_{\phi(k)} = f_k$. In particular, $g_0 = g_{\phi(0)} = f_0$. We also let $w_0 := v_0$ and $\widetilde{w}_j := v_k$ for $\phi(k-1) < j \leqslant \phi(k)$, so that this is measurable with respect to $\mathscr{F}_{k-1} = \mathscr{G}_{\phi(k-1)} \subseteq \mathscr{G}_{j-1}$. (The tilde signifies that we still need to modify this construction slightly to get the final w_j as in the assertion.) We have

$$v_k df_k = v_k(g_{\phi(k)} - g_{\phi(k-1)}) = \sum_{j:\phi(k-1)<j\leqslant\phi(k)} \widetilde{w}_j dg_j$$

and summing over k we deduce that $(v \star f)_k = (\widetilde{w} \star g)_{\phi(k)}$.

Finally, we want to see that $\widetilde{w}_j$ may be replaced by a constant, without changing the other properties already achieved. To this end, from the condition that $\#\mathscr{G}_j^* = \#\mathscr{G}_{j-1}^* + 1$ it follows that $\mathscr{G}_j$ and $\mathscr{G}_{j-1}$ share the same atoms, except that one atom $A \in \mathscr{G}_{j-1}^*$ is a union of two atoms $A_0, A_1 \in \mathscr{G}_j^*$. From $\mathbb{E}(dg_j|\mathscr{G}_{j-1}) = 0$ and the $\mathscr{G}_{j-1}$-measurability of $\widetilde{w}_j$, it follows that dg_j is supported on A, and $\widetilde{w}_j$ is constant on this set. Thus $\widetilde{w}_j dg_j = \lambda_j dg_j$, where λ_j is the constant value of $\widetilde{w}_j$ on A. Hence $\sum_{j=1}^N \widetilde{w}_j dg_j = \sum_{j=1}^N \lambda_j dg_j$ as claimed, and we may set $w_0 := \widetilde{w}_0$, $w_j := \lambda_j$ to complete the construction.

If v_0 (and hence $w_0 = v_0$) is a constant, then we define $\mathscr{G}_{-1}$ as the trivial σ-algebra. Then

$$w_0 g_0 = w_0(g_{-1} + dg_0) =: w_{-1} g_{-1} + w_0 dg_0, \ \ g_{-1} := \mathbb{E}(g_0|\mathscr{G}_{-1}), \ \ dg_0 := g_0 - g_{-1},$$

where $w_{-1} := w_0$ is a constant, thus $\mathscr{G}_{-1}$-measurable. It follows that the values of the martingale transform $(w \star g)_k$, for $k \geqslant 0$, remain unchanged on replacing w and g by the extended sequences starting at $j = -1$; we also have $g_{-1} = 0$ if $g_0 = 0$. Obviously we can then re-index these sequences to start from 0, so that the assertions of the Lemma are satisfied with (ϕ, N) replaced by $(\phi(\cdot) + 1, N + 1)$. $\square$

Proof of Proposition 3.6.16. We apply Lemma 3.6.17 to produce an approximating martingale and predictable sequences relative to a filtration of finite σ-algebras of a subset $E \in \mathscr{F}_0$ of finite measure, and then Lemma 3.6.18 to make them a sampling of a martingale relative to an incremental filtration, and a constant predictable sequence. The detailed properties are immediate from the combination of the corresponding statements in the two lemmas. □

3.7 Notes

The theory of martingales is the creation of Doob, whose monograph Doob [1953] contains already most of the classical results. A full account is given in the monumental treatise Doob [1984]. Standard introductions to the subject include Neveu [1975] and Williams [1991]. More advanced treatments can be found in Liptser and Shiryayev [1989], Revuz and Yor [1999], Rogers and Williams [1994].

Section 3.1

The study of martingales with values in a Banach space goes back to Chatterji [1960, 1964]. A detailed treatment is given in Pisier [2016]. An account on the history of vector-valued martingales and its connections with the Radon–Nikodým theorem can be found in Diestel and Uhl [1977].

Section 3.2

Doob's maximal inequality (Theorem 3.2.2) is due to Doob [1953], where it is stated that the L^1-inequality for martingales (rather than submartingales) "was first used by Lévy and Ville" (Doob [1953, p. 630]).

The dual version of Doob's inequality in Proposition 3.2.8 is a special case of a convexity result in Burkholder, Davis, and Gundy [1972, Theorem 3.2] (also see Kallenberg [2002, Proposition 25.21] for a continuous generalisation and alternative proof). It admits a far-reaching generalisation in the context of non-commutative martingales (see Junge [2002]).

The contraction principle of Proposition 3.2.10 was obtained in Kahane [1985]. To obtain the extension to the complex setting with constant $\pi/2$ we follow the presentation in Pietsch and Wenzel [1998, 3.5.4].

The John–Nirenberg inequality (Theorem 3.2.17) for adapted sequences is due to Herz [1974]; the classical form of the inequality (Theorem 3.2.30), which we have derived from the martingale version, is from John and Nirenberg [1961]. The present treatment of the BMO spaces, both in the martingale and Euclidean settings, should be seen as an illustration of stopping time techniques, not as a systematic discussion of the theory. Most notably, we have bypassed the duality of BMO with the Hardy space H^1, whose martingale version is also studied by Herz [1974]. For a systematic treatment of the martingale Hardy and BMO spaces, we refer to the books of Garsia [1973], Long [1993] and Müller [2005]. In particular, Chapter 7 of Long [1993] also has an extensive discussion of the regular filtrations mentioned in Remark 3.2.18.

The Kahane–Khintchine inequality (Theorem 3.2.23) is Kahane's vector-valued extension of the classical scalar-valued version due to Khintchine [1923]. A modern version of Khintchine's original proof is in Garling [2007]. A tail estimate that captures the essence of the vector-valued estimate is contained in Kahane [1964]; see also the book of Kahane [1985]. We shall revisit this fundamental inequality with a different approach in Volume II, where the problem of finding the correct qualitative behaviour of the constants will also be addressed.

The shifted dyadic systems of Definition 3.2.25 and the related Covering Lemma 3.2.26 appear in various versions in the literature, and there has been confusion about their somewhat misty origins. A nice compilation of the relevant history has been collected by Cruz-Uribe [2014] (see the lengthy footnote following Theorem 3.4), who gives the primary credit to Okikiolu [1992] and, in a slightly weaker version, to Chang, Wilson, and Wolff [1985]. Several variants of Lemma 3.2.26 in the literature take the following generic form:

> In $\mathbb{R}^d$, there are N_d shifted dyadic systems $\mathscr{D}^k$, $k = 1, \dots, N$ (not necessarily of exactly the same form as in Definition 3.2.25) such that every cube $Q \subseteq \mathbb{R}^d$ is contained in $D \in \bigcup_{k=1}^N \mathscr{D}^k$ with $\ell(Q) \leqslant C_d \ell(D)$.

Our statement with $N_d = 3^d$ and $C_d = 3$ is equal to that of Cruz-Uribe [2014, Theorem 3.4], but one can make a trade-off between the size of these two numbers. The present choice is reasonably efficient for applications like Theorem 3.2.27, where the factor that enters is $N_d \cdot (C_d)^d = 9^d$ (cf. (3.36)).

Implicit in the proof of Theorem 3.2.30 is the fact that the standard BMO space on $\mathbb{R}^d$ is an intersection of finitely many dyadic BMO spaces, a representation studied by Mei [2003, 2007]. In such a context it is of some interest to minimise the number N_d of the dyadic versions, and a somewhat more elaborate argument shows the possibility of taking $N_d = d+1$. See Mei [2007, Proposition 3.1] for $d = 1$ and Hytönen [2010, Lemma 5.3] or Conde [2013] for general d; the key idea and a periodic version goes back to Mei [2003]. However, this comes at the cost of a large C_d, estimated as $C_d = 10d$ in Hytönen [2010], which would lead to a rather fast growth of $N_d \cdot (C_d)^d$.

While Definition 3.2.25 of (shifted or not) dyadic cubes relies heavily on the coordinate geometry of the Euclidean space $\mathbb{R}^d$, it admits a far-reaching generalisation in the context of general *spaces of homogeneous type*, an abstract setting for harmonic analysis introduced by Coifman and Weiss [1971]. These are quasi-metric spaces S equipped with a Borel measure μ for which the quasi-metric balls $B(x,r)$ obey the *doubling property*

$$\mu(B(x,2r)) \leqslant C_\mu \mu(B(x,r)) \tag{3.69}$$

with a constant C_μ independent of $x \in S$ and $r > 0$. In this generality, a construction of sets that share the essential properties of Euclidean dyadic cubes has been given by Christ [1990]. Abstract analogues of shifted dyadic systems, which satisfy a version of the Covering Lemma 3.2.26, were first constructed by Hytönen and Kairema [2012], which contains versions of the results of Subsection 3.2.d in spaces of homogeneous type.

Theorem 3.2.28 is from Fefferman and Stein [1971]. A generalisation where the role of ℓ^q is replaced by a UMD Banach space with a normalised unconditional basis X is given in Bourgain [1984b] and, for an arbitrary UMD Banach function space X, in Rubio de Francia [1986]. We will take up this topic in the Notes of Chapter 4 (in particular see Theorem 4.6.4). For more information and results in weighted spaces we refer to García-Cuerva, Macías, and Torrea [1993] and Tozoni [1996]. We have showed how to deduce Doob's maximal inequality in ℓ^q (see Theorem 3.2.7) from Theorem 3.2.28, which then shows that both result are equivalent. Indeed, this simply follows by using the fact that it suffices to obtain the estimate for Paley–Walsh martingales (see Corollary 3.6.7) which then directly follows from the fact that the dyadic maximal function is dominated by the Hardy–Littlewood maximal function. The same procedure can be followed in any UMD Banach function space, yielding a version of Doob's maximal inequality in UMD Banach spaces.

The alternative approach to Theorem 3.2.28 presented in Example 3.5.8, based on the martingale extrapolation theorem, goes back to Martínez and Torrea [2000]. Yet another approach, via Muckenhoupt A_p-weights, is due to Rubio de Francia [1984] and was hinted at in the Notes of Chapter 2. The main idea consists of the following steps:

1. Proving that for $1 < p < \infty$ the Hardy–Littlewood maximal function is bounded on $L^p(\mathbb{R}^d, w)$ for all A_p-weights w on $\mathbb{R}^d$;
2. Proving that if the functions F and G satisfy, for some $1 < p < \infty$, then
$$\|F\|_{L^p(\mathbb{R}^d,w)} \leqslant C\|G\|_{L^p(\mathbb{R}^d,w)}$$
for all A_p-weights w on $\mathbb{R}^d$, with a constant $C = C_{A_p(w)}$ only depending on the A_p-constant of w in a suitable 'monotone' way, then they satisfy, for all $1 < q < \infty$, the estimate
$$\|F\|_{L^q(\mathbb{R}^d,w)} \leqslant C'\|G\|_{L^q(\mathbb{R}^d,w)}$$
for all A_q-weights w on $\mathbb{R}^d$, with a constant C' depending on C, p and on the A_q-constant of w.
3. Applying this to $F := (\sum_{n\geqslant 1}(|Mf^{(n)}|^q)^{1/q}$ and $G := (\sum_{n\geqslant 1}(|f^{(n)}|^q)^{1/q}$.

We recommend Cruz-Uribe, Martell, and Pérez [2011] for an especially lucid presentation of this circle of ideas.

Section 3.3

Parts of Theorem 3.3.2 and its corollary are contained in Chatterji [1960, 1964], the almost sure convergence assertion being due to Neveu [1965] after previous contributions by Chatterji [1960, 1964], Scalora [1961]. The principal difficulty in establishing almost everywhere convergence was that the standard proof involving up-crossings in the scalar-valued case doesn't make sense in the Banach space-valued context. The paper Chatterji [1964] also contains the essentials of Theorem 3.3.5.

Theorem 3.3.9, for $X = \mathbb{K}$, is due to Jessen [1934]. The introduction of Ross and Stromberg [1967] has a discussion of both positive and negative results related to possible extensions of the theorem. Our martingale proof of the strong law of large numbers (Theorem 3.3.10) goes back to Doob [1953]. Several variants of the strong law of large numbers in Banach spaces can be found in Ledoux and Talagrand [1991, Chapter 7].

The Itô-Nisio theorem for sums of independent random variables is due to Itô and Nisio [1968] and does not require any integrability assumptions. It will be discussed in Volume II. Its version for martingales is due to Davis, Ghoussoub, Johnson, Kwapien, and Maurey [1990]. In this paper two proofs are presented, one of which we follow here.

Theorem 3.3.16 goes back to Chatterji [1960, 1964], Ionescu Tulcea and Ionescu Tulcea [1963], Scalora [1961], who obtained it for martingales with values in reflexive Banach spaces. Ionescu Tulcea and Ionescu Tulcea [1963] also obtained the result for separable duals. The extension to spaces with the Radon–Nikodým property is due to Chatterji [1968]. Further results are contained in Métivier [1967], Uhl [1969a,b]. Our proof of Theorem 3.3.16 is taken from Burkholder [2001].

Section 3.4

The Gundy decomposition (Theorem 3.4.1) originates from Gundy [1968]. It is a probabilistic counterpart of the celebrated decomposition of Calderón and Zygmund [1952] from their pioneering work on singular integrals. The Gundy decomposition in the σ-finite case is used in Stein [1970b] to derive Littlewood–Paley inequalities. A noteworthy contribution of Gundy's decomposition is that, like many results in probability, it makes essentially no assumptions on the structure of the underlying filtered space, whereas the Calderón–Zygmund decomposition, in its original form, depends quite critically on the doubling property (3.69) of the Lebesgue measure, or, what is essentially the same, the *regularity* of the dyadic filtration, in the sense of Remark 3.2.18. In view of this early achievement of Gundy [1968], it is interesting that the development of appropriate Calderón–Zygmund type decompositions, suited for the study of singular integrals (instead of martingale transforms) with respect to non-doubling measures on $\mathbb{R}^d$, has taken much longer. A prototypical form of such decompositions is due to Tolsa [2001]. A pleasingly explicit decomposition closely related to Gundy's is as recent as López-Sánchez, Martell, and Parcet [2014]; it was devised for the analysis of certain dyadic operators in the extended family of dyadic martingale transforms.

The Davis decomposition originates from Davis [1970], where it was introduced in order to prove the square function estimates for discrete martingales of Burkholder [1966] in the limiting case $p = 1$. An extension of these square function estimates for martingales was given in Burkholder [1973] and since then it has become the standard tool for obtaining such estimates.

Section 3.5

The systematic theory of martingale transforms was pioneered by Burkholder [1966], who notes that "Such transforms, particularly in the case in which the v_n may take only 0 and 1 as possible values, have a long history and an interesting gambling interpretation."

Burkholder [1966] already contains proofs of some of the main statements of Theorem 3.5.4 for scalar-valued functions (in which case the L^2-boundedness can be easily verified, so that the several boundedness properties formulated in Theorem 3.5.4 are not only equivalent to each other but simply true). The so-called good-λ inequality and application to extrapolation theory for martingale inequalities have been developed by Burkholder and Gundy [1970]. In the vector-valued setting this has been studied in Burkholder [1981a] and more details on this will be discussed in Chapter 4. The extrapolation result Theorem 3.5.4 in the stated generality is essentially due to Martínez and Torrea [2000] (see also Tozoni [1995] where a weighted setting is considered). The explicit documentation of the constants is probably new. The operator-valued setting has the advantage that several types of martingale inequalities can be put in a martingale transform framework (e.g., maximal estimates, square function estimates). A characterisation of the boundedness of a fixed operator-valued martingale transform T_v has been obtained in Girardi and Weis [2005] in terms of the R-boundedness of the range of v, a notion that we will discuss in Volume II. An extrapolation result for martingale transforms from BMO-L^∞ estimates toward L^p-L^p estimates can be found in Geiss [1997].

The notions of martingale type and martingale cotype were introduced in Pisier [1975a] where deep connections with convexity and smoothness properties of the Banach space were uncovered. We shall review them briefly.

A Banach space X is said to be *uniformly smooth* if for all $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x, d \in X$ with $\|x\| \leqslant 1$ and $\|d\| < \delta$ we have

$$\|x + d\| + \|x - d\| \leqslant 2 + \varepsilon\|d\|.$$

The space X is called *p-smooth*, where $p \in [1, 2]$, if its *modulus of smoothness*

$$\rho_X(t) := \sup\Big\{\frac{1}{2}(\|x + ty\| + \|x - ty\|) - 1 \,:\, \|x\| = \|y\| = 1\Big\}$$

satisfies $\rho_X(t) \leqslant Ct^p \quad \forall t > 0$. A Banach space X is said to be *uniformly convex* if every $0 < \varepsilon \leqslant 2$ there exists a $\delta > 0$ such that for any two vectors

$x, y \in X$ with $\|x\| = \|y\| = 1$ satisfying $\|x - y\| < \varepsilon$ we have $\frac{1}{2}\|x + y\| < 1 - \delta$. The space X is called *q-convex*, where $q \in [2, \infty)$, if its *modulus of convexity*

$$\delta_X(t) := \inf \Big\{ 1 - \frac{1}{2}\|x + y\| \, : \, \|x\| \leqslant 1, \ \|y\| \leqslant 1, \ \|x - y\| \geqslant t \Big\}$$

satisfies $\delta_X(t) \geqslant Ct^q \quad \forall 0 < t \leqslant 2$.

Theorem 3.7.1 (Pisier). *Let X be a Banach space.*

(1) *For any $p \in [1, 2]$ the following assertions are equivalent:*
  (a) *X has martingale type p;*
  (b) *X has Paley–Walsh martingale type p;*
  (c) *X admits an equivalent p-smooth norm.*
(2) *For any $q \in [2, \infty]$ the following assertions are equivalent:*
  (a) *X has martingale cotype q;*
  (b) *X has Paley–Walsh martingale cotype q;*
  (c) *X admits an equivalent q-convex norm.*

In both parts (b), the Paley–Walsh versions are obtained by restricting to Paley–Walsh martingales in the relevant definitions. In Pisier [2016], a martingale characterisation of super-reflexivity (the definition of this notion can be found in Section 4.3.a) is included as well. In this connection one should also mention the classical papers James [1972a] and Enflo [1972].

Wenzel [2005b] introduces the notions of *strong martingale type* and *strong martingale cotype* and shows that these are equivalent to uniform smoothness and uniform convexity, respectively.

The class of 2-smooth Banach spaces plays a distinguished role in various problems in stochastic analysis. Their usefulness was probably first noticed by Neidhardt [1978], who extended Itô's theory of stochastic integration to the setting of 2-smooth Banach spaces. This theory was further developed in a series of papers by Dettweiler (see, e.g., Dettweiler [1991]) and Brzeźniak (see, e.g., Brzeźniak [1995, 1997]). Exponential inequalities for martingales with values in 2-smooth spaces have been obtained by Pinelis [1994]. Maximal inequalities for stochastic convolutions in 2-smooth spaces were obtained in Brzeźniak and Peszat [2000], Seidler [2010], Van Neerven and Zhu [2011]; see Veraar and Weis [2011] for related results.

In Example 3.5.7 we have seen that if $(v_n)_{n \in \mathbb{Z} \cup \{-\infty\}}$ is a strongly predictable uniformly bounded sequence with values in $\mathscr{L}(H, K)$, where H and K are Hilbert spaces, then the corresponding martingale transform T_v is L^2-bounded. For Banach spaces X and Y one has the following generalisation:

Proposition 3.7.2. *Let X and Y be non-zero Banach spaces and let $p \in (1, \infty)$. The following assertions are equivalent:*

(1) *for every strongly predictable uniformly bounded sequence $(v_n)_{n \in \mathbb{Z} \cup \{-\infty\}}$ with values in $\mathscr{L}(X, Y)$, the corresponding martingale transform T_v is L^p-bounded;*

(2) *X has martingale cotype* 2 *and Y has martingale type* 2*.*

In particular, X has type 2 and Y has cotype 2. If $X = Y$, a theorem of Kwapień [1972a], which will be proved in Volume II, then implies that X is isomorphic to a Hilbert space. In this form the result is equivalent to Burkholder [2001, Theorem 14(i)].

The proof of the proposition makes use of the fact that the class of martingale (co)type spaces remains unchanged if in their definitions one only considers Paley–Walsh martingales; this follows from Theorem 3.7.1.

Proof of Proposition 3.7.2. (2)⇒(1): For $f \in L^p(S;X)$ with finitely non-zero difference sequence and v as in (1), it follows from Propositions 3.5.25 and 3.5.27 that

$$\begin{aligned}\|T_v(f)\|_{L^p(S;Y)} &\leqslant C_Y \Big\| \Big(\|v_{-\infty} f_{-\infty}\|_X^2 + \sum_{n\in\mathbb{Z}} \|v_n df_n\|_X^2 \Big)^{1/2} \Big\|_{L^p(S)} \\ &\leqslant C_Y \Big\| \Big(\|f_{-\infty}\|_X^2 + \sum_{n\in\mathbb{Z}} \|df_n\|_X^2 \Big)^{1/2} \Big\|_{L^p(S)} \\ &\leqslant C_Y C_X \|f\|_{L^p(S;X)}.\end{aligned}$$

(1)⇒(2): By the extrapolation result Theorem 3.5.4 it suffices to consider the case $p = 2$. Moreover it suffices to prove that X has Paley–Walsh martingale cotype 2 and Y has Paley–Walsh martingale type 2.

Let $(r_n)_{n\geqslant 1}$ be a Rademacher sequence on a probability space $(S, \mathscr{A}, \mu)$ with $\mathscr{A} = \sigma(r_n : n \geqslant 1)$. By the closed graph theorem there exists a constant C such that for all sequences $v \in \ell^\infty(L^\infty_{\rm so}(S; \mathscr{L}(X,Y)))$,

$$\|T_v\|_{L^p(S;X)} \leqslant C \sup_{n\geqslant 0} \|v_n\|_{L^\infty_{\rm so}(S;\mathscr{L}(X,Y))}. \tag{3.70}$$

In order to derive that X has martingale cotype 2 let $f : S \to X$ be a Paley–Walsh martingale. By Proposition 3.1.10 we may write $df_n = \phi_n(r_1, \dots, r_{n-1}) r_n$ where $\phi_n : \{-1,1\}^{n-1} \to X$ for $n \geqslant 1$ and let $\phi_0 = f_0$. Let $\psi_n : \{-1,1\}^{n-1} \to X^*$ satisfy $\|\psi_n\| = 1$ and $\langle \phi_n, \psi_n\rangle = \|\phi_n\|$ pointwise on $\{-1,1\}^{n-1}$. Observe that $\|df_n\| = \|\phi_n(r_1, \dots, r_{n-1})\|$ because of the Paley–Walsh structure. Fix an arbitrary norm one vector $y_0 \in Y$. Define $v_n : S \to \mathscr{L}(X,Y)$ by

$$v_n(x) := \langle x, \phi_n(r_1, \dots, r_{n-1})\rangle y_0.$$

Then $\|v_n\| \leqslant 1$ and therefore, by (3.70) we obtain

$$\|f_0\| + \sum_{n=1}^N \mathbb{E}\|df_n\|^2 = \mathbb{E}\|T_v(f_N)\|^2 \leqslant C\mathbb{E}\|f_N\|^2.$$

In order to derive that Y has martingale type 2 let $f : S \to Y$ be a Paley–Walsh martingale and let ϕ_n as before. Fix an arbitrary $x_0 \in X$ of norm one and pick a norm one element $x_0^* \in X^*$ such that $\langle x_0, x_0^*\rangle = 1$. Define $v_n : \Omega \to \mathscr{L}(X,Y)$ by

$$v_n(x) := \langle x, x_0^*\rangle \frac{\phi_n(r_1,\dots,r_n)}{\|\phi_n(r_1,\dots,r_n)\|}.$$

Then $\|v_n\| \leqslant 1$. Let g be the martingale with $g_0 = \|f_0\|x_0$ and difference sequence $dg_n = r_n\|\phi_n(r_1,\dots,r_{n-1})\|x_0$. By (3.70) we obtain

$$\mathbb{E}\|f_N\|^2 = \mathbb{E}\|T_v(g_N)\|^2 \leqslant C\mathbb{E}\|g_N\|^2 = C\Big(\mathbb{E}\|f_0\|^2 + \sum_{n=1}^N \mathbb{E}\|df_n\|^2\Big).$$

□

Section 3.6

The essential idea of Theorem 3.6.1 comes from Maurey [1975, Remarque 3], who indicates the “grandes lignes” of the “argument fastidieux”. We have tried to present the details of the proof as painlessly as possible, although a certain level of technicality seemed unavoidable.

The Rademacher maximal function and the related RMF property (Definitions 3.6.8 and 3.6.10) were introduced by Hytönen, McIntosh, and Portal [2008]. Most of the related results in Subsection 3.6.b are from the same paper, although Corollary 3.6.12 is essentially due to Kemppainen [2011]. Example 3.6.15 on the failure of RMF for ℓ^1 gives a quantitative improvement, as compared to Hytönen, McIntosh, and Portal [2008], on the growth of the RMF constants of the finite-dimensional spaces ℓ^1_n; nevertheless, the construction is only a slight variant of the original once.

Any of the following conditions is known to imply the RMF property for a Banach space X:

(a) X has type 2.
(b) X is a Banach function space with the UMD property.
(c) X is a non-commutative L^p-space on a von Neumann algebra $\mathscr{M}$ with a normal semi-finite faithful trace τ.

It is also shown in Hytönen, McIntosh, and Portal [2008] that every Banach space with the RMF property has non-trivial type. Further results on the RMF property are due to Hytönen and Kemppainen [2011] and Kemppainen [2011, 2013]. A non-tangential version of the RMF property has been introduced in Di Plinio and Ou [2015]. The main open question related to this notion is whether the UMD property studied in Chapter 4 implies RMF (see Problem O.13).

4

UMD spaces

In this chapter we will apply the methods developed so far to the study of a class of Banach spaces, the so-called UMD spaces, which in many ways provides the correct setting for vector-valued analysis. By definition, a Banach space X is a UMD space if X-valued martingale differences are unconditional in $L^p(\Omega; X)$ for all (equivalently, for some) $p \in (1, \infty)$.

Although the UMD property is defined probabilistically, it turns out to be equivalent to a wealth of purely analytic statements. This can be explained by the fact that many results in analysis depend on dyadic techniques, and the UMD property comes in helpfully as an unconditionality property of the martingales associated with dyadic filtrations. This idea will be pursued in depth in the next chapter. In the present chapter we limit ourselves to a presentation of the general theory of UMD spaces, mostly from a probabilistic point of view.

After an introduction in Section 4.1 developing the themes of unconditionality and Rademacher averages, we reap the fruits of our work on martingale transforms to obtain, in Section 4.2, various equivalent formulations of the UMD property, in particular the p-independence of its definition and the fact that it follows already from unconditionality of Paley–Walsh martingales. From these results one easily derives the basic examples of UMD spaces: Hilbert spaces, the spaces $L^p(S)$ for $1 < p < \infty$, and various spaces constructed from spaces that are already known to be UMD. We also show by means of direct quantitative arguments that any Banach space containing either the spaces ℓ^1_n or ℓ^∞_n uniformly fails the UMD property. As a direct consequence we obtain that many classical Banach spaces such as $C(K)$, $L^1(S)$, and $L^\infty(S)$ fail the UMD property (unless they are finite-dimensional). We also present some first consequences of the UMD property, among them Bourgain's extension of Stein's inequality for conditional expectations in UMD spaces and the boundedness of general martingale transforms in UMD spaces.

In Section 4.3 we show that the UMD property in fact implies reflexivity, along with various other Banach space properties, such as K-convexity and non-trivial martingale type.

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 63, DOI 10.1007/978-3-319-48520-1_4

Section 4.4 presents several decoupling inequalities for various classes sums of random variables with values in a UMD space. These will play an important role in the theory of stochastic integration in UMD spaces in Volume III. In order to go beyond the setting of random variables defined on a probability space, we also develop a more general approach to decoupling based on the notion of tangency, which can be applied in the context of σ-finite measure spaces typically arising in harmonic analysis.

The final section, Section 4.5, presents Burkholder's theory of zigzag-concave functions and its application to the determination of the exact values of the UMD constants $\beta_{p,\mathbb{R}}$, $\beta_{p,\mathbb{C}}$, and, more generally, $\beta_{p,H}$ for Hilbert spaces H.

4.1 Motivation

In order to motivate the introduction of the class of UMD spaces we begin by taking a look at two classical circles of results, both of which contain the idea of the UMD property in an embryonic form: square function estimates for Rademacher sums and unconditionality results for Schauder decompositions.

4.1.a Square functions for martingale difference sequences

Let $(\varepsilon_n)_{n\geqslant 1}$ be a Rademacher sequence defined on some probability space Ω. In Corollary 3.2.24 we have proved Khintchine's inequality, which asserts that for all $p \in (0,\infty)$ and all choices of scalars $c_1, \dots, c_N$ the following estimate holds:

$$A_p\Big(\sum_{n=1}^N |c_n|^2\Big)^{1/2} \leqslant \Big\| \sum_{n=1}^N c_n\varepsilon_n\Big\|_{L^p(\Omega)} \leqslant B_p\Big(\sum_{n=1}^N |c_n|^2\Big)^{1/2},$$

with constants $0 < A_p \leqslant B_p < \infty$ depending only on p. For $1 < p < \infty$, Khintchine's inequality admits the following far-reaching generalisation to a martingale setting. Recall that for a martingale $(f_n)_{n\geqslant 0}$ we define its martingale differences by $df_n := f_n - f_{n-1}$ for $n \geqslant 1$.

Theorem 4.1.1 (Burkholder). *For all $p \in (1,\infty)$ there exist constants $0 < A_p \leqslant B_p < \infty$ such that whenever $(f_n)_{n=0}^N$ is a scalar-valued L^p-martingale defined on a σ-finite measure space $(S, \mathscr{A}, \mu)$, then*

$$A_p\|f_N\|_{L^p(S)} \leqslant \Big\|\Big(|f_0|^2 + \sum_{n=1}^N |df_n|^2\Big)^{1/2}\Big\|_{L^p(S)} \leqslant B_p\|f_N\|_{L^p(S)}. \tag{4.1}$$

The middle expression is often referred to as the *square function* associated with $(f_n)_{n=0}^N$. Estimates of the type proved in this theorem are called *square function estimates*.

Proof. The right-hand side estimate in (4.1) can equivalently be formulated as the boundedness of the martingale transform $T_v : L^p(S) \to L^p(S;\ell^2_N)$ associated with $v_n : \mathbb{K} \to \ell^2_N$ given by $v_n(x) = xe_n$ for $n = 0, 1, \dots, N$. For $p = 2$, the boundedness of T_v follows from the orthogonality of the differences df_n (Proposition 3.5.6):

$$\|T_v f_N\|_{L^2(S;\ell^2_N)} = \Big\| \Big(|f_0|^2 + \sum_{n=1}^N |df_n|^2 \Big)^{1/2} \Big\|_{L^2(S)} = \|f_N\|_{L^2(S)}.$$

Theorem 3.5.4 implies that T_v is bounded for all $p \in (1,\infty)$. This proves the second estimate in (4.1), with $B_p = \|T_v\|_{\mathscr{L}(L^p(S),L^p(S;\ell^2_N))}$.

The first estimate in (4.1) follows via a duality argument from the second. Indeed, by Hölder's inequality we find, for $g_N \in L^{p'}(S,\mathscr{F}_N)$,

$$\begin{aligned}
|\langle f_N, g_N\rangle| &= \Big| \langle f_0, g_0\rangle + \sum_{n=1}^N \langle df_n, dg_n\rangle \Big| \\
&\leqslant \Big\| \Big(|f_0|^2 + \sum_{n=1}^N |df_n|^2 \Big)^{1/2} \Big\|_{L^p(S)} \Big\| \Big(|g_0|^2 + \sum_{n=1}^N |dg_n|^2 \Big)^{1/2} \Big\|_{L^{p'}(S)} \\
&\leqslant B_{p'} \Big\| \Big(|f_0|^2 + \sum_{n=1}^N |df_n|^2 \Big)^{1/2} \Big\|_{L^p(S)} \|g_N\|_{L^{p'}(S)},
\end{aligned}$$

where we applied (4.1) to g_N. The result now follows by taking the supremum over all $g_N \in L^{p'}(S,\mathscr{F}_N)$ of norm one. □

4.1.b Unconditionality

Various non-trivial results in probability theory and harmonic analysis can be rephrased as stating the unconditionality of some decomposition of the underlying space. For instance, Burkholder's square function estimate gives the unconditionality of martingale difference sequences in $L^p(S)$ (see Theorem 4.1.11).

Before we develop further this line of thought, we pause to take a closer look at the phenomenon of unconditionality in the present subsection.

Definition 4.1.2. *An indexed family $(x_i)_{i\in I}$ of elements in a Banach space X is said to be* unconditionally summable *or, more briefly,* summable, *if there exists an $x \in X$ with the following property: for all $\varepsilon > 0$ there is a finite subset $F_\varepsilon \subseteq I$ such that if $F \subseteq I$ is a finite set containing F_ε, then*

$$\Big\| x - \sum_{i\in F} x_i \Big\| < \varepsilon.$$

In this situation we say that $(x_i)_{i\in I}$ is unconditionally summable to x *or* converges unconditionally to x *and write*

$$\sum_{i\in I} x_i = x.$$

By abuse of notation we will sometimes say that '$\sum_{i\in I} x_i$ is summable' instead of ' $(x_i)_{i\in I}$ is summable'.

It is clear that if $(x_i)_{i\in I}$ is summable to both x and x', then $x = x'$.

Example 4.1.3. If $(h_i)_{i\in I}$ is an orthonormal family in a Hilbert space H then $(c_i h_i)_{i\in I}$ is summable if and only if

$$\sup_F \sum_{i\in F} |c_i|^2 < \infty,$$

the supremum being taken over all finite subsets F of I. In the same vein, in $\ell^p(0,1)$ with $1 \leqslant p < \infty$ we may consider the family of unit functions $(u_t)_{t\in(0,1)}$, defined by $u_t(s) = \delta_{s,t}$ for $s, t \in (0,1)$. Then $(c_t u_t)_{t\in(0,1)}$ is summable if and only if

$$\sup_F \sum_{t\in F} |c_t|^p < \infty,$$

the supremum being taken over all finite subsets F of $(0,1)$.

It is of course not this type of simple example that we have in mind, but rather a class of examples associated with decompositions of Banach spaces to which we will turn shortly. It will be convenient to first develop the notion of summability a bit by proving its equivalence with some conditions that are easier to check and which, as a matter of fact, justify the term "unconditional".

We begin with a Cauchy criterion for summability:

Lemma 4.1.4. *The family $(x_i)_{i\in I}$ is summable if and only if for all $\varepsilon > 0$ there is a finite set $G_\varepsilon \subseteq I$ such that if $G \subseteq I$ is finite and $G_\varepsilon \cap G = \varnothing$, then*

$$\Big\|\sum_{i\in G} x_i\Big\| < \varepsilon.$$

Proof. 'Only if': Suppose that $(x_i)_{i\in I}$ is summable to x and let $\varepsilon > 0$ be given. Let F_ε be as in the definition of summability. If G is finite and disjoint from F_ε, then $F_\varepsilon \subseteq F_\varepsilon \cup G$ implies that

$$\Big\|\sum_{i\in G} x_i\Big\| \leqslant \Big\|x - \sum_{i\in F_\varepsilon\cup G} x_i\Big\| + \Big\|x - \sum_{i\in F_\varepsilon} x_i\Big\| < 2\varepsilon.$$

'If': Considering singletons G, the assumption implies that for each $n \geqslant 1$ there are at most finitely many $i \in I$ (namely the elements of $G_{1/n}$) for which $\|x_i\| \geqslant 1/n$. It follows that the set $J := \{i \in I : \ x_i \neq 0\}$ is countable. Let $J = \{\eta(0), \eta(1), \dots\}$ be an enumeration. We claim that the partial sums of the series $\sum_{n\geqslant 0} x_{\eta(n)}$ form a Cauchy sequence in X. To this end fix $\varepsilon > 0$ and choose $N \geqslant 1$ so large that $G_\varepsilon \subseteq \{\eta(j) : 0 \leqslant j \leqslant N-1\}$. For $n \geqslant m \geqslant N$ the set $G_{m,n} := \{\eta(j) : m \leqslant j \leqslant n\}$ is disjoint from G_ε and therefore

$$\Big\| \sum_{j=m}^{n} x_{\eta(j)} \Big\| = \Big\| \sum_{i \in G_{m,n}}^{n} x_i \Big\| < \varepsilon.$$

This proves that the sequence of partial sums of $\sum_{j \geqslant 0} x_{\eta(j)}$ is Cauchy. Let $x := \sum_{j \geqslant 0} x_{\eta(j)}$. We claim that $(x_i)_{i \in I}$ is summable to x. If G_ε and N are as before and if $F \supseteq G_\varepsilon$ is finite, then for $n \geqslant N$ we find

$$\Big\| x - \sum_{i \in F} x_i \Big\| \leqslant \varepsilon + \Big\| \sum_{j=0}^{N-1} x_{\eta(j)} - \sum_{i \in F} x_i \Big\| < 2\varepsilon,$$

since $\{\phi(j) : 0 \leqslant j \leqslant N-1\} \setminus F$ is disjoint from G_ε. □

From the above proof we extract the following further information:

Lemma 4.1.5. *If $(x_i)_{i \in I}$ is summable to x, then the set $J := \{i \in I : x_i \neq 0\}$ is countable, and for every enumeration $J = \{\eta(0), \eta(1), \dots\}$ we have*

$$\sum_{j \geqslant 0} x_{\eta(j)} = x.$$

Now we are ready for the main characterisation of summability. The property alluded to in the fifth condition below is usually referred to as the *unconditionality* of the series.

Proposition 4.1.6 (Summability and unconditionality). *Let I be a countably infinite set. For a family $(x_i)_{i \in I}$ in a Banach space X, the following assertions are equivalent:*

(1) *$\sum_{i \in I} x_i$ is summable;*
(2) *$\sum_{i \in I} c_i x_i$ is summable for all choices of scalars $|c_i| \leqslant 1$;*
(3) *for some enumeration $\eta : \mathbb{N} \to I$, $\sum_{n \in \mathbb{N}} a_{\eta(n)} x_{\eta(n)}$ converges for all choices $a_i \in \{0, 1\}$;*
(4) *for some enumeration $\eta : \mathbb{N} \to I$, $\sum_{n \in \mathbb{N}} \epsilon_{\eta(n)} x_{\eta(n)}$ converges for all choices of signs $\epsilon_i \in \{-1, 1\}$;*
(5) *for every enumeration $\eta : \mathbb{N} \to I$, $\sum_{n \in \mathbb{N}} x_{\eta(n)}$ converges.*

Proof. The implication (1)⇒(5) has already been shown. We will prove the chain of implications (5)⇒(3)⇒(1)⇒(2)⇒(4)⇒(3).

(5)⇒(3): If (3) fails, for any enumeration $\eta : \mathbb{N} \to I$ we can find $\varepsilon > 0$, a sequence $(a_i)_{i \in I} \subseteq \{0, 1\}$, and a strictly increasing sequence of integers $(N_j)_{j \geqslant 1}$ such that $\| \sum_{k=N_j}^{N_{j+1}-1} a_{\eta(k)} x_{\eta(k)} \| \geqslant \varepsilon$. Let

$$A_j = \{N_j, \dots, N_{j+1} - 1\} \text{ and } B_j = \{n \in A_j : a_{\eta(n)} = 1\}.$$

Let π be a permutation of $\mathbb{N}$ such that

$$\pi(A_j) = A_j, \quad \pi(B_j) = \{N_j, \dots, N_j + b_j - 1\}$$

for all $j \geqslant 0$, where b_j is the number of elements of B_j. Then

$$\Big\| \sum_{k=N_j}^{N_j+b_j-1} x_{\eta(\pi(k))} \Big\| = \Big\| \sum_{k=N_j}^{N_{j+1}-1} a_{\eta(\pi(k))} x_{\eta(\pi(k))} \Big\| = \Big\| \sum_{k=N_j}^{N_{j+1-1}} a_{\eta(k)} x_{\eta(k)} \Big\| \geqslant \varepsilon,$$

and therefore $\sum_{k\in\mathbb{N}} x_{\eta(\pi(k))}$ cannot converge.

(3)⇒(1): If (1) fails, by Lemma 4.1.4 there exists $\varepsilon > 0$ as well as a sequence $N_1 < N_2 < \dots$ and finite sets $F_n \subseteq \{N_n + 1, \dots, N_{n+1}\}$ such that $\|\sum_{k\in F_n} x_{\eta(k)}\| \geqslant \varepsilon$. Let $a_{\eta(k)} := 1$ if $k \in \bigcup_{n\geqslant 1} F_n$ and $a_{\eta(k)} := 0$ otherwise. Then

$$\Big\| \sum_{k=N_n+1}^{N_{n+1}} a_{\eta(k)} x_{\eta(k)} \Big\| = \Big\| \sum_{k\in F_n} x_{\eta(k)} \Big\| \geqslant \varepsilon$$

and $\sum_{k\geqslant 1} a_{\eta(k)} x_{\eta(k)}$ cannot converge.

(1)⇒(2): By considering real and imaginary parts separately, and then positive and negative parts, we may assume that $c_i \in [0, 1]$ for all $i \in I$. Fix $\varepsilon > 0$. Using Lemma 4.1.4, choose a finite set $F_\varepsilon \subseteq I$ such that $\|\sum_{i\in F} x_i\| < \varepsilon$ for all finite sets $F \subseteq I$ disjoint with F_ε. Let $G \subseteq I$ be a finite set disjoint from F_ε, say $G = \{i_1, \dots, i_n\}$. Writing $(c_{i_1}, \dots c_{i_n})$ as a convex combination of the 2^n sequences $\alpha = (\alpha_{i_1}, \dots, \alpha_{i_n}) \in \{0, 1\}^n$ (using Lemma 3.2.11 and a scaling argument), we find

$$\Big\| \sum_{m=1}^{n} c_{i_m} x_{i_m} \Big\| \leqslant \sup_{\alpha\in\{0,1\}^n} \Big\| \sum_{m=1}^{n} \alpha_{i_m} x_{i_m} \Big\| = \sup_{G'\subseteq G} \Big\| \sum_{i\in G'} x_i \Big\| < \varepsilon.$$

(2)⇒(4): This is trivial.

(4)⇒(3): This is immediate by observing that any sequence in $\{0, 1\}$ is the sum of two sequences in $\{-1, 1\}$. □

Example 4.1.7 (Unconditionality of Rademacher sums). Let $(\varepsilon_n)_{n\geqslant 1}$ be a Rademacher sequence defined on a probability space Ω. Suppose that for a given sequence $(x_n)_{n\geqslant 1}$ in X the sum

$$\sum_{n\geqslant 1} \varepsilon_n x_n$$

converges in $L^2(\Omega; X)$, or equivalently (by the Kahane–Khintchine inequality, Theorem 3.2.23) in $L^p(\Omega; X)$ for some/all $p \in [1, \infty)$. We claim that the convergence in each of these spaces is unconditional. Indeed, for any sequence signs $\epsilon_n \in \{-1, 1\}$, the sequence $(\epsilon_n \varepsilon_n)_{n\geqslant 1}$ is again a Rademacher sequence. Since the convergence of a Rademacher sum only depends on the joint distribution of its terms, it follows that $\sum_{n\geqslant 1} \epsilon_n \varepsilon_n x_n$ converges in $L^2(\Omega; X)$, respectively in $L^p(\Omega; X)$. The claim now follows from Proposition 4.1.6.

Unconditional sums in Banach spaces typically arise in connection with coordinate expansions in terms of a given basis or expansions in terms of more general decompositions of the Banach space. In such situations one would like to know whether or not the sums involved are unconditional.

Definition 4.1.8 (Unconditional decompositions). *A* pre-decomposition *of a Banach space X is a family $(D_i)_{i\in I}$ of bounded projections on X satisfying*

$$D_i D_j = 0 \quad \textit{whenever} \quad i \neq j.$$

A pre-decomposition $(D_i)_{i\in I}$ is called an unconditional decomposition *if*

$$\sum_{i\in I} D_i x = x$$

unconditionally for all $x \in X$.

The definition might suggest an intermediate notion of 'unconditional pre-decomposition', where the summability $\sum_{i\in I} D_i x = x$ is only assumed on the closed linear span of the ranges of the projections D_i; we will not need this notion in the present context, but introduce a variant of it later in Definition 4.1.14.

If each D_i of an unconditional decomposition has rank one, i.e., $D_i(x) = \langle x, x_i^*\rangle x_i$ for some $x_i \in X$ and $x_i^* \in X^*$, the family $(x_i)_{i\in I}$ (unique up to scalar factors) is called an *unconditional basis*. It is usually normalised so that $\|x_i\| = 1$ for all $i \in I$.

Example 4.1.9. Examples 4.1.3 and 4.1.7 provide unconditional bases for their closed linear spans. In particular, the coordinate projections associated with orthonormal bases in Hilbert spaces and the standard unit bases of c_0 and ℓ^p, $1 \leqslant p < \infty$, are unconditional. The Haar basis is an unconditional basis of $L^p(0,1)$ for $1 < p < \infty$ (see Corollaries 4.5.8 and 4.5.16).

Before turning to some non-trivial examples it is convenient to state a simple criterion for a pre-decomposition to be unconditional.

Proposition 4.1.10. *A pre-decomposition $(D_i)_{i\in I}$ in a Banach space X is an unconditional decomposition if and only if the following two conditions are satisfied:*

(i) *the sums $\sum_{i\in F} D_i x$, where $x \in X$ and $F \subseteq I$ is finite, are dense in X;*
(ii) *there exists a constant $C \geqslant 0$ such that*

$$\Big\|\sum_{i\in F} \epsilon_i D_i x\Big\| \leqslant C\Big\|\sum_{i\in F} D_i x\Big\|$$

for all families of signs $(\epsilon_i)_{i\in I}$, all $x \in X$, and all finite $F \subseteq I$.

In this situation we have the estimate

$$\Big\|\sum_{i\in F} a_i D_i x\Big\| \leqslant C\|a\|_{\ell^\infty(I)}\|x\|$$

for all $a \in \ell^\infty(I)$, all $x \in X$, and all finite $F \subseteq I$. Moreover, $\sum_{i\in I} a_i D_i x$ is summable for all $a \in \ell^\infty(I)$ and $x \in X$, and the operator

$$T : x \mapsto \sum_{i\in I} a_i D_i x$$

is bounded with norm $\|T\| \leqslant C\|a\|_{\ell^\infty(I)}$.

The least admissible constants C in (ii) and the subsequent estimate are equal. This common constant is called the *unconditionality constant* of the decomposition.

The proof will also show that a pre-decomposition is unconditional if and only if (ii) holds for all elements in the closed linear span of the ranges of the projections D_i.

Proof. We begin with the 'only if' part. Assertion (i) is immediate from the definitions. For any $a \in \ell^\infty(I)$ and $x \in X$, the series $\sum_{i\in I} a_i D_i x$ is summable by Proposition 4.1.6, and therefore for each $a \in \ell^\infty$ the linear operator $T_a : x \mapsto \sum_{i\in I} a_i D_i x$ is well defined. We claim that T_a is closed, and hence bounded by the closed graph theorem. Suppose that $x_n \to x$ in X and $T_a x_n \to y$ in X. By the boundedness of D_i,

$$D_i y = \lim_{n\to\infty} D_i T_a x_n = \lim_{n\to\infty} a_i D_i x_n = a_i D_i x = D_i T_a x.$$

Summing over i gives $y = T_a x$ and the closedness of T_a follows.

Next, for fixed $x \in X$ we consider the linear operator $S_x : \ell^\infty \to X$ given by $S_x a := T_a x$, and claim that this operator is closed, hence bounded, as well. If $a^{(n)} \to a$ in ℓ^∞ and $S_x a^{(n)} \to y$ in X, then

$$D_i y = \lim_{n\to\infty} D_i S_x a^{(n)} = \lim_{n\to\infty} a_i^{(n)} D_i x = a_i D_i x = D_i S_x a.$$

Summing over i gives $y = S_x a$ and the closedness of S_x follows.

We have shown, for each fixed $x \in X$, that $T_a x$ remains bounded when a ranges over the unit ball of $\ell^\infty(I)$. Hence the uniform boundedness principle implies that the operators T_a are uniformly bounded for a in this range, which gives the required estimate

$$\Big\|\sum_{i\in I} a_i D_i x\Big\| \leqslant C\|a\|_{\ell^\infty}\|x\|$$

by scaling with $\|a\|_{\ell^\infty}$. Applying this to the vector $\sum_{i\in F} D_i x$ we find the desired estimate in (ii). The remaining assertions follow from the above.

Turning to the 'if' part, we first claim that for all finite $F \subseteq I$ and $x \in X$,

$$\Big\|\sum_{i\in F} D_i x\Big\| \leqslant C\|x\|. \tag{4.2}$$

Fix a finite $F \subseteq I$. By (i) it suffices to prove (4.2) for $x = \sum_{i\in G} D_i y$ with $y \in X$ and $G \subseteq I$ finite. Now let $\epsilon_i = 1$ if $i \in F \cap G$ and $\epsilon_i = -1$ if $i \in G \setminus F$. Then by (ii) and using that $D_i D_j = 0$ for $i \neq j$ we find that

$$\Big\|\sum_{i\in F} D_i x\Big\| = \Big\|\sum_{i\in F\cap G} D_i y\Big\| = \frac{1}{2}\Big\|\sum_{i\in G} \epsilon_i D_i y + \sum_{i\in G} D_i y\Big\| \leqslant C\Big\|\sum_{i\in G} D_i y\Big\| = C\|x\|$$

and hence (4.2) follows.

It remains to prove the finals assertions. Fix $x \in X$ and $\varepsilon > 0$, and use (i) to find $y \in X$ and a finite set $F_\varepsilon \subseteq I$ such that

$$\Big\|x - \sum_{i\in F_\varepsilon} D_i y\Big\| < \varepsilon.$$

Let $F \subseteq I$ be a finite set containing F_ε. Using that $D_i D_j = 0$ for $i \neq j$, together with (4.2) we find that

$$\begin{aligned}
\Big\|x - \sum_{i\in F} D_i x\Big\| &\leqslant \Big\|x - \sum_{i\in F_\varepsilon} D_i y\Big\| + \Big\|\sum_{i\in F_\varepsilon} D_i y - \sum_{i\in F} D_i x\Big\| \\
&= \Big\|x - \sum_{i\in F_\varepsilon} D_i y\Big\| + \Big\|\sum_{i\in F} D_i\Big(\sum_{j\in F_\varepsilon} D_j y - x\Big)\Big\| \\
&\leqslant \varepsilon + C\Big\|\Big(\sum_{j\in F_\varepsilon} D_j y - x\Big)\Big\| < (1+C)\varepsilon.
\end{aligned}$$

□

This proposition reduces the task of checking whether a given pre-decomposition is unconditional to checking a density condition and proving an estimate on finite sums.

Theorem 4.1.1 implies the unconditionality of L^p-martingale differences for $1 < p < \infty$:

Theorem 4.1.11 (Unconditionality of martingale differences, Burkholder). *Let $(S, \mathscr{A}, \mu)$ be a measure space with a σ-finite filtration $(\mathscr{F}_n)_{n\geqslant 0}$, and let $\mathscr{F}_\infty := \sigma(\mathscr{F}_n : n \geqslant 0)$. Set*

$$\begin{aligned}
D_0 &:= \mathbb{E}(\cdot|\mathscr{F}_0), \\
D_n &:= \mathbb{E}(\cdot|\mathscr{F}_n) - \mathbb{E}(\cdot|\mathscr{F}_{n-1}), \quad n \geqslant 1.
\end{aligned}$$

Then $(D_n)_{n\geqslant 0}$ is an unconditional decomposition of $L^p(S, \mathscr{F}_\infty)$, $1 < p < \infty$.

Proof. The linear span of the ranges of the projections D_n, $n \geqslant 0$, is dense (by the L^p-convergence $\mathbb{E}(f|\mathscr{F}_n) = \sum_{k=0}^n D_k f \to f$ for $f \in L^p(S, \mathscr{F}_\infty)$; see Theorem 3.3.2). Moreover, given a function $f \in L^p(S, \mathscr{F}_\infty)$, Burkholder's theorem applied first to the martingale $g_n = \sum_{k=0}^n a_k D_k f$ and then to the martingale $f_n = \sum_{k=0}^n D_k f$, gives

$$\Big\| \sum_{k=0}^n a_k D_k f \Big\|_p \eqsim_p \Big\| \Big(\sum_{k=0}^n |a_k D_k f|^2 \Big)^{1/2} \Big\|_p$$
$$\leqslant \|a\|_{\ell^\infty_N} \Big\| \Big(\sum_{k=0}^n |D_k f|^2 \Big)^{1/2} \Big\|_p \eqsim_p \|a\|_{\ell^\infty_N} \Big\| \sum_{k=0}^n D_k f \Big\|_p.$$

This argument also gives an upper bound on the unconditionality constant of this decomposition, namely $A_p^{-1} B_p$, where A_p and B_p are the constants in Burkholder's theorem 4.1.1. Now the result follows from Proposition 4.1.10. □

We continue with two further examples, a negative one and a positive one.

Example 4.1.12 (Trigonometric system in $L^p(\mathbb{T})$). Recall that the trigonometric functions $e_n : \mathbb{T} \to \mathbb{C}$ have been defined by $e_n(t) = e^{2\pi i n t}$. The coordinate projections associated with the trigonometric system $(e_n)_{n\in\mathbb{Z}}$ fail to be unconditional in $L^p(\mathbb{T})$ for $p \in [1, \infty] \setminus \{2\}$. More generally, let S be any finite measure space such that $L^2(S)$ is infinite-dimensional. Given $c, C > 0$, if $(f_n)_{n\in\mathbb{N}}$ is any orthonormal basis for $L^2(S)$ such that $c \leqslant |f_n| \leqslant C$ almost everywhere for all $N \in \mathbb{N}$, then $(f_n)_{n\in\mathbb{N}}$ fails to be unconditional in $L^p(S)$ for $p \in [1, \infty] \setminus \{2\}$.

To see this let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a Rademacher sequence on a probability space Ω. Unconditionality of the system in $L^p(S)$ with $p \in [1, \infty)$ would imply that for all $\omega \in \Omega$,

$$\Big\| \sum_{n=1}^N a_n f_n \Big\|_{L^p(S)} \eqsim \Big\| \sum_{n=1}^N \varepsilon_n(\omega) a_n f_n \Big\|_{L^p(S)}.$$

Taking $L^p(\Omega)$-norms and applying Khintchine's inequality (3.34), this gives

$$\Big\| \sum_{n=1}^N a_n f_n \Big\|_{L^p(S)} \eqsim_{p,c,C} \Big(\sum_{n=1}^N |a_n|^2 \Big)^{1/2} = \Big\| \sum_{n=1}^N a_n f_n \Big\|_{L^2(S)}.$$

This would lead to equality $L^p(S) = L^2(S)$ with equivalent norms. By considering simple functions and using the infinite dimensionality, this can be true only if $p = 2$. The case $p = \infty$ can be obtained by a duality argument (see Proposition 4.1.16 below).

It turns out that an appropriate blocking of the trigonometric system does form an unconditional decomposition, and a similar result holds for the real line. This is the content of the next example.

Example 4.1.13 (Littlewood–Paley decomposition). In Theorem 5.3.24 it will be shown that for UMD Banach spaces X and $1 < p < \infty$, the Littlewood–Paley decomposition is an unconditional decomposition of $L^p(\mathbb{R}; X)$.

To define the Littlewood–Paley decomposition in the scalar-valued case of $L^p(\mathbb{R})$, for $n \in \mathbb{Z}$ let

$$I_n := [-2^n, -2^{n-1}] \cup [2^{n-1}, 2^n]$$

and define X_n to be the closure in $L^p(\mathbb{R})$ of all functions in $L^2(\mathbb{R}) \cap L^p(\mathbb{R})$ whose Fourier–Plancherel transform is supported in I_n. The associated projections Δ_n onto X_n are given, for $f \in L^2(\mathbb{R}) \cap L^p(\mathbb{R})$, by

$$\mathscr{F}(\Delta_n f) = \mathbf{1}_{I_n} \mathscr{F}(f).$$

By the boundedness of the Hilbert transform on $L^p(\mathbb{R})$, the operators Δ_n extend to bounded projections in $L^p(\mathbb{R})$ onto X_n (see Lemma 5.3.9).

The classical Littlewood–Paley inequality asserts that there are constants $0 < a_p \leqslant b_p < \infty$ such that for all $f \in L^p(\mathbb{R})$ the square function estimate

$$a_p \Big\| \Big(\sum_{n\in\mathbb{Z}} |\Delta_n f|^2 \Big)^{1/2} \Big\|_p \leqslant \|f\|_p \leqslant b_p \Big\| \Big(\sum_{n\in\mathbb{Z}} |\Delta_n f|^2 \Big)^{1/2} \Big\|_p$$

holds; this result will follow from Theorem 5.3.24 and Khintchine's inequality (3.34). Clearly, the unconditionality of $(\Delta_n)_{n\in\mathbb{Z}}$ follows from the above estimate.

This example can be extended to $L^p(\mathbb{R}^d)$ by replacing the dyadic intervals by dyadic cubes (see Theorem 5.7.10), and by transference a similar result holds for $L^p(\mathbb{T}^n)$.

Our next aim is to connect the notion of unconditional decompositions with Rademacher sums. This necessitates some preliminary definitions and observations. We shall restrict our attention to countable index sets I for the ease of presentation. For proving properties of unconditional decompositions this limitation is by no means essential and the reader will have no difficulty in extending the relevant proofs to arbitrary index sets.

We start with the more general situation of a general pre-decomposition with countable index set I. Unless we impose unconditionality, the order of summation does matter and we have to make a choice with regard to the ordering of I. As a rule we shall formulate results for general decompositions for the index set $I = \{n \geqslant 1\}$ with the natural ordering; versions for $I = \mathbb{Z}$ hold as well and are left to the reader.

Definition 4.1.14. *A pre-decomposition $(D_n)_{n\geqslant 1}$ of X is called a* Schauder decomposition *if*

$$\lim_{n\to\infty} \sum_{k=1}^{n} D_k x = x$$

for all $x \in X$. A Schauder decomposition consisting of rank one projections is called a Schauder basis*.*

For decompositions indexed by $\mathbb{Z}$ the defining condition should be replaced by the requirement

$$\lim_{m,n\to\infty}\sum_{k=-m}^{n} D_k x = x$$

for all $x \in X$.

Every unconditional decomposition is a Schauder decomposition, but the converse does not hold:

Example 4.1.15 (Trigonometric system in $L^p(\mathbb{T})$ revisited). The trigonometric system $(e_n)_{n\in\mathbb{Z}}$ is a Schauder basis for $L^p(\mathbb{T})$, $1 < p < \infty$ which, by the result of Example 4.1.12, fails to be unconditional for $p \neq 2$. To see that $(e_n)_{n\in\mathbb{Z}}$ is indeed a Schauder basis we recall the fact, mentioned in Section 2.1.c, the periodic Hilbert transform is bounded on $L^p(\mathbb{T})$ (a proof of this fact is obtained by combining Corollary 5.2.11 and the fact that the scalar field has the UMD property). As in Lemma 5.3.9, from this it follows that the mappings

$$R_{k_0,k_1} : e_k \mapsto \begin{cases} e_k & \text{if } k_0 \leqslant k \leqslant k_1, \\ 0 & \text{otherwise,} \end{cases}$$

are bounded, uniformly in $k_0 \leqslant k_1$. Set $D_k := R_{k,k}$. The projections $\sum_{k=-m}^{n} D_k = R_{-m,n}$ are bounded, uniformly in $m, n \geqslant 0$. On the dense subspace of finite linear combinations of the form $f = \sum_{j=-k}^{k} c_k e_k$ we have $\lim_{m,n\to\infty}\sum_{k=-m}^{n} D_k f = f$, and by uniform boundedness this convergence extends to arbitrary $f \in L^p(\mathbb{T})$.

In $L^1(\mathbb{T})$ the trigonometric system does not even form a Schauder basis. For otherwise the partial sum projections $\sum_{k=-m}^{n} D_k$ would be uniformly bounded, and hence so would be the projections $\sum_{k=0}^{n+m} D_k$, by the same routine arguments used above. By passing to the strong limit, the Riesz projection $\sum_{k\geqslant 0} D_k$ would be bounded on $L^1(\mathbb{T})$ – which is not the case.

If $(D_n)_{n\geqslant 1}$ is a Schauder decomposition of X, the *partial sum projections*

$$P_n := \sum_{k=1}^{n} D_k$$

are uniformly bounded by the uniform boundedness principle.

The next proposition connects the notions of unconditionality with randomisation by Rademacher variables.

Proposition 4.1.16. *Let $(D_i)_{i\in I}$ be a pre-decomposition of a Banach space X with the following property:*

$$\text{the sums } \textstyle\sum_{i\in F} D_i x, \text{ with } x \in X \text{ and } F \subseteq I \text{ finite, are dense in } X. \tag{4.3}$$

Let $(\varepsilon_i)_{i\in I}$ be a family of independent Rademacher variables on a probability space Ω. For any exponent $1 \leqslant p < \infty$, the following assertions are equivalent:

(1) $(D_i)_{i\in I}$ *is unconditional;*
(2) *there exist constants* $0 < C^{\pm} < \infty$ *such that for all* $x \in X$ *we have*

$$\frac{1}{C^-}\|x\| \leqslant \sup_{F\subseteq I \text{ finite}} \Big\| \sum_{i\in F} \varepsilon_i D_i x\Big\|_{L^p(\Omega;X)} \leqslant C^+\|x\|. \tag{4.4}$$

When these equivalent conditions are satisfied, $\sum_{i\in I} \varepsilon_i D_i x$ *is summable in* $L^p(\Omega; X)$ *for all* $x \in X$ *and*

$$\frac{1}{C^-}\|x\| \leqslant \Big\| \sum_{i\in I} \varepsilon_i D_i x\Big\|_{L^p(\Omega;X)} \leqslant C^+\|x\|. \tag{4.5}$$

If $I = \{n \geqslant 1\}$ *and* $(D_n)_{n\geqslant 1}$ *is a Schauder decomposition of* X*, then* (4.3) *holds, and the conditions* (1) *and* (2) *are equivalent to:*

(3) *there exists a constant* $C \geqslant 0$ *such that for all* $x \in X$*,* $x^* \in X^*$*, and* $N \geqslant 1$*,*

$$\begin{aligned}\Big\| \sum_{n=1}^N \varepsilon_n D_n x\Big\|_{L^p(\Omega;X)} &\leqslant C\Big\| \sum_{n=1}^N D_n x\Big\|,\\ \Big\| \sum_{n=1}^N \varepsilon_n D_n^* x^*\Big\|_{L^{p'}(\Omega;X^*)} &\leqslant C\Big\| \sum_{n=1}^N D_n^* x^*\Big\|.\end{aligned} \tag{4.6}$$

Proof. (1)⇒(2): Suppose that $(D_i)_{i\in I}$ is unconditional, with unconditionality constant C_D. We will prove that (4.4) holds with constants $C^- = C^+ = C_D$.

By Proposition 4.1.10, for all finite $F \subseteq I$ and all $\omega \in \Omega$ we have

$$\frac{1}{C_D}\Big\| \sum_{i\in F} \varepsilon_i(\omega) D_i x\Big\| \leqslant \Big\| \sum_{i\in F} D_i x\Big\| \leqslant C_D\Big\| \sum_{i\in F} \varepsilon_i(\omega) D_i x\Big\|.$$

Indeed, the first inequality is immediate from the proposition and the second follows from the first by applying it to the vector $\sum_{i\in F} \varepsilon_i(\omega) D_i x$. Raising both sides to the power p and integrating gives

$$\frac{1}{C_D}\Big\| \sum_{i\in F} \varepsilon_i D_i x\Big\|_{L^p(\Omega;X)} \leqslant \Big\| \sum_{i\in F} D_i x\Big\| \leqslant C_D\Big\| \sum_{i\in F} \varepsilon_i D_i x\Big\|_{L^p(\Omega;X)}. \tag{4.7}$$

It is evident from the definition of summability that

$$\|x\| \leqslant \sup_{F\subseteq I \text{ finite}} \Big\| \sum_{i\in F} D_i x\Big\|,$$

and hence the first estimate in (4.4) follows from (4.7). To prove the second estimate in (4.4) it suffices to note that, by Proposition 4.1.10, for all finite $F \subseteq I$ we have

$$\frac{1}{C_D}\Big\| \sum_{i\in F} \varepsilon_i(\omega) D_i x\Big\| \leqslant \|x\|,$$

and to take L^p-norms as before.

(2)⇒(1): Fix $x \in X$ and let $(\epsilon_i)_{i\in I}$ be a family of signs. Let $G \subseteq I$ be finite. Applying the first inequality of (4.4) to $\sum_{i\in G} \epsilon_i D_i x$ and the second to x and using that $(\epsilon_i\varepsilon_i)_{i\in I}$ is a family of independent Rademacher variables, we obtain

$$\begin{aligned}\frac{1}{C^-}\Big\|\sum_{i\in G}\epsilon_i D_i x\Big\| &\leqslant \sup_F \Big\|\sum_{i\in F\cap G}\epsilon_i\varepsilon_i D_i x\Big\|_{L^p(\Omega;X)} \\ &= \sup_F \Big\|\sum_{i\in F\cap G}\varepsilon_i D_i x\Big\|_{L^p(\Omega;X)} \\ &\leqslant \sup_F \Big\|\sum_{i\in F}\varepsilon_i D_i x\Big\|_{L^p(\Omega;X)} \leqslant C^+\|x\|,\end{aligned}$$

the suprema being taken over all finite subsets $F \subseteq I$. Applying this to $x = \sum_{i\in G} D_i y$ gives condition (ii) of Proposition 4.1.10, with constant C^-C^+. Since condition (i) of Proposition 4.1.10 is precisely the assumption (4.3), the result follows from the proposition. This gives the unconditionality of $(D_i)_{i\in I}$, with constant C^-C^+.

The Cauchy criterion of Lemma 4.1.4, (4.7), and the summability of $\sum_{i\in I} D_i x = x$ further imply that $\sum_{i\in I}\varepsilon_i D_i x$ is summable in $L^p(\Omega;X)$. The inequalities (4.5) follow from (4.7).

Suppose now that $I = \{n \geqslant 1\}$ and that $(D_n)_{n\geqslant 1}$ is a Schauder decomposition. Then (4.3) is clearly satisfied.

(1)⇒(3): Suppose first that $(D_n)_{n\geqslant 1}$ is unconditional, with unconditionality constant C_D. By the estimate of Proposition 4.1.10, for all $\omega \in \Omega$ we have

$$\Big\|\sum_{n=1}^N \varepsilon_n(\omega) D_n x\Big\| \leqslant C_D \Big\|\sum_{n=1}^N D_n x\Big\|.$$

Raising both sides to the power p and taking expectations, this gives the first inequality of (4.6), with constant C_D. By the observation on duality in the text preceding the proposition, the same argument gives the second inequality of (4.6), again with constant C_D.

(3)⇒(1): We apply (4.6) to $P_N x$ to obtain, for all choices of $|\epsilon_n| = 1$ in $\mathbb{K}$ and all $x^* \in X^*$,

$$\begin{aligned}\Big|\Big\langle \sum_{n=1}^N \epsilon_n D_n x, x^*\Big\rangle\Big| &= \Big|\mathbb{E}\Big\langle \sum_{n=1}^N \epsilon_n\varepsilon_n D_n x, \sum_{m=1}^N \varepsilon_m D_m^* x^*\Big\rangle\Big| \\ &\leqslant C^2 \Big\|\sum_{n=1}^N D_n x\Big\| \Big\|\sum_{m=1}^N D_m^* x^*\Big\| \leqslant C^2 C_P \Big\|\sum_{n=1}^N D_n x\Big\|\, \|x^*\|,\end{aligned}$$

where $C_P = \sup_{k\geqslant 1}\|P_k\|$. Here we used that $(\epsilon_n\varepsilon_n)_{n\geqslant 1}$ is a Rademacher sequence. Taking the supremum over all $x^* \in X^*$ of norm one, by Proposition 4.1.10 this gives the unconditionality, with constant $C_D \leqslant C^2 C_P$. □

4.2 The UMD property

In the previous section we have discussed Burkholder's square function estimate for L^p-martingale difference sequences and observed its unconditionality. This raises the natural question whether these results generalise to the vector-valued situation. Here we take up this question in the following format: Under what conditions is it true that vector-valued L^p-martingale difference sequences are unconditional? This question leads to the introduction of what is arguably the most important class of Banach spaces for Analysis, the class of UMD spaces, defined as consisting of those spaces for which the question just posed has an affirmative answer.

The unconditionality of vector-valued martingale differences in this class of Banach spaces has surprisingly far-reaching consequences in various domains of analysis. Moreover, it turns out that this property, and thereby its consequences, can also be verified for many concrete Banach spaces of interest.

4.2.a Definition and basic properties

Definition 4.2.1 (UMD property). *A Banach space X is said to have the* property of unconditional martingale differences (UMD property) *if for all $p \in (1, \infty)$ there exists a finite constant $\beta \geqslant 0$ (depending on p and X) such that the following holds. Whenever $(S, \mathscr{A}, \mu)$ is a σ-finite measure space, $(\mathscr{F}_n)_{n=0}^N$ is a σ-finite filtration, and $(f_n)_{n=0}^N$ is a finite martingale in $L^p(S; X)$, then for all scalars $|\epsilon_n| = 1$, $n = 1, \dots, N$, we have*

$$\Big\| \sum_{n=1}^N \epsilon_n df_n \Big\|_{L^p(S;X)} \leqslant \beta \Big\| \sum_{n=1}^N df_n \Big\|_{L^p(S;X)}. \tag{4.8}$$

If this condition holds, then X is said to be a UMD space*.*

The dyadic UMD property *is defined by requiring the same condition for Paley–Walsh martingales only.*

Recall that a *Paley–Walsh martingale* is a martingale with respect to a Paley–Walsh filtration, i.e., a filtration in which each σ-algebra $\mathscr{F}_n$ consists of 2^n sets, each of which is a union of atoms of size 2^{-n}. In Proposition 3.1.10 it was shown that every Paley–Walsh martingale is of the form $f_0 =$ constant and

$$df_n = \phi_n(r_1, \dots, r_{n-1}) r_n \tag{4.9}$$

for a suitable function $\phi_n : \{-1, 1\}^{n-1} \to X$ and a real Rademacher sequence $(r_n)_{n \geqslant 1}$.

Example 4.2.2. Theorem 4.1.1 expresses the fact that the scalar field $\mathbb{K}$ is a UMD space.

The least admissible constant β in (4.8) will be denoted by $\beta_{p,X}$ and is called the *UMD constants* of X; the least constant for the dyadic UMD property is denoted by $\beta^{\Delta}_{p,X}$.

Condition (4.8) asserts that the martingale differences df_n are unconditional, with unconditionality constant β. As an immediate consequence we therefore obtain the following corollary to Proposition 4.1.16:

Proposition 4.2.3 (Randomised UMD property). *A Banach space X has the UMD property (dyadic UMD property) if and only if for all $p \in (1,\infty)$ there exist constants $0 < \beta^{\pm} < \infty$ (depending on p and X) such that whenever $(f_n)_{n=0}^N$ is an L^p-martingale (Paley–Walsh martingale) on a σ-finite measure space (probability space) $(S, \mathscr{A}, \mu)$, we have*

$$\frac{1}{\beta^-}\Big\| \sum_{n=1}^N df_n \Big\|_{L^p(S;X)} \leqslant \Big\| \sum_{n=1}^N \varepsilon_n df_n \Big\|_{L^p(S\times\Omega;X)} \leqslant \beta^+ \, \Big\| \sum_{n=1}^N df_n \Big\|_{L^p(S;X)}. \tag{4.10}$$

Here, $(\varepsilon_n)_{n=1}^N$ is a Rademacher sequence on a probability space Ω.

The least admissible constants β^- and β^+ in (4.10) are denoted by $\beta^-_{p,X}$ and $\beta^+_{p,X}$. Tracing constants in the proof of Proposition 4.1.16 reveals that

$$\max\{\beta^-_{p,X}, \beta^+_{p,X}\} \leqslant \beta_{p,X} \leqslant \beta^-_{p,X}\beta^+_{p,X}.$$

Remark 4.2.4. One may wonder whether just one of the two inequalities in (4.10) is enough to recover the full UMD property. For the left-hand inequality, this turns out not to be the case. It is an open problem whether the right-hand side in (4.10) implies the UMD property. We refer the reader to the Notes at the end of this chapter for a discussion of this issue; see also Problem O.9.

The flexibility of the notion of UMD is already reflected by the following two fundamental theorems:

Theorem 4.2.5. *For a Banach space X, the following assertions are equivalent:*

(1) *X has the UMD property;*
(2) *X has the dyadic UMD property.*

Moreover, for such X we have $\beta_{p,X} = \beta^{\Delta}_{p,X}$.

Since real Rademacher sequences can be modelled on the unit interval $[0,1)$ (see Example 3.1.9) and keeping (4.9) in mind, an easy corollary of this theorem is that the measure space $[0,1)$ is universal for the UMD property. (By this we mean that, in the definition of the UMD property, it suffices to consider martingales defined on the probability space $[0,1)$ with the dyadic filtration.)

Corollary 4.2.6. *If a Banach space X satisfies the defining UMD estimate for every dyadic X-valued martingale on $[0,1)$, then it satisfies the UMD estimate for X-valued martingales on arbitrary σ-finite measure spaces with a σ-finite filtration, with the same UMD constants.*

For the randomised UMD constants $\beta^{\pm}_{p,X}$, a corresponding reduction to Paley–Walsh martingales seems impossible (see the Notes for a discussion).

Theorem 4.2.7. *A Banach space X is UMD if and only if it satisfies the defining condition for* some $p \in (1,\infty)$, *and in fact for all $p, q \in (1,\infty)$ we have*

$$\beta_{p,X} \leqslant 100\Big(\frac{p}{q} + \frac{p'}{q'}\Big)\beta_{q,X}.$$

Note in particular that

$$\frac{1}{200}\beta_{2,X} \leqslant \beta_{p,X} \leqslant 50(p+p')\beta_{2,X},$$

so that the constant $\beta_{2,X}$ is always 'quasi-minimal' among the constants $\beta_{p,X}$, and $\beta_{p,X}$ increases at the rate of $O(p+p')$ as $p \to 1$ or $p \to \infty$.

As we shall see, these are rather quick consequences of the general theory of martingale transforms and inequalities developed in Chapter 3. However, it is worthwhile observing some easy equivalences first:

Lemma 4.2.8. *The defining condition (4.8), of either UMD or dyadic UMD, may be equivalently replaced by either*

$$\Big\|\sum_{n=1}^N \epsilon_n df_n\Big\|_{L^p(S;X)} \leqslant \beta\|f_N\|_{L^p(S;X)} \tag{4.11}$$

or

$$\Big\|\epsilon_0 f_0 + \sum_{n=1}^N \epsilon_n df_n\Big\|_{L^p(S;X)} \leqslant \beta\|f_N\|_{L^p(S;X)} \tag{4.12}$$

where the objects range over the same families as in Definition 4.2.1. The best constants in (4.11) and (4.12) coincide with the UMD constants of X.

The same result holds for each of the individual inequalities in Proposition 4.2.3; thus in the definition of the constants $\beta^+_{p,X}$ and $\beta^-_{p,X}$ one could also use martingales which do not start at zero and add the terms with f_0 on each side of the defining inequalities without changing the constants.

Proof. Denote the best constants in (4.11) and (4.12) by $\beta^{(1)}_{p,X}$ and $\beta^{(2)}_{p,X}$, respectively.

Averaging (4.12) with signs $\epsilon_0 = \pm 1$ gives (4.11) with constant $\beta^{(1)}_{p,X} \leqslant \beta^{(2)}_{p,X}$. Applying (4.11) to the martingale $f_n - f_0$ we see that it reduces to

(4.8); this also gives $\beta_{p,X} \leqslant \beta_{p,X}^{(1)}$. So the main point is to prove that (4.8) implies (4.12), with $\beta_{p,X}^{(2)} \leqslant \beta_{p,X}$. This uses the flexibility that the underlying measure space and the length N are also allowed to vary.

Indeed, let r be a real Rademacher variable on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and let $g_n := f_{n-1} \otimes r$ for $n = 1, \ldots, N+1$ and $g_0 := 0$. This is a martingale with respect to the filtration given by $\mathscr{G}_n := \mathscr{F}_{n-1} \times \sigma(r)$ for $n \geqslant 1$ and $\mathscr{G}_0 := \mathscr{F}_0 \times \{\Omega, \varnothing\}$; in particular, $\mathbb{E}(g_1|\mathscr{G}_0) = \mathbb{E}(f_0|\mathscr{F}_0) \otimes \mathbb{E}r = 0$, since $\mathbb{E}r = 0$. Moreover, $dg_n = df_{n-1} \otimes r$ for $n \geqslant 2$ and $dg_1 = f_0 \otimes r$. If f_n is a dyadic martingale, so is g_n, perhaps after relabelling the Rademacher sequence $\{r\} \cup \{r_n\}_{n=1}^N$ if one so insists.

Writing (4.8) with $(g_n)_{n=0}^{N+1}$ in place of $(f_n)_{n=0}^N$, and re-indexing the ϵ_n's, we arrive at an estimate like (4.12), except that there is a common factor r inside both norms. However, since $|r| \equiv 1$, this factor is invisible to the norms and may be cancelled, and thus we get (4.12) exactly as stated, with the inequality on the best constants as claimed. $\square$

It turns out that the equivalent condition (4.12) is perhaps most amenable to the applications of the general theory. In fact, this condition plainly states the requirement of uniform boundedness of all finite martingale transform operators

$$T_\epsilon f := \epsilon_0 f_0 + \sum_{n=1}^N \epsilon_n df_n$$

for sequences of signs $\epsilon = (\epsilon_n)_{n=0}^N$ in $\mathbb{K}$.

We now turn to the proofs of Theorems 4.2.5 and 4.2.7.

Proof of Theorem 4.2.5. It suffices to show that (4.12) for Paley–Walsh martingales implies (4.12) for all martingales, with the same constant.

Let $(f_n)_{n=0}^N$ be an $L^p(S;X)$-martingale. For a given $\varepsilon > 0$, we apply Theorem 3.6.1 to produce another $L^p(S;X)$-martingale $(g_j)_{j=0}^N$ and an increasing function $\phi : \{0, 1, \ldots, n\} \to \{0, 1, \ldots, \phi(n) = N\}$ such that:

(i) $\|f_j - g_{\phi(j)}\|_p < \varepsilon$ for all $j = 0, 1, \ldots, n$,
(ii) $(g_n)_{n=0}^N$ is supported on a set $E \in \mathscr{A}$ of finite measure,
(iii) $(g_n)_{n=0}^N$ is a Paley–Walsh martingale in $L^p(E, \mu_E; X)$, with $\mu_E = \frac{1}{\mu(E)} \mu|_E$ the restriction of μ on E normalised to a probability measure.

The proof is now a direct computation:

$$\begin{aligned}
&\Big\| \epsilon_0 f_0 + \sum_{j=1}^n \epsilon_j (f_j - f_{j-1}) \Big\|_{L^p(S;X)} \\
&\qquad \leqslant \Big\| \epsilon_0 g_{\phi(0)} + \sum_{j=1}^n \epsilon_j (g_{\phi(j)} - g_{\phi(j-1)}) \Big\|_{L^p(S;X)} + 2(n+1)\varepsilon
\end{aligned}$$

$$
\begin{aligned}
&= \mu(E)^{1/p}\Big\|\epsilon_0 g_0 + \sum_{j=0}^{n} \epsilon_j \sum_{k=\phi(j-1)+1}^{\phi(j)} dg_k\Big\|_{L^p(E,\mu_E;X)} + 2(n+1)\varepsilon \\
&\qquad \text{where} \quad \phi(-1) := 0 \\
&= \mu(E)^{1/p}\Big\|\epsilon_0 g_0 + \sum_{i=1}^{N} \epsilon^{(i)} dg_i\Big\|_{L^p(E,\mu_E;X)} + 2(n+1)\varepsilon \\
&\qquad \text{where} \quad \epsilon^{(i)} := \epsilon_j \quad \text{for} \quad \phi(j-1) < i \leqslant \phi(j) \\
&\leqslant \mu(E)^{1/p}\beta^{\Delta}_{p,X}\|g_N\|_{L^p(E,\mu_E;X)} + 2(n+1)\varepsilon \\
&= \beta^{\Delta}_{p,X}\|g_{\phi(n)}\|_{L^p(S;X)} + 2(n+1)\varepsilon \\
&\leqslant \beta^{\Delta}_{p,X}\|f_n\|_{L^p(S;X)} + (\beta^{\Delta}_{p,X} + 2(n+1))\varepsilon.
\end{aligned}
$$

Since $\varepsilon > 0$ was arbitrary, Lemma 4.2.8 implies that X is UMD with constant $\beta_{p,X} \leqslant \beta^{\Delta}_{p,X}$. □

Proof of Theorem 4.2.7. Again, we use the formulation of UMD as in (4.12), which requires the uniform boundedness of the finite martingale transform

$$T_\epsilon f = \epsilon_0 f_0 + \sum_{n=1}^{N} \epsilon_n df_n.$$

Restricting ourselves to the case of Paley–Walsh martingales, Theorem 3.5.4 guarantees that these satisfy

$$\|T_\epsilon\|_{\mathscr{L}(L^p(S;X))} \leqslant 100\Big(\frac{p}{q} + \frac{p'}{q'}\Big)\|T_\epsilon\|_{\mathscr{L}(L^q(\mu;X))},$$

and taking the supremum over all such transforms, we arrive at

$$\beta^{\Delta}_{p,X} \leqslant 100\Big(\frac{p}{q} + \frac{p'}{q'}\Big)\beta^{\Delta}_{q,X}.$$

Recalling from Theorem 4.2.5 that $\beta_{p,X} = \beta^{\Delta}_{p,X}$, this gives the assertion of Theorem 4.2.7 as stated. □

Remark 4.2.9. Theorem 4.2.5 was only needed in the previous proof to get the sharp quantitative assertion of Theorem 4.2.7: applying a similar argument on the level of arbitrary (rather than Paley–Walsh) martingales, Theorem 3.5.4 would have given the qualitatively equivalent but quantitatively weaker estimate $\beta_{p,X} \leqslant 1400pp'\beta_{q,X}$ instead.

The role of the scalar field

Recall that when X is a complex Banach space, $X_{\mathbb{R}}$ denotes the real Banach space obtained by restricting the scalar multiplication of X to the reals. Below we write $\beta^{\mathbb{R}}_{p,X}$ for the constant one obtains in the definition of UMD if one restricts to the real scalars $(\epsilon_n)_{n=1}^N$ in $\{-1,1\}$ rather than taking them from the possibly larger set $\{z \in \mathbb{K} : |z| = 1\}$.

Proposition 4.2.10. *Let X be a complex Banach space and let $1 < p < \infty$. The following assertions are equivalent:*

(1) *X is a UMD space;*
(2) *$X_{\mathbb{R}}$ is a UMD space.*

In this situation, for all $p \in (1, \infty)$ we have $\beta^{\mathbb{R}}_{p,X} = \beta_{p,X_{\mathbb{R}}}$ and

$$\beta_{p,X_{\mathbb{R}}} \leqslant \beta_{p,X} \leqslant \frac{\pi}{2}\beta_{p,X_{\mathbb{R}}}.$$

Proof. The identity $\beta^{\mathbb{R}}_{p,X} = \beta_{p,X_{\mathbb{R}}}$ is obvious.

Suppose now that X is UMD with constant $\beta_{p,X}$. Then, by considering only real signs in the definition of the UMD property, it is immediate that $X_{\mathbb{R}}$ is UMD with constant $\beta_{p,X}$ as well. This gives the inequality $\beta_{p,X_{\mathbb{R}}} \leqslant \beta_{p,X}$.

Now suppose that $X_{\mathbb{R}}$ is UMD with constant $\beta_{p,X_{\mathbb{R}}}$. If $\alpha = (\alpha_n)_{n=0}^N$ is a sequence of complex numbers of modulus at most one, then α is a convex combination of elements in $(\pi/2)\,\mathrm{abco}\{-1,1\}^{N+1}$ by Proposition 3.2.14, say $\alpha = (\pi/2)\sum_{j=1}^k \lambda^{(j)}\epsilon^{(j)}$ with $\sum_{j=1}^k |\lambda_j| \leqslant 1$. Then, using the triangle inequality in $L^p(S;X)$ and the UMD property of $X_{\mathbb{R}}$, we obtain

$$\begin{aligned}
\Big\| \sum_{n=0}^N \alpha_n d_n \Big\|_{L^p(S;X)} &= \frac{\pi}{2}\Big\| \sum_{j=1}^k \lambda^{(j)} \sum_{n=0}^N \epsilon_n^{(j)} d_n \Big\|_{L^p(S;X)} \\
&\leqslant \frac{\pi}{2}\sum_{j=1}^k |\lambda^{(j)}| \Big\| \sum_{n=0}^N \epsilon_n^{(j)} d_n \Big\|_{L^p(S;X)} \\
&\leqslant \frac{\pi}{2}\beta_{p,X_{\mathbb{R}}} \sum_{j=1}^k |\lambda^{(j)}| \Big\| \sum_{n=0}^N d_n \Big\|_{L^p(S;X)} \\
&\leqslant \frac{\pi}{2}\beta_{p,X_{\mathbb{R}}} \Big\| \sum_{n=0}^N d_n \Big\|_{L^p(S;X)}.
\end{aligned}$$

This gives the inequality $\beta_{p,X} \leqslant (\pi/2)\beta_{p,X_{\mathbb{R}}}$. □

Of course, an easier proof with constant 2 instead of $\pi/2$ could be given.

4.2.b Unconditionality of the Haar decomposition

Here we prove a convenient characterisation of the UMD property in terms of analysis of functions on the real line $\mathbb{R}$ or the unit interval $[0,1)$; besides its intrinsic interest, it will play a major role in the forthcoming study of singular integrals in Chapter 5.

A *dyadic system* (of intervals on $\mathbb{R}$) is a family $\mathscr{D} = \bigcup_{j\in\mathbb{Z}} \mathscr{D}_j$, where each $\mathscr{D}_j$ is a partition of $\mathbb{R}$ consisting of intervals of the form $[x, x+2^{-j})$, and each interval $I \in \mathscr{D}_j$ is a union of two intervals I_- and I_+ (its left and right halves) from $\mathscr{D}_{j+1}$. The particular example $\mathscr{D}^0 = \bigcup_{j\in\mathbb{Z}} \mathscr{D}^0_j$ with $\mathscr{D}^0_j = \{2^{-j}[k,k+1) : k \in \mathbb{Z}\}$ is called the *standard dyadic system* and has already been introduced in Chapter 3.

If $J \in \mathscr{D}$, we denote

$$\mathscr{D}(J) := \big\{ I \in \mathscr{D} : I \subseteq J \big\}.$$

In particular, $\mathscr{D}^0([0,1))$ gives the (standard) dyadic intervals of the unit interval $[0,1)$.

To every interval I, we associate the *Haar function*

$$h_I := |I|^{-1/2}\big(\mathbf{1}_{I_-} - \mathbf{1}_{I_+}\big),$$

where I_- and I_+ are the left and right halves of I. It is easy to check that

$$\langle h_I, h_J\rangle = \int_{\mathbb{R}} h_I(t) h_J(t)\,\mathrm{d}t = \delta_{IJ} \qquad \forall I, J \in \mathscr{D},$$

where $\mathscr{D}$ is any dyadic system; in fact, two such intervals are either disjoint or one is contained in the other, and if, say, $I \subsetneq J$, then h_J is constant on the support of h_I. It follows that the family $(D_I)_{I\in\mathscr{D}}$ of *Haar projections*

$$D_I f := h_I \langle f, h_I\rangle$$

is a pre-decomposition of $L^p(\mathbb{R};X)$ and of $L^p(0,1;X)$ (in the latter case replacing $\mathscr{D}$ by $\mathscr{D}([0,1))$) for any $p \in [1,\infty]$ and any Banach space X.

We denote by $\langle f\rangle_J$ the average of f over J,

$$\langle f\rangle_J := \fint_J f\,\mathrm{d}x = \frac{1}{|J|}\int_J f\,\mathrm{d}x. \tag{4.13}$$

Lemma 4.2.11. *Let $\mathscr{D}$ be a dyadic system on $\mathbb{R}$, and $\mathscr{F}_k = \sigma(\mathscr{D}_k)$, $k \in \mathbb{Z}$, be the associated filtration. Then we have the identities*

$$\begin{aligned} \mathbf{1}_I(\mathbb{E}(f|\mathscr{F}_{k+1}) - \mathbb{E}(f|\mathscr{F}_k)) &= D_I f, && I \in \mathscr{D}_k,\ k \in \mathbb{Z}, \\ \mathbf{1}_J(\mathbb{E}(f|\mathscr{F}_n) - \langle f\rangle_J) &= \sum_{\substack{I\in\mathscr{D}(J)\\ \ell(I)>2^{-n}}} D_I f, && J \in \mathscr{D}_k,\ k \leqslant n. \end{aligned}$$

Proof. Concerning the first claim, we compute (cf. Example 2.6.13)

$$\begin{aligned} &\mathbf{1}_I\big(\mathbb{E}(f|\mathscr{F}_{k+1}) - \mathbb{E}(f|\mathscr{F}_k)\big) \\ &\quad = \frac{\mathbf{1}_{I_-}}{|I_-|}\int_{I_-} f\,\mathrm{d}x + \frac{\mathbf{1}_{I_+}}{|I_+|}\int_{I_+} f\,\mathrm{d}x - \frac{\mathbf{1}_I}{|I|}\int_I f\,\mathrm{d}x \\ &\quad = \mathbf{1}_{I_-}\frac{2}{|I|}\int_{I_-} f\,\mathrm{d}x + \mathbf{1}_{I_+}\frac{2}{|I|}\int_{I_+} f\,\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&\qquad - (\mathbf{1}_{I_-} + \mathbf{1}_{I_+})\frac{1}{|I|}\Big\{\int_{I_-} f\,\mathrm{d}x + \int_{I_+} f\,\mathrm{d}x\Big\}\\
&= \mathbf{1}_{I_-}\Big\{\frac{1}{|I|}\int_{I_-} f\,\mathrm{d}x - \frac{1}{|I|}\int_{I_+} f\,\mathrm{d}x\Big\}\\
&\qquad\qquad + \mathbf{1}_{I_+}\Big\{\frac{1}{|I|}\int_{I_+} f\,\mathrm{d}x - \frac{1}{|I|}\int_{I_-} f\,\mathrm{d}x\Big\}\\
&= \Big(\mathbf{1}_{I_-} - \mathbf{1}_{I_+}\Big)\frac{1}{|I|}\int_{\mathbb{R}} (\mathbf{1}_{I_-} - \mathbf{1}_{I_+}) f\,\mathrm{d}x\\
&= h_I \int_{\mathbb{R}} h_I f\,\mathrm{d}x = D_I f.
\end{aligned}
$$

For the second claim, we expand in a telescopic fashion to deduce

$$
\begin{aligned}
\mathbf{1}_J(\mathbb{E}(f|\mathscr{F}_n) - \langle f\rangle_J) &= \mathbf{1}_J(\mathbb{E}(f|\mathscr{F}_n) - \mathbb{E}(f|\mathscr{F}_k))\\
&= \mathbf{1}_J \sum_{j=k}^{n-1} \big(\mathbb{E}(f|\mathscr{F}_{j+1}) - \mathbb{E}(f|\mathscr{F}_j)\big)\\
&= \sum_{j=k}^{n-1} \sum_{I\in\mathscr{D}_j(J)} \mathbf{1}_I\big(\mathbb{E}(f|\mathscr{F}_{j+1}) - \mathbb{E}(f|\mathscr{F}_j)\big)\\
&= \sum_{j=k}^{n-1} \sum_{I\in\mathscr{D}_j(J)} D_I f = \sum_{\substack{I\in\mathscr{D}(J)\\ \ell(I)>2^{-n}}} D_I f.
\end{aligned}
$$

□

Lemma 4.2.12. *Let X be a Banach space and $p \in (1,\infty)$, and let $\mathscr{D}$ be a dyadic system on $\mathbb{R}$ or $[0,1)$. Then finite linear combinations of $D_I f$, $I \in \mathscr{D}$, are dense in $L^p(\mathbb{R};X)$ or $L^p(0,1;X)$, respectively.*

This lemma cannot be extended to $p \in \{1,\infty\}$. Concerning $p = 1$, we note that the bounded linear operator $f \mapsto \int_{\mathbb{R}} f\,\mathrm{d}x$ annihilates all functions $D_I f$ but not general L^1 functions. For $p = \infty$, counterexamples are also easy and left to the reader.

Proof. Suppose that $f \in L^p(\mathbb{R};X) \cap L^1(\mathbb{R};X)$ has bounded support. Then for all large enough N, the support of f is contained in at most two intervals $J \in \mathscr{D}_{-N}$. Moreover, we have

$$\|\mathbf{1}_J\langle f\rangle_J\|_p \leqslant |J|^{1/p-1}\|f\|_1,$$

which tends to zero as $|J| = 2^N \to \infty$. Thus any f of the stated form, and by their density any $f \in L^p(\mathbb{R};X)$, can be approximated in the norm by a sum of at most two functions of the form $f' = \mathbf{1}_J(f - \langle f\rangle_J)$.

Consider the conditional expectation of such a function with respect to $\mathscr{F}_n = \sigma(\mathscr{D}_n)$ for $n \geqslant -N$. This is given by $\mathbb{E}(f'|\mathscr{F}_n) = \mathbf{1}_J(\mathbb{E}(f|\mathscr{F}_n) - \langle f \rangle_J)$, which is a finite combination of $D_I f$ according to Lemma 4.2.11. On the other hand, from general martingale convergence we know that $\mathbb{E}(f'|\mathscr{F}_n) \to f'$ in $L^p(\mathbb{R}; X)$, which proves the desired density. (As an alternative to martingale convergence, one could also have applied Lebesgue's differentiation theorem combined with a pointwise maximal function bound and dominated convergence to arrive at the same conclusion.)

The case of $L^p(0,1;X)$ is similar and slightly simpler, since bounded support is automatic, and we can simply consider the interval $J = [0,1)$; the rest of the argument is the same. $\square$

Theorem 4.2.13. *Let X be a Banach space and $p \in (1,\infty)$. Then the following conditions are equivalent:*

(1) *X is a UMD space;*
(2) *for every dyadic system $\mathscr{D}$ of $\mathbb{R}$, the Haar projections $(D_I)_{I\in\mathscr{D}}$ form an unconditional decomposition of $L^p(\mathbb{R}; X)$;*
(3) *the Haar projections $(D_I)_{I\in\mathscr{D}^0([0,1))}$ form an unconditional decomposition of $L^p(0,1;X)$.*

Moreover, the unconditionality constant is equal to $\beta_{p,X}$ in each case. Under the equivalent conditions, for all $f \in L^p(\mathbb{R}; X)$ we have the following estimates:

$$\Big\| \sum_{I\in\mathscr{D}} \epsilon_I h_I \langle f, h_I\rangle \Big\|_{L^p(\mathbb{R};X)} \leqslant \beta_{p,X} \|f\|_{L^p(\mathbb{R};X)} \tag{4.14}$$

and

$$\frac{1}{\beta^-_{p,X}} \|f\|_{L^p(\mathbb{R};X)} \leqslant \Big\| \sum_{I\in\mathscr{D}} \varepsilon_I h_I \langle f, h_I\rangle \Big\|_{L^p(\mathbb{R}\times\Omega;X)} \leqslant \beta^+_{p,X} \|f\|_{L^p(\mathbb{R};X)} \tag{4.15}$$

for all signs $\epsilon_I \in \mathbb{K}$ and all Rademacher sequences $(\varepsilon_I)_{I\in\mathscr{D}}$ on any probability space Ω.

Proof. (1)$\Rightarrow$(2): Consider a dyadic system $\mathscr{D}$ of $\mathbb{R}$. Thanks to Lemma 4.2.12 and Proposition 4.1.10, it remains to prove the estimate

$$\Big\| \sum_{I\in F} \epsilon_I D_I f \Big\|_{L^p(\mathbb{R};X)} \leqslant C \Big\| \sum_{I\in F} D_I f \Big\|_{L^p(\mathbb{R};X)} \tag{4.16}$$

for any finite collection $F \subseteq \mathscr{D}$ and signs $\epsilon_I \in \mathbb{K}$, with $C \leqslant \beta_{p,X}$.

Consider the set $S := \bigcup_{I\in F} I$, which has finite Lebesgue measure, and fix an enumeration $F = (I_n)_{n=1}^N$ such that $|I_1| \geqslant |I_2| \geqslant \ldots \geqslant |I_N|$. We define an atomic filtration of S by $\mathscr{G}_n := \sigma(D_{I_k} f : 1 \leqslant k \leqslant n)$. If A is an atom of $\mathscr{G}_n$, it is either a dyadic interval of the form $J_\pm$ for some $J = I_k$ with $k \leqslant n$, or the complement in S of the union of all these intervals.

Consider the interval I_{n+1} that supports $D_{I_{n+1}}f$. By the decreasing ordering of the distinct dyadic intervals, if I_{n+1} intersects some $J = I_k$, $k \leqslant n$, then $I_{n+1} \subsetneq J$, and thus I_{n+1} is a subset of either half, J_- or J_+, of J. The minimal interval $J_\pm$ of this form is an atom A of $\mathscr{G}_n$ that contains I_{n+1}. If, on the other hand, I_{n+1} does not intersect any I_k, then it is contained in their complement in S, which again is an atom A of $\mathscr{G}_n$. In either case, we have

$$\mathbb{E}(D_{I_{n+1}}f|\mathscr{G}_n) = \mathbf{1}_A \langle D_{I_{n+1}f}\rangle_A = 0,$$

and hence $(D_{I_n}f)_{n=1}^N$ is a martingale difference sequence adapted to $(\mathscr{G}_n)_{n=1}^N$. Thus (4.16) follows from the definition of the UMD constant.

(2)⇒(3): We need to prove (4.16) for any finite subset $F \subseteq \mathscr{D}^0([0,1))$. But this is also a finite subset of the dyadic system $\mathscr{D}^0$ of $\mathbb{R}$, so the estimate is immediate from (2).

(3)⇒(1): From Lemma 4.2.11 it follows that

$$\sum_{I\in\mathscr{D}^0_j([0,1))} D_I f = \mathbb{E}(f|\mathscr{F}_{j+1}) - \mathbb{E}(f|\mathscr{F}_j) = df_{j+1}$$

is a martingale difference with respect to the dyadic filtration $(\mathscr{F}_j)_{j\geqslant 0}$ of $[0,1)$, where $\mathscr{F}_j = \sigma(\mathscr{D}^0_j([0,1)))$. Choosing $F := \{I \in \mathscr{D}^0([0,1)) : |I| > 2^{-n}\}$ and factors ϵ_I that depend only on the length $|I|$ of I, (4.16) takes the form

$$\Big\|\sum_{j=0}^{n-1} \epsilon_{j+1} df_{j+1}\Big\|_{L^p(0,1;X)} \leqslant C\Big\|\sum_{j=0}^{n-1} df_{j+1}\Big\|_{L^p(0,1;X)}.$$

After trivial re-indexing, this is the defining inequality of the dyadic UMD property, on a canonical model of Paley–Walsh filtrations given by the dyadic filtration $(\mathscr{F}_j)_{j=0}^\infty$ of $[0,1)$. (It is clear that any finite Paley–Walsh martingale on any probability space has an equally distributed analogue on this canonical model.) Thus we find that the dyadic UMD constant satisfies $\beta^\Delta_{p,X} \leqslant C$, and we conclude by recalling that $\beta_{p,X} = \beta^\Delta_{p,X}$.

(4.14) and (4.15): In the course of proving (1)⇒(2) we already established (4.14) whenever f has a finite Haar expansion. In proving this, we actually found that $D_I f = h_I\langle f, h_I\rangle$, where I ranges over a finite set, can be enumerated as a martingale difference sequence. Whence also (4.15) is immediate from the definition of one-sided UMD constants, for functions f with a finite Haar expansion. Since we already know the summability of the infinite expansion $f = \sum_{I\in\mathscr{D}} h_I\langle f, h_I\rangle$, estimates (4.14) and (4.15) for all $f \in L^p(\mathbb{R}; X)$ follow by convergence. □

4.2.c Examples and constructions

We now proceed to provide several examples of Banach spaces with the UMD property. In this respect, Theorem 4.2.7 is extremely helpful, as it allows us to check the defining condition for a single value of $p \in (1,\infty)$. For a concrete Banach space X, there is often a distinguished value of p, depending on X, for which this is much easier than for other values. Theorem 4.1.1 in particular implies that $X = \mathbb{K}$ is a UMD space. It is quite simple to extend Theorem 4.1.1 to the Hilbert space setting; this is left to the reader.

Proposition 4.2.14 (Pythagoras). *Every Hilbert space H is a UMD space and*

$$\beta_{2,H} = 1.$$

Proof. We use that every martingale difference sequence $(df_n)_{n=1}^N$, and therefore also $(\epsilon_n df_n)_{n=1}^N$, is orthogonal in $L^2(S;H)$ for any σ-finite measure space $(S, \mathscr{A}, \mu)$ (see Proposition 3.5.6). From this it follows that

$$\Big\| \sum_{n=1}^N \epsilon_n df_n \Big\|_{L^2(S;H)} = \Big(\sum_{n=1}^N \|df_n\|_{L^2(S;H)}^2 \Big)^{1/2} = \Big\| \sum_{n=1}^N df_n \Big\|_{L^2(S;H)}.$$

□

Proposition 4.2.15 (Fubini). *Let $(T, \mathscr{B}, \nu)$ be a measure space. For all $p \in (1,\infty)$,*

(1) *$L^p(T)$ is UMD and*

$$\beta_{p,L^p(T)} = \beta_{p,\mathbb{K}}.$$

(2) *$L^p(T;X)$ is UMD whenever if X is UMD and*

$$\beta_{p,L^p(T;X)} = \beta_{p,X}.$$

Proof. We give the proof for the equivalent (by Theorem 4.2.5) dyadic UMD property. This is not essential, but allows us to avoid some tiresome issues of joint measurability and zero sets.

Let X be a UMD space and consider functions $\phi_j : \{-1,1\}^{j-1} \to L^p(T;X)$. We may view them as functions $\phi_j : \{-1,1\}^{j-1} \times T \to X$ such that $\phi_j(\eta_1,\ldots,\eta_{j-1};\cdot) \in L^p(T;X)$ for each $\eta_1,\ldots,\eta_{j-1} \in \{-1,1\}$. Using the representation (4.9),

$$\begin{aligned}
&\Big\| \sum_{j=1}^N \epsilon_j \phi_j(r_1,\ldots,r_{j-1}) r_j \Big\|_{L^p(\Omega;L^p(T;X))} \\
&= \Big(\int_T \Big\| \sum_{j=1}^N \epsilon_j \phi_j(r_1,\ldots,r_{j-1};y) r_j \Big\|_{L^p(\Omega;X)}^p \,\mathrm{d}\nu(y) \Big)^{1/p} \\
&\leqslant \Big(\int_T (\beta_{p,X}^{\Delta})^p \Big\| \sum_{j=1}^N \phi_j(r_1,\ldots,r_{j-1};y) r_j \Big\|_{L^p(\Omega;X)}^p \,\mathrm{d}\nu(y) \Big)^{1/p} \\
&= \beta_{p,X}^{\Delta} \Big\| \sum_{j=1}^N \phi_j(r_1,\ldots,r_{j-1}) r_j \Big\|_{L^p(\Omega;L^p(T;X))},
\end{aligned}$$

where the first and last steps are consequences of Fubini's theorem, since the integral over Ω is just a finite sum.

This proves that $\beta^{\Delta}_{p,L^p(T;X)} \leqslant \beta^{\Delta}_{p,X}$, and the converse bound is immediate by considering ϕ_j that map into a subspace of $L^p(T;X)$ of the form $\{\phi \otimes x : x \in X\}$, where $\phi \in L^p(T) \setminus \{0\}$ is a fixed scalar-valued function. The first claim follows from the fact that $L^p(\nu) = L^p(T;\mathbb{K})$, where the scalar field $\mathbb{K}$ is a UMD space as a one-dimensional Hilbert space by Proposition 4.2.14. □

Remark 4.2.16. It is also true that

$$\beta^{+}_{p,L^p(\Omega;X)} = \beta^{+}_{p,X}, \qquad \beta^{-}_{p,L^p(\Omega;X)} = \beta^{-}_{p,X}.$$

For Paley–Walsh martingales this can be seen as in the above proof. However, it is not clear whether one can reduce to this setting for these cases, and for this reason we need to argue slightly differently. A possible approach is to use Lemma 3.6.17 to reduce to the case of finite σ-algebras. Once this has been done, it easy to extend the above Fubini-type argument.

In the following we collect several basic constructions of UMD spaces.

Proposition 4.2.17 (New UMD spaces from old). *Let X, X_i $(i \in I)$ be UMD spaces. Then each of the following spaces is also UMD:*

(1) *every Banach space $\widetilde{X}$ isomorphic to X, with*

$$\beta_{p,\widetilde{X}} \leqslant \|J\| \cdot \|J^{-1}\| \cdot \beta_{p,X},$$

where $J : X \to \widetilde{X}$ denotes the isomorphism.

(2) *the dual space X^* of X, with*

$$\beta_{p',X^*} = \beta_{p,X}.$$

(3) *every closed subspace $Y \subseteq X$ and its quotient X/Y, with*

$$\beta_{p,Y} \leqslant \beta_{p,X}, \qquad \beta_{p,X/Y} \leqslant \beta_{p,X}.$$

(4) *the ℓ^p-direct sum*

$$\bigoplus_{i\in\mathscr{I}}^{p} X_i := \Big\{x = (x_i)_{i\in\mathscr{I}} : x_i \in X_i, \|x\| := \Big(\sum_{i\in\mathscr{I}} \|x_i\|^p_{X_i}\Big)^{1/p} < \infty\Big\},$$

for any index set $\mathscr{I}$, with

$$\beta_{p,\bigoplus^p_{i\in\mathscr{I}} X_i} = \sup_{i\in\mathscr{I}} \beta_{p,X_i},$$

provided the supremum on the right side is finite.

(5) *the complex and real interpolation spaces* $[X_0, X_1]_\theta$ *and* $(X_0, X_1)_{\theta,p_0,p_1}$, *with*

$$\beta_{p,[X_0,X_1]_\theta} \leqslant \beta_{p_0,X_0}^{1-\theta}\beta_{p_1,X_1}^{\theta},$$
$$\beta_{p,(X_0,X_1)_{\theta,p_0,p_1}} \leqslant \beta_{p_0,X_0}^{1-\theta}\beta_{p_1,X_1}^{\theta},$$

provided that (X_0, X_1) *is an interpolation couple of Banach spaces,* $\theta \in (0,1)$, *and* $1 < p_0 < p < p_1 < \infty$ *satisfy* $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$.

In all these assertions, the estimates for the UMD constants hold for $p \in (1,\infty)$.

With regard to (5) we recall that $(X_0, X_1)_{\theta,p_0,p_1} = (X_0, X_1)_{\theta,p}$ up to equivalent norms, where $(X_0, X_1)_{\theta,p_\theta}$ is the usual real interpolation space and $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$ (see Theorem C.3.14).

Proof. The cases $\widetilde{X} \eqsim X$ and $Y \subseteq X$ are immediate from the definition, as $\widetilde{X}$ or Y-valued martingales are immediately identified with X-valued ones, either via the isomorphism J for $\widetilde{X}$ or directly by the embedding. A direct computation gives "$\leqslant$" for $\bigoplus_{i\in\mathscr{I}}^p X_i$, and "$\geqslant$" follows from the previous cases by identifying each X_j as a closed subspace of $\bigoplus_{i\in\mathscr{I}}^p X_i$ in the natural way.

As for the dual space X^*, recall that $\beta_{p,X}$ is the supremum of the norms of certain martingale transforms T on $L^p(S;X)$, and the formal adjoint T^* on $L^{p'}(\mu;X^*)$ has exactly the same form. Taking the supremum over all T it follows that $\beta_{p',X^*} \leqslant \beta_{p,X}$, and by the same argument $\beta_{p,X^{**}} \leqslant \beta_{p',X^*}$. Since X is isomorphic to a closed subspace of X^{**}, we have $\beta_{p,X} \leqslant \beta_{p,X^*}$ by the cases already considered.

For the quotient X/Y, recall that the *annihilator* $Y^\perp$ of $Y \subseteq X$, given by

$$Y^\perp = \{x^* \in X^* : \langle y, x^*\rangle = 0 \text{ for all } y \in Y\},$$

is a closed subspace of X^* isometric to $(X/Y)^*$. Combining the previous estimates it then follows that

$$\beta_{p,X/Y} = \beta_{p',(X/Y)^*} = \beta_{p',Y^\perp} \leqslant \beta_{p',X^*} = \beta_{p,X}.$$

Concerning the interpolation results, let T be one of the martingale transforms appearing in the equivalent definition of UMD in Lemma 4.2.8. Then Theorems 2.2.6 and C.2.6 show that

$$\begin{aligned}\|T\|_{\mathscr{L}(L^p(S;[X_0,X_1]_\theta))} &= \|T\|_{\mathscr{L}([L^{p_0}(\mu;X_0),L^{p_1}(\mu;X_1)]_\theta)} \\ &\leqslant \|T\|_{\mathscr{L}(L^{p_0}(\mu;X_0))}^{1-\theta}\|T\|_{L^{p_1}(\mu;X_1)}^{\theta}.\end{aligned}$$

Taking the supremum over all relevant T, this gives $\beta_{p,[X_0,X_1]_\theta} \leqslant \beta_{p_0,X_0}^{1-\theta}\beta_{p_1,X_1}^{\theta}$. The argument for $(X_0, X_1)_{\theta,p_0,p_1}$ is similar, now using Theorems 2.2.10 and C.3.16 instead; this gives

$$\begin{aligned}\|T\|_{\mathscr{L}(L^p(S;(X_0,X_1)_{\theta,p_0,p_1}))} &= \|T\|_{\mathscr{L}((L^{p_0}(\mu;X_0),L^{p_1}(\mu;X_1))_{\theta,p_0,p_1})} \\ &\leqslant \|T\|^{1-\theta}_{\mathscr{L}(L^{p_0}(\mu;X_0))}\|T\|^{\theta}_{L^{p_1}(\mu;X_1)}.\end{aligned}$$

□

Example 4.2.18. Let $1 \leqslant p \leqslant \infty$. The Sobolev space

$$W^{k,p}(\mathbb{R}^d) := \Big\{ f \in L^p(\mathbb{R}^d) : \ \partial^\alpha f \in L^p(\mathbb{R}^d) \text{ for all multi-indices } |\alpha| \leqslant k \Big\},$$

with norm

$$\|f\|_{W^{k,p}(\mathbb{R}^d)} := \sum_{|\alpha|\leqslant k} \|\partial^\alpha f\|_{L^p(\mathbb{R}^d)}$$

(see Definition 2.5.4) is a UMD space for each $p \in (1,\infty)$ and $k \in \mathbb{N}$, by identifying it with the closed subspace of all $(g_\alpha)_{|\alpha|\leqslant k}$ in $\bigoplus_{|\alpha|\leqslant k} L^p(\mathbb{R}^d)$ for which $\partial^\alpha g_0 = g_\alpha$ for all $|\alpha| \leqslant k$. The closedness of this subspace depends on the fact that each weak partial derivative ∂^α is a closed operator in $L^p(\mathbb{R}^d)$. A similar identification may be used for several other function spaces built from the L^p spaces (such as Besov spaces), thereby establishing their UMD property.

In the scale of L^p spaces, the UMD property cannot be pushed beyond the exponents $1 < p < \infty$ covered by Proposition 4.2.15:

Proposition 4.2.19. *The spaces c_0, ℓ^1, and ℓ^∞ fail to be UMD, and their finite-dimensional versions satisfy*

$$\tfrac{1}{2}n \leqslant \beta_{p',\ell^1_{2^n}} = \beta_{p,\ell^\infty_{2^n}} \leqslant 200p^* n, \quad n \geqslant 1, \ p \in (1,\infty). \tag{4.17}$$

The randomised UMD constants introduced in (4.10) *satisfy*

$$\beta^-_{p,\ell^\infty_{2^n}} \geqslant c_p\sqrt{n} \ \ \textit{and} \ \ \beta^+_{p,\ell^\infty_{2^n}} \geqslant c_p\sqrt{n}, \quad n \geqslant 1, \ p \in (1,\infty).$$

With tools to be developed below, the failure of UMD for c_0, ℓ^1, and ℓ^∞ is immediate from a general result that UMD spaces are reflexive (Theorem 4.3.3). The quantitative information provided by (4.17), however, is of independent interest.

Proof. The core of the proof is a concrete example of a (Paley–Walsh) martingale that provides the estimate $\beta_{p,\ell^\infty(I)} \geqslant \frac{1}{2}n$, where $I := \{-1,1\}^n$. This also proves the estimate for $\beta_{p,\ell^\infty_{2^n}}$, since $\ell^\infty(I)$ can be identified with $\ell^\infty_{2^n}$. The underlying space I will be endowed with the uniform measure μ which gives each of the 2^n points mass 2^{-n}.

On $I \times I$, consider the functions

$$d_j(s,t) := \mathbf{1}_{\{s_1=t_1,\dots,s_{j-1}=t_{j-1}\}} \cdot s_j \cdot t_j$$

$$= \begin{cases} +1, & \text{if } s_1 = t_1, \dots, s_{j-1} = t_{j-1}, \text{ and } s_j = t_j, \\ -1, & \text{if } s_1 = t_1, \dots, s_{j-1} = t_{j-1}, \text{ and } s_j \neq t_j, \\ 0, & \text{if } s_i \neq t_i \text{ for some } i < j. \end{cases}$$

It is evident from the formula that $(s \mapsto d_j(s,t))_{j=1}^n$ is a Paley–Walsh martingale difference sequence in $L^p(I)$ for every fixed $t \in I$, and therefore $(s \mapsto (d_j(s,t))_{t\in I})_{j=1}^n$ is a Paley–Walsh martingale difference sequence in $L^p(I;\ell^\infty(I))$. It follows from the definition of UMD (applied to the difference sequence $\epsilon_j d_j$ in place of d_j, with the choice $\epsilon_j = (-1)^j$) that

$$\begin{aligned} n = \inf_{s\in I} \Big| \sum_{j=1}^n d_j(s,s) \Big| &\leqslant \Big(\int_I \Big\| \sum_{j=1}^n d_j(s,\cdot) \Big\|_{\ell^\infty(I)}^p \, \mathrm{d}\mu(s) \Big)^{1/p} \\ &\leqslant \beta_{p,\ell^\infty(I)} \Big(\int_I \Big\| \sum_{j=1}^n (-1)^j d_j(s,\cdot) \Big\|_{\ell^\infty(I)}^p \, \mathrm{d}\mu(s) \Big)^{1/p} \\ &\leqslant \beta_{p,\ell^\infty(I)} \sup_{s,t\in\{-1,1\}^n} \Big| \sum_{j=1}^n (-1)^j d_j(s,t) \Big| \\ &\overset{(*)}{\leqslant} 2\beta_{p,\ell^\infty(I)}, \end{aligned}$$

where the first identity comes from observing that $d_j(s,s) = 1$ for all j and s and $(*)$ is proved as follows. For given $s,t \in I$, let $k \in \{1,\dots,n\}$ be the largest index such that $s_i = t_i$ for $i = 1,\dots,k$ (so that $s_{k+1} \neq t_{k+1}$ if $k < n$) and set $k = 0$ if $s_1 \neq t_1$. Thus $d_j(s,t) = 1$ for $j = 1,\dots,k$, $d_{k+1}(s,t) = -1$ (if $k < n$), $d_j(s,t) = 0$ for $j > k+1$, and

$$\begin{aligned} \sum_{j=1}^n (-1)^j d_j(s,t) &= \sum_{j=1}^k (-1)^j \cdot 1 + (-1)^{k+1} \cdot (-1) \cdot \mathbf{1}_{\{k<n\}} \\ &= (-1)\mathbf{1}_{\{k \text{ odd}\}} + (-1)^k \mathbf{1}_{\{k<n\}} \in \{-2,-1,0,1\}, \end{aligned} \tag{4.18}$$

and the absolute value is bounded by 2, as claimed.

This immediately shows the failure of UMD for c_0 and ℓ^∞, since the same finite-dimensional example could obviously have been set on the first 2^n coordinates of either of these infinite-dimensional spaces (or one could argue by Proposition 4.2.17(3)).

Next we prove the upper bound for β_{p,ℓ_d^∞} with $d = 2^n$. For this we will use that $\beta_{p,\ell_d^p} = \beta_{p,\mathbb{K}} \leqslant 100p^*$ with $p^* = \max\{p,p'\}$, which follows from Theorem 4.2.7 and Proposition 4.2.14. This estimate will be improved to $\beta_{p,\mathbb{K}} = p^* - 1$ in Theorem 4.5.7 for $\mathbb{K} = \mathbb{R}$ and in Corollary 4.5.15 for $\mathbb{K} = \mathbb{C}$. In order to avoid a forward reference we just use that $\beta_{p,\ell_d^p} \leqslant Kp^*$ and the reader may choose whether to take $K = 100$ (which gives (4.17) with constant 20 000)) or $K = 1$.

For any L^p-martingale $(f_j)_{j=0}^k$ with values in ℓ_d^∞ with $f_0 = 0$ and any sequence of signs $(\epsilon_j)_{j=1}^k$ in $\mathbb{K}$ we have the estimate

$$\Big\|\sum_{j=1}^k \epsilon_j df_j\Big\|_{L^p(S;\ell^\infty_d)} \leqslant \Big\|\sum_{j=1}^k \epsilon_j df_j\Big\|_{L^p(S;\ell^p_d)}$$
$$\leqslant Kp^*\|f_k\|_{L^p(S;\ell^p_d)} \leqslant Kp^* d^{1/p}\|f_k\|_{L^p(S;\ell^\infty_d)}.$$

It follows that $\beta_{p,\ell^\infty_d} \leqslant Kp^* d^{1/p}$.

Set $p_d := n$ and assume that $n \geqslant 2$, so that $p_d^* = p_d = n$. Then $\beta_{p_d,\ell^\infty_d} \leqslant 2Kn$, so that Theorem 4.2.7 yields the estimate

$$\beta_{p,\ell^\infty_d} \leqslant 100\Big(\frac{p}{p_d} + \frac{p'}{p'_d}\Big)\beta_{p_d,\ell^\infty_d} \leqslant 100p^*\Big(\frac{1}{n} + \frac{1}{n'}\Big)2Kn = 200Kp^*n.$$

For $n = 1$ (i.e., $d = 2$) we note that ℓ^∞_2 is isomorphic to ℓ^2, and by Proposition 4.2.17(1) $\beta_{p,\ell^\infty_2} \leqslant \sqrt{2}\beta_{p,\ell^2_2} \leqslant \sqrt{2}p^* \leqslant 200Kp^*n$.

The claims concerning ℓ^1 and $\ell^1_{2^n}$ follow from Proposition 4.2.17(2) and the fact that $(\ell^1)^* = \ell^\infty$ and $(\ell^1_{2^n})^* = \ell^\infty_{2^n}$.

To conclude we prove the estimates for $\beta^\pm_{p,\ell^\infty_{2^n}}$. Let $(\varepsilon_j)_{j\geqslant 1}$ be a Rademacher sequence. As in (4.18) we find, using symmetry,

$$\Big|\sum_{j=1}^n \varepsilon_j d_j(s,t)\Big| \leqslant \sup_{1\leqslant m\leqslant n}\Big|\sum_{j=1}^m \varepsilon_j + \varepsilon_{m+1}\mathbf{1}_{\{m<n\}}\Big| \leqslant S_n^\star + 1,$$

where $S_n^\star = \max_{1\leqslant m\leqslant n}|S_m|$ with $S_m = \sum_{j=1}^m \varepsilon_j$. Taking $L^p(\mathbb{P}\times\mu;\ell^\infty(I))$-norms,

$$\Big(\int_I \mathbb{E}\Big\|\sum_{j=1}^n \varepsilon_j d_j(s,\cdot)\Big\|^p_{\ell^\infty(I)}\,\mathrm{d}\mu(s)\Big)^{1/p} \leqslant (\mathbb{E}|S_n^\star|^p)^{1/p} + 1$$
$$\leqslant p'(\mathbb{E}|S_n|^p)^{1/p} + 1 \leqslant C_p\sqrt{n},$$

where we applied Doob's maximal inequality and Khintchine's inequality (Corollary 3.2.24). Similarly, since $d_j(s,s) = 1$, we find

$$c_p\sqrt{n} \leqslant (\mathbb{E}|S_n|^p)^{1/p} \leqslant \Big(\int_I \mathbb{E}\Big\|\sum_{j=1}^n \varepsilon_j d_j(s,\cdot)\Big\|^p_{\ell^\infty(I)}\,\mathrm{d}\mu(s)\Big)^{1/p}.$$

By symmetry the same estimates hold with $\varepsilon_j d_j$ replaced by $\varepsilon_j(-1)^j d_j$. As in the proof of the estimates for $\beta_{p,\ell^\infty(I)}$ it follows that $\beta^-_{p,\ell^\infty(I)} \geqslant C_p^{-1}\sqrt{n}$ and $\beta^+_{p,\ell^\infty(I)} \geqslant c_p\sqrt{n}/2$. □

Example 4.2.20 (Classical spaces without UMD). The spaces

$$C([0,1]), \quad L^1(0,1), \quad L^\infty(0,1), \quad \mathscr{K}(\ell^2), \quad \mathscr{C}^1$$

fail to be UMD spaces. For $C([0,1])$, $L^\infty(0,1)$, and $\mathscr{K}(\ell^2)$, this follows from Proposition 4.2.19 and the fact that all these spaces contain isometric copies of c_0. For $\mathscr{K}(\ell^2)$, such a copy of c_0 is obtained by embedding

$$a = (a_n)_{n=1}^\infty \mapsto \operatorname{diag} a := \sum_{n=1}^\infty a_n e_n(\cdot|e_n) \tag{4.19}$$

into the family of diagonal matrices in a given orthonormal basis $(e_n)_{n=1}^\infty$.

The spaces $L^1(0,1)$ and the Schatten class $\mathscr{C}^1$ similarly contain isometric copies of ℓ^1, where the latter case can be realised by the same diagonal map (4.19). Alternatively, this could be deduced from the previous cases via the isometric dualities $(L^1(0,1))^* \eqsim L^\infty(0,1)$ and $\mathscr{C}^1 \eqsim (\mathscr{K}(\ell^2))^*$ (see Theorem D.2.6 for the latter).

Example 4.2.21. The Schatten classes $\mathscr{C}^p$ are UMD spaces for all $p \in (1,\infty)$. This will be proved in Proposition 5.4.2. The case $p=2$ already follows from Proposition 4.2.14 and the fact that $\mathscr{C}^2$ is a Hilbert space. The case of general $p \in (1,\infty)$ will be deduced from this by an extrapolation argument along the "ladder" of dyadic powers $p = 2^n$, completed by interpolation (Proposition 4.2.17(5)) for other $p \in (2,\infty)$, and duality (Proposition 4.2.17(2)) for $p \in (1,2)$.

We conclude with some further cases where the UMD constant $\beta_{p,X}$ of a Banach space X reduces to $\beta_{p,\mathbb{K}}$. Later on, in Corollary 4.5.15, we will show that $\beta_{p,\mathbb{K}} = p^* - 1$ where $p^* = \max\{p,p'\}$.

Proposition 4.2.22. *The UMD constants of a Hilbert space H coincide with the scalar case:*

$$\beta_{p,H} = \beta_{p,\mathbb{K}} \qquad \forall\, p \in (1,\infty).$$

The UMD constants of any L^q space coincide with the scalar case for q between 2 and p:

$$\beta_{p,L^q} = \beta_{p,\mathbb{K}} \qquad \forall\, 1 < p \leqslant q \leqslant 2, \quad \forall\, 2 \leqslant q \leqslant p < \infty.$$

Proof. In both cases, "$\geqslant$" is clear by identifying $\mathbb{K}$ with an arbitrary one-dimensional subspace of X. The bounds "$\leqslant$" are proved by an application of some general operator results to the martingale transforms T, whose norms give the relevant UMD constants of interest.

For the first assertion, we recall the theorem of Paley and Marcinkiewicz–Zygmund (Theorem 2.1.9), which asserts that $\|T\otimes I_H\|_{\mathscr{L}(L^p(H))} = \|T\|_{\mathscr{L}(L^p)}$. Taking the supremum over the relevant T, we arrive at $\beta_{p,H} = \beta_{p,\mathbb{K}}$.

For the second assertion, note that we already know the cases that $q=2$ (Proposition 4.2.14) and $q=p$ (Proposition 4.2.15). For the intermediate values, we can write $1/q = (1-\theta)/2 + \theta/p$ so that $L^q = [L^2, L^p]_\theta$, and apply Proposition 4.2.17(5) to deduce that

$$\beta_{p,L^q} \leqslant \beta_{p,L^2}^{1-\theta}\beta_{p,L^p}^{\theta} = \beta_{p,\mathbb{K}}^{1-\theta}\beta_{p,\mathbb{K}}^{\theta} = \beta_{p,\mathbb{K}}.$$

$\square$

4.2.d Stein's inequality for conditional expectations

In this section we will prove the fundamental fact that for UMD spaces X, the conditional expectations with respect to an increasing sequence of σ-algebras satisfy a certain randomised inequality known as *Stein's inequality*. This result marked the starting point of the theory of R-boundedness, which will be briefly touched upon in Chapter 5 and fully developed in Volume II.

Recall that the constants $\beta^{\pm}_{p,X}$ have been defined as the best constants in (4.10).

Theorem 4.2.23 (Stein's inequality for conditional expectations). *Let X be a UMD space and let $p \in (1,\infty)$. Let $(S,\mathscr{A},\mu)$ be a measure space equipped with a σ-finite filtration $(\mathscr{F}_n)_{n=1}^N$. Then for all finite sequences $(f^n)_{n=1}^N$ in $L^p(S;X)$ we have*

$$\Big\| \sum_{n=1}^N \varepsilon_n \mathbb{E}(f^n|\mathscr{F}_n) \Big\|_{L^p(S\times\Omega;X)} \leqslant \beta^+_{p,X} \Big\| \sum_{n=1}^N \varepsilon_n f^n \Big\|_{L^p(S\times\Omega;X)},$$

where $(\varepsilon_n)_{n=1}^N$ is a Rademacher sequence on a probability space $(\Omega,\mathscr{B},\mathbb{P})$.

It is important to observe that we do not assume the sequence $(f^n)_{n=1}^N$ to be adapted to $(\mathscr{F}_n)_{n=1}^N$. Also notice that the constant in this estimate is independent of N.

Proof of Theorem 4.2.23. We consider the sub-σ-algebras

$$\mathscr{G}_0 = \mathscr{F}_0, \qquad \mathscr{G}_{2k-1} = \sigma(\mathscr{F}_k, \varepsilon_0, \dots, \varepsilon_{k-1}),$$
$$\mathscr{G}_{2k} = \sigma(\mathscr{F}_k, \varepsilon_0, \dots, \varepsilon_k), \quad k = 1,\dots,N$$

of $\mathscr{A}\times\mathscr{B}$ and $F := \sum_{n=1}^N \varepsilon_n f^n$. Consider the martingale $F_k := \mathbb{E}(F|\mathscr{G}_k)$, $k = 0,\dots,2N$. By Proposition 2.6.36, we find that

$$F_{2k} = \sum_{n=1}^k \varepsilon_n \mathbb{E}(f^n|\mathscr{F}_k), \qquad F_{2k-1} = \sum_{n=1}^{k-1} \varepsilon_n \mathbb{E}(f^n|\mathscr{F}_k),$$

so that the even part of the difference sequence is given by

$$dF_{2k} = \varepsilon_k \mathbb{E}(f^k|\mathscr{F}_k).$$

(We could also compute dF_{2k-1}, but an explicit expression for this will not be relevant below.)

Randomising with another Rademacher sequence $(\varepsilon'_k)_{k=0}^{2N}$ on another probability space Ω', we compute

$$
\begin{aligned}
&\Big\| \sum_{n=1}^{N} \varepsilon_n \mathbb{E}(f^n | \mathscr{F}_n) \Big\|_{L^p(S\times\Omega;X)} \\
&= \Big\| \sum_{n=1}^{N} \varepsilon'_{2n} dF_{2n} \Big\|_{L^p(S\times\Omega\times\Omega';X)} \leqslant \Big\| \sum_{k=1}^{2N} \varepsilon'_k dF_k \Big\|_{L^p(S\times\Omega\times\Omega';X)} \\
&\leqslant \beta^+_{p,X} \|F_{2N}\|_{L^p(S\times\Omega;X)} \leqslant \beta^+_{p,X} \Big\| \sum_{n=1}^{N} \varepsilon_n f^n \Big\|_{L^p(S\times\Omega;X)},
\end{aligned}
$$

where the first inequality uses the contraction principle, the second one is the definition of the constant $\beta^+_{p,X}$, and the last one uses the contractivity of the conditional expectation $\mathbb{E}(\cdot|\mathscr{G}_{2N})$. □

Remark 4.2.24. Stein's inequality can be rephrased as the boundedness of a suitable operator-valued martingale transform as defined in Section 3.5. In particular, Theorem 3.5.4 can be used to find equivalent formulations of Theorem 4.2.23 using weak L^p spaces and Hardy spaces.

In order to explain this, let $\varepsilon^p_N(X)$ be the product space $X^N = X \times \cdots \times X$ endowed with the norm

$$
\Big\| \sum_{n=1}^{N} e_n \otimes x^n \Big\|_{\varepsilon^p_N(X)} := \Big\| \sum_{n=1}^{N} \varepsilon_n x^n \Big\|_{L^p(\Omega;X)}.
$$

Here we write $\sum_{n=1}^{N} e_n \otimes x^n$ for the vector $(x^1, \dots, x^N)$, since we think of X^N as the tensor product $\mathbb{K}^N \otimes X$. The estimate of Theorem 4.2.23 is equivalent to the boundedness of the operator-valued martingale transform

$$
T_v : f \mapsto v_0 f_0 + \sum_{k=1}^{N} v_k df_k
$$

on $L^p(S; \varepsilon^p_N(X))$ with $\|T_v\| \leqslant \beta^+_{p,X}$, where $v_0 = I$ and the operators $v_k \in \mathscr{L}(\varepsilon^p_N(X))$ for $k = 1, \dots, N$ are given by the projections

$$
v_k \sum_{n=1}^{N} e_n \otimes x^n = \sum_{n=k}^{N} e_n \otimes x^n.
$$

Indeed, this follows from

$$
\begin{aligned}
T_v \sum_{n=1}^{N} e_n \otimes f^n &= \sum_{n=1}^{N} e_n \otimes f^n_0 + \sum_{k=1}^{N} \sum_{n=k}^{N} e_n \otimes df^n_k \\
&= \sum_{n=1}^{N} e_n \otimes \Big(f^n_0 + \sum_{k=1}^{N} df^n_k \Big) = \sum_{n=1}^{N} e_n \otimes \mathbb{E}(f^n | \mathscr{F}_n).
\end{aligned}
$$

4.2.e Boundedness of martingale transforms

We now return to the general martingale transforms

$$Tf = v \star f := v_{-\infty} f_{-\infty} + \sum_{k \in \mathbb{Z}} v_k df_k \tag{4.20}$$

introduced in Section 3.5. In Theorem 3.5.4, we showed the equivalence (i.e., the simultaneous truth or failure) of various L^p inequalities for these operators. We now provide a condition under which some, and hence by Theorem 3.5.4 all, of these norm estimates are actually valid.

As we have observed above, the very definition of UMD postulates the boundedness of certain particular martingale transform on $L^p(S; X)$. It turns out that these special transforms are extremal, in the sense that their uniform boundedness already controls the norms of a much wider class of martingale transforms. This is the content of the following:

Theorem 4.2.25 (Burkholder). *Let X be a UMD space and $p \in (1, \infty)$. Then all martingale transforms (4.20) with a scalar-valued predictable sequence $v = (v_k)_{k \in \{-\infty\} \cup \mathbb{Z}}$ such that $\|v\|_\infty \leqslant 1$ are bounded on $L^p(S; X)$ with norm at most $\beta_{p,X}$.*

We remind the reader that predictability has been defined in Example 3.1.7.

Proof. Recall that the boundedness of such transforms was defined as the boundedness on test functions $f \in L^p(S; X)$ with finitely non-zero difference sequence. Thus (after re-indexing a fixed finite sum) it suffices to consider the uniform boundedness of

$$Tf = v_0 f_0 + \sum_{k=1}^{n} v_k df_k,$$

where $f = f_n = f_0 + \sum_{k=1}^{n} df_k$. We now apply Proposition 3.6.16 to produce another martingale $(g_j)_{j=0}^N$ and a predictable sequence $(u_j)_{j=0}^N$ such that:

- (i) $(g_j)_{j=0}^N$ and $(u_j)_{j=0}^N$ are supported on a set $E \subseteq S$ of finite μ-measure,
- (ii) their restrictions to E form a martingale and a predictable sequence with respect to an incremental filtration $(\mathscr{G}_j)_{j=0}^N$ of σ-algebras of E,
- (iii) $\|u\|_\infty \leqslant \|v\|_\infty$ and $u_j = \mathbf{1}_E \lambda_j$ for constants λ_j, for all $j = 1, \ldots, N$,
- (iv) $\|f_k - g_{\phi(k)}\|_p < \varepsilon$ and $\|(v \star f)_k - (u \star g)_{\phi(k)}\|_p < \varepsilon$ for some increasing $\phi : \{0, 1, \ldots, n\} \to \{0, 1, \ldots, \phi(n) = N\}$.

Note that $(\mathbf{1}_A g_j)_{j=0}^\infty$ is also a martingale for every atom $A \in \mathscr{G}_0^*$, and u_0 is constant on A, say $u_0 \equiv \lambda_A$, by $\mathscr{G}_0$-measurability. Moreover, $(\lambda_A, \lambda_1, \ldots, \lambda_N)$, where each component is a scalar of absolute value at most one, belongs to the convex hull of $\epsilon = (\epsilon_0, \epsilon_1, \ldots, \epsilon_N)$, where $\epsilon_j \in S_{\mathbb{K}}$. Thus

$$\mathbf{1}_A(u \star g)_N = \lambda_A(\mathbf{1}_A g)_0 + \sum_{j=1}^{N} \lambda_j d(\mathbf{1}_A g)_j$$

belongs to the convex hull of the transforms $T_\epsilon(\mathbf{1}_A g)_N$ for $\epsilon \in S_{\mathbb{K}}^{N+1}$, which appear in the very definition of UMD spaces. It follows from the definition and convexity of the norm that

$$\|(u \star g)_N\|_{L^p(A;X)} \leqslant \beta_{p,X} \|g_N\|_{L^p(A;X)}.$$

Thus, using (iv),

$$\begin{aligned}
\|(v \star f)_n\|_{L^p(S;X)} &\leqslant \|(u \star g)_{\phi(n)}\|_{L^p(E,\mu_E;X)} + \varepsilon \\
&= \Big(\sum_{A \in \mathscr{G}_0} \|(u \star g)_N\|_{L^p(A;X)}^p \Big)^{1/p} + \varepsilon \\
&\leqslant \Big(\sum_{A \in \mathscr{G}_0} (\beta_{p,X} \|g_N\|_{L^p(A;X)})^p \Big)^{1/p} + \varepsilon \\
&= \beta_{p,X} \|g_{\phi(n)}\|_{L^p(E,\mu_E;X)} + \varepsilon \\
&\leqslant \beta_{p,X} \|f_n\|_{L^p(S;X)} + (\beta_{p,X} + 1)\varepsilon,
\end{aligned}$$

and letting $\varepsilon \to 0$ concludes the estimate. □

Remark 4.2.26. While the given proof of Theorem 4.2.25 was not overly long, it depended on the somewhat technical approximation result of Proposition 3.6.16. It turns out that the qualitative conclusion of Theorem 4.2.25, with a weaker constant, can also be achieved by an alternative argument avoiding the martingale approximations, but using the randomised form of the UMD inequality instead. This proof method has some independent interest, and runs as follows. Under the assumptions of Theorem 4.2.25 we have, given a Rademacher sequence $(\varepsilon_k)_{k=1}^n$,

$$\begin{aligned}
\Big\| v_0 f_0 + \sum_{k=1}^{n} v_k df_k \Big\|_{L^p(S;X)} &\leqslant \beta_{p,X}^- \Big\| \varepsilon_0 v_0 f_0 + \sum_{k=1}^{n} \varepsilon_k v_k df_k \Big\|_{L^p(S \times \Omega;X)} \\
&\leqslant \beta_{p,X}^- \|v\|_\infty \Big\| \varepsilon_0 f_0 + \sum_{k=1}^{n} \varepsilon_k df_k \Big\|_{L^p(S \times \Omega;X)} \\
&\leqslant \beta_{p,X}^- \beta_{p,X}^+ \Big\| f_0 + \sum_{k=1}^{n} df_k \Big\|_{L^p(S;X)}.
\end{aligned}$$

This produces the constant $\beta_{p,X}^- \beta_{p,X}^+$ (which can be bounded by $(\beta_{p,X})^2$) in place of $\beta_{p,X}$ in Theorem 4.2.25.

4.3 Banach space properties implied by UMD

4.3.a Reflexivity

The present subsection is devoted to a proof that UMD spaces are reflexive. The main ingredient is an ingenious characterisation of reflexive Banach spaces due to James to which we turn first. After having completed the proof of our main result, Theorem 4.3.3, we shall explain how the equivalence of the UMD property and its dyadic counterpart implies a self-improvement of Theorem 4.3.3: UMD spaces are super-reflexive.

James's theorem

We start with an auxiliary criterion concerning solubility of abstract linear equations. Observe that reflexivity is about solving, given any $x^{**} \in X^{**}$, for $x \in X$ the uncountable system of equations $\langle x, x^* \rangle = \langle x^*, x^{**} \rangle$, $x^* \in X^*$. The next criterion only involves a finite system.

Proposition 4.3.1 (Helly's condition). *Let X be a normed linear space and let $x_n^* \in X^*$ and $c_n \in \mathbb{K}$, $n = 1, \dots, N$, and $M > 0$ be given. The following assertions are equivalent:*

(1) *for every $\varepsilon > 0$ there exists $x \in X$ such that $\|x\| \leqslant M + \varepsilon$ and*

$$\langle x, x_n^* \rangle = c_n, \quad \forall n = 1, \dots, N;$$

(2) *we have*

$$\Big| \sum_{n=1}^N a_n c_n \Big| \leqslant M \Big\| \sum_{n=1}^N a_n x_n^* \Big\| \qquad \forall a_n \in \mathbb{K}, \quad n = 1, \dots, N. \tag{4.21}$$

Observe, for later use, that in the estimate (4.21) we can always normalise one of the a_n, say a_j, to unity and vary the others, yielding an equivalent version of Helly's condition. Indeed, the case of general $(a_n)_{n=1}^N$ follows by a simple scaling argument if $a_j \neq 0$, and by a limiting argument in the case that $a_j = 0$.

Any constant $M > 0$ for which Helly's condition holds will be called a *Helly constant* for the vectors x_n^* and scalars c_n, $1 \leqslant n \leqslant N$.

Proof. If (1) holds, then for all $\varepsilon > 0$ we can find an $x \in X$ such that

$$(M + \varepsilon) \Big\| \sum_{n=1}^N a_n x_n^* \Big\| \geqslant \Big| \Big\langle x, \sum_{n=1}^N a_n x_n^* \Big\rangle \Big| = \Big| \sum_{n=1}^N a_n c_n \Big|,$$

and (4.21) follows upon letting $\varepsilon \to 0$.

Suppose now that (2) holds, and also assume for the moment that the x_n^* are linearly independent. Consider the mapping

$$T : X \to \mathbb{K}^N : x \mapsto (\langle x, x_n^* \rangle)_{n=1}^N$$

which is obviously linear and continuous (since each of the components are).

Suppose, for a contradiction, that our linear system in (1) is not soluble in $\overline{B}(0, M + \varepsilon)$ for a certain $\varepsilon > 0$. This is equivalent to saying that the point $c = (c_n)_{n=1}^N \in \mathbb{K}^N$ is not in the convex subset $S := T(\overline{B}(0; M + \varepsilon))$ of $\mathbb{K}^N$. But then we can find a separating hyperplane through c having the whole of S on one of its sides, i.e., for a certain non-zero $p \in \mathbb{K}^N$ (perpendicular to the hyperplane) we have $\Re p \cdot y \leqslant \Re p \cdot c$ for all $y \in S$. Using the definition of S, this means that

$$\Re \sum_{n=1}^N p_n \langle x, x_n^* \rangle \leqslant \Re \sum_{n=1}^N p_n c_n$$

for all $x \in \overline{B}(0; M + \varepsilon)$. Since this holds equally well with ζx, $|\zeta| = 1$, in place of x, we actually have an inequality for the absolute values,

$$\Big| \sum_{n=1}^N p_n \langle x, x_n^* \rangle \Big| \leqslant \Big| \sum_{n=1}^N p_n c_n \Big|.$$

But then

$$\begin{aligned}(M + \varepsilon) \Big\| \sum_{n=1}^N p_n x_n^* \Big\| &= \sup_{x \in \overline{B}(0; M+\varepsilon)} \Big| \Big\langle x, \sum_{n=1}^N p_n x_n^* \Big\rangle \Big| \\ &\leqslant \Big| \sum_{n=1}^N p_n c_n \Big| \leqslant M \Big\| \sum_{n=1}^N p_n x_n^* \Big\|,\end{aligned}$$

where we used (2) in the last inequality. This can only be the case if $\sum_{n=1}^N p_n x_n^* = 0$, which is impossible, since $p \neq 0$ and the x_n^* are linearly independent. This completes the proof with the additional assumption of linear independence of the x_n^*.

For general x_n^*, $n = 1, \ldots, N$, we can first extract a maximal linearly independent sub-collection. Setting some a_n to 0 in (4.21), it is obvious that this sub-collection satisfies Helly's condition as well. The previous part of the proof then gives us an $x \in X$, so that the equations $\langle x, x_n^* \rangle = c_n$ are satisfied for x_n^* in the independent collection chosen. Due to dependence, these conditions already determine $\langle x, x_n^* \rangle$ for the rest of the x_n^*. To see that this yields correct values, express any one of the remaining dependent functionals as $x_n^* = \sum_j a_j x_j^*$, in which case $\langle x, x_n^* \rangle = \sum_j a_j c_j$, and using (2) we obtain

$$\Big| c_n - \sum_j a_j c_j \Big| \leqslant M \Big\| x_n^* - \sum_j a_j x_j^* \Big\| = 0.$$

So $\langle x, x_n^* \rangle = c_n$ is satisfied for all $n = 1, \ldots, N$. □

With the help of Helly's condition we now establish the following characterisations of reflexivity:

Theorem 4.3.2 (James). *For a Banach space X, the following conditions are equivalent:*

(1) *X is not reflexive.*
(2) *there exist a real number $0 < \vartheta < \frac{1}{2}$ and sequences $(x_n)_{n\geqslant 1}$ in $\overline{B}_X$ and $(x_n^*)_{n\geqslant 1}$ in $\overline{B}_{X^*}$, such that*

$$\langle x_j, x_i^* \rangle = \begin{cases} \vartheta & \textit{if } i \leqslant j, \\ 0 & \textit{if } i > j. \end{cases}$$

(3) *there exist a real number $\vartheta > 0$ and a bounded sequence $(x_n)_{n\geqslant 1}$ in X, such that*

$$d(\mathrm{conv}(x_n)_{n=1}^N, \mathrm{conv}(x_n)_{n=N+1}^\infty) \geqslant \vartheta, \tag{4.22}$$

for all $N \geqslant 1$.

Here, the distance between two subsets A and B of X is defined as their separation, i.e., $d(A, B) := \inf_{x\in A, y\in B} \|x - y\|$.

Proof. (1)⇒(2): Suppose 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 $\|x^{**}\| < 1$ and also fix a $\vartheta \in (0, 1)$ such that $d(x^{**}, X) > \vartheta$. Observe that this implies, since $0 \in X$, that $\|x^{**}\| > \vartheta$.

We inductively construct sequences $(x_n)_{n\geqslant 1}$ in $\overline{B}_X$ and $(x_n^*)_{n\geqslant 1}$ in $\overline{B}_{X^*}$ meeting the conditions of (2); the sequence $(x_n^*)_{n\geqslant 1}$ will in addition have the property $\langle x_n^*, x^{**}\rangle = \vartheta$ for all $n \geqslant 1$.

For the initial step, since $\|x^{**}\| > \vartheta$, there is an $x_1^* \in X^*$ with norm at most one such that $|\langle x_1^*, x^{**}\rangle| = \vartheta$; multiplying x_1^* with an appropriate scalar of modulus one if necessary, we obtain the same equality without the absolute values. Since $\|x^{**}\| < 1$, we must then have $\|x_1^*\| > \vartheta$, and we can repeat the argument just given to find $x_1 \in X$ of norm at most one such that $\langle x_1, x_1^*\rangle = \vartheta$. This completes the initial step.

Suppose 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 theorem are satisfied for this range of indices. We now seek an $x_n^* \in X^*$ of norm at most one so that $\langle x_n^*, x^{**}\rangle = \vartheta$ and $\langle x_i, x_n^*\rangle = 0$ for $i = 1, \dots, n-1$, and an $x_n \in X$ of norm at most one so that $\langle x_n, x_j^*\rangle = \vartheta$ for $j = 1, \dots, n$.

The existence of a solution of the former system of equations 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 normalised to unity). Indeed, $\|x^{**} + \sum_{i=1}^{n-1} a_i x_i\| \geqslant d(x^{**}, X) > 0$ implies

$$\Big|\vartheta + \sum_{i=1}^{n} a_i \cdot 0\Big| = \vartheta \leqslant \frac{\vartheta}{d(x^{**}, X)} \Big\|x^{**} + \sum_{i=1}^{n-1} a_i x_i\Big\|$$

for all choices of scalars a_i. Since $M := \vartheta/d(x^{**}, X) < 1$ is a Helly constant, we can choose $\varepsilon > 0$ such that $M+\varepsilon \leqslant 1$ to deduce the existence of x_n^* of norm at most one and such that $\langle x_n^*, x^{**}\rangle = \vartheta$ and $\langle x_i, x_n^*\rangle = 0$ for $i = 1, \dots, n-1$.

We still need to find an $x_n \in X$ of norm at most one which satisfies $\langle x_n, x_j^*\rangle = \vartheta$ for $j = 1, \dots, n$. Using that $\langle x_j^*, x^{**}\rangle = \vartheta$ and $\|x^{**}\| < 1$,

$$\Big|\sum_{j=1}^n a_j\vartheta\Big| = \Big|\Big\langle\sum_{j=1}^n a_j x_j^*, x^{**}\Big\rangle\Big| \leqslant \|x^{**}\|\Big\|\sum_{j=1}^n a_j x_j^*\Big\|,$$

so Helly's condition is satisfied again, with Helly constant $\|x^{**}\| < 1$ and therefore we can find x_n with the required properties. This completes the induction step.

(2)⇒(3): Assume the existence of the sequences $(x_n)_{n\geqslant 1}$ and $(x_n^*)_{n\geqslant 1}$ as in condition (2). Then $(x_n)_{n\geqslant 1}$ is an appropriate sequence also for condition (3), with the same ϑ. Indeed, let $\sum_{n=1}^N \lambda_n = \sum_{n=N+1}^\infty \mu_n = 1$, with $\lambda_n, \mu_n \geqslant 0$ and only finitely many of the μ_n non-zero. Then

$$\begin{aligned}\Big\|\sum_{n=1}^N \lambda_n x_n - \sum_{n=N+1}^\infty \mu_n x_n\Big\| &\geqslant \Big|\Big\langle\sum_{n=1}^N \lambda_n x_n - \sum_{n=N+1}^\infty \mu_n x_n, x_{N+1}^*\Big\rangle\Big| \\ &= \Big|\sum_{n=1}^N \lambda_n\cdot 0 + \sum_{n=N+1}^\infty \mu_n\theta\Big| = \theta,\end{aligned}$$

so $d(\mathrm{conv}(x_n)_{n=1}^N, \mathrm{conv}(x_n)_{n=N+1}^\infty) \geqslant \vartheta$.

(3)⇒(1): Suppose there is a bounded sequence in X with the property (4.22). Then, given any $x \in X$, we claim that, for a large enough $n \geqslant 1$, $d(x, \mathrm{conv}(x_k)_{k=n+1}^\infty) \geqslant \frac12\vartheta$. Indeed, if $d(x, \mathrm{conv}(x_k)_{k=n+1}^\infty) < \frac12\vartheta$, then $\|x-y\| < \frac12\vartheta$ for some finite convex combination y of the x_k, $k > n+1$. Due to the finiteness, we can actually find an $m \geqslant n+1$ such that $y \in \mathrm{conv}(x_k)_{k=1}^m$, and this implies $d(y, \mathrm{conv}(x_k)_{k=m+1}^\infty) \geqslant \vartheta$. But then

$$d(x, \mathrm{conv}(x_k)_{k=m+1}^\infty) \geqslant d(y, \mathrm{conv}(x_k)_{k=m+1}^\infty) - \|x-y\| \geqslant \vartheta - \frac12\vartheta = \frac12\vartheta,$$

and this proves the claim.

Fix for the moment an $x \in X$ and set $C := \overline{\mathrm{conv}}\{x_k\}_{k=n+1}^\infty$, where n is so large that $d(x, C) \geqslant \frac12\vartheta$. The Hahn–Banach theorem implies that there exists a linear functional $x^* \in X^*$ such that $\Re\langle x_k, x^*\rangle < \alpha < \Re\langle x, x^*\rangle$ for all $k \geqslant n+1$ and some $\alpha \in \mathbb{R}$.

Now consider $x^{**} \in X^{**}$ defined by $\Re\langle y^*, x^{**}\rangle := \Lambda(\Re\langle x_k, y^*\rangle)_{k=1}^\infty$, where $\Lambda \in \mathscr{L}(\ell^\infty; \mathbb{R})$ is a Banach limit, i.e., a linear functional on bounded real sequences with the property $\liminf_{k\to\infty}\lambda_k \leqslant \Lambda\lambda \leqslant \limsup_{k\to\infty}\lambda_k$.

Recall that it is legitimate to define a linear functional in terms of its real part, because every real-linear functional is the real part of a unique complex-linear functional. From the definition it is immediate that $\Re\langle y^*, x^{**}\rangle \leqslant$

$\limsup_{k\to\infty} \Re\langle x_k, y^*\rangle$. This property makes it impossible that x^{**} belongs to X, as we will see next.

Given any $x \in X$, let $x^* \in X^*$ be the linear functional with the property $\Re\langle x_k, x^*\rangle < \alpha < \Re\langle x, x^*\rangle$ for all $k \geqslant n+1$ as above. Then clearly

$$\Re\langle x^*, x^{**}\rangle \leqslant \limsup_k \Re\langle x_k, x^*\rangle \leqslant \alpha < \Re\langle x, x^*\rangle,$$

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 therefore X is not reflexive. $\square$

The main result

Theorem 4.3.3. *Every UMD space is reflexive.*

The proof relies on the possibility of constructing certain 'pathological' Paley–Walsh martingales in non-reflexive Banach spaces.

Proposition 4.3.4. *If X is a non-reflexive Banach space X, there exists a constant $\theta > 0$ such that the following holds: For every $N \geqslant 1$ there is an X-valued Paley–Walsh martingale $(f_n)_{n=0}^N$ such that, pointwise almost everywhere,*

$$\|f_n\| \leqslant 1, \quad 0 \leqslant n \leqslant N$$

and

$$\|df_n\| \geqslant \theta, \quad 1 \leqslant n \leqslant N.$$

Proof. By Theorem 4.3.2 there is a number $0 < \theta < \frac{1}{2}$ and a sequence $(x_n)_{n\geqslant 1}$ in the closed unit ball of X such that $d(A_n, B_n) \geqslant 2\theta$ for all $n \geqslant 1$, where

$$A_n = \mathrm{conv}(\{x_i : 1 \leqslant i \leqslant n\}), \quad B_n = \mathrm{conv}(\{x_i : i \geqslant n+1\}).$$

Define $f_N : [0,1) \to X$ by $f_N := \sum_{n=1}^{2^N} \mathbf{1}_{[(n-1)2^{-N}, n2^{-N})} x_n$. Then f_N is strongly measurable with respect to the dyadic filtration $\mathscr{D}_N$ generated by the intervals $I_n^N = [(n-1)2^{-N}, n2^{-N})$, $1 \leqslant n \leqslant 2^N$. Define the martingale $(f_n)_{n=0}^N$ by $f_n = \mathbb{E}(f_N|\mathscr{D}_n)$. On the interval I_j^n, almost surely f_n takes the value

$$y_j^n := \frac{1}{|I_j^n|}\int_{I_j^n} f_N \,\mathrm{d}t = 2^{n-N} \sum_{k=(j-1)2^{N-n}+1}^{j2^{N-n}} x_k$$

and hence $y_j^n \in \mathrm{conv}(\{x_k : (j-1)2^{N-n}+1 \leqslant k \leqslant j2^{N-n})\}$. Therefore, $\|y_j\| \leqslant 1$ and $\|y_i^n - y_j^n\| \geqslant 2\theta$ whenever $i \neq j$ and $1 \leqslant i, j \leqslant 2^n$.

From $I_j^{n-1} = I_{2j-1}^n \cup I_{2j}^n$ we infer that $y_j^{n-1} = \frac{1}{2}(y_{2j-1}^n + y_{2j}^n)$. It follows that on I_{2j-1}^n

$$\|df_n\| = \Big\|y_{2j-1}^n - \frac{1}{2}(y_{2j-1}^n + y_{2j}^n)\Big\| = \frac{1}{2}\|y_{2j-1}^n - y_{2j}^n\| \geqslant \theta$$

and similarly on I_{2j}^n

$$\|df_n\| = \Big\|y_{2j}^n - \frac{1}{2}(y_{2j-1}^n + y_{2j}^n)\Big\| = \frac{1}{2}\|y_{2j}^n - y_{2j-1}^n\| \geqslant \theta.$$

□

Proof of Theorem 4.3.3. Arguing by contradiction, suppose that X is a non-reflexive UMD space. Fix $p \in (1,\infty)$ and an integer $N \geqslant 1$, and choose $(f_n)_{n=0}^N$ and $0 < \theta < \frac{1}{2}$ as in Proposition 4.3.4, and let $(\varepsilon_n)_{n=1}^N$ be a Rademacher sequence on a probability space Ω. By Proposition 4.2.3 and the Kahane contraction principle (Proposition 3.2.10), for all integers $n \in \{1,\dots,N\}$ and scalars $a_1,\dots,a_N$ we have

$$\begin{aligned}
\theta|a_n| \leqslant \|a_n df_n\|_{L^p(0,1;X)} &= \|a_n\varepsilon_n df_n\|_{L^p(\Omega\times[0,1);X)} \\
&\leqslant \Big\|\sum_{n=1}^N a_n\varepsilon_n df_n\Big\|_{L^p(\Omega\times[0,1);X)} \\
&\leqslant \|a\|_{\ell_N^\infty}\Big\|\sum_{n=1}^N \varepsilon_n df_n\Big\|_{L^p(\Omega\times[0,1);X)} \\
&\leqslant \beta_{p,X}^+\|a\|_{\ell_N^\infty}\|f_N\|_{L^p(0,1;X)} \leqslant \beta_{p,X}^+\|a\|_{\ell_N^\infty}.
\end{aligned}$$

This shows that the mapping $J : a \mapsto \sum_{n=1}^N a_n\varepsilon_n df_n$ is an isomorphic embedding of ℓ_N^∞ into $L^q(\Omega\times[0,1);X)$ that satisfies $\theta \leqslant \|J\| \leqslant \beta_{p,X}^+$. Hence, by (the randomised version of) part (1) of Proposition 4.2.17, $\beta_{p,\ell_N^\infty}^+ \leqslant \frac{1}{\theta}\beta_{p,X}^+$. This gives the desired contradiction, since by Proposition 4.2.19 we know that $\lim_{N\to\infty}\beta_{p,\ell_N^\infty}^+ = \infty$. □

By the same method we obtain the following result on martingale type and cotype (see Section 3.5.d).

Proposition 4.3.5. *If X has martingale type $p \in (1,2]$ or martingale cotype $q \in [2,\infty)$, then X is reflexive.*

Proof. First assume X has martingale cotype $q \in [2,\infty)$. If X is non-reflexive, given any $N \geqslant 1$ we choose $(f_n)_{n=0}^N$ and $0 < \theta < \frac{1}{2}$ as in Proposition 4.3.4. Then the martingale cotype q property implies that

$$\theta N^{1/q} \leqslant \Big(\sum_{n=1}^N \|df_n\|_{L^q(S;X)}^q\Big)^{1/q} \leqslant c_{q,X}^{\mathrm{mart}}\|f_N\|_{L^q(S;X)} \leqslant c_{q,X}^{\mathrm{mart}},$$

yielding a contradiction for large enough N.

If X has martingale type $p \in (1,2]$, then by Proposition 3.5.29 the dual space X^* has martingale cotype p' and is therefore reflexive. This implies the reflexivity of X. □

UMD implies super-reflexivity

The equivalence of the UMD property and its dyadic counterpart in terms of Paley–Walsh martingales (Theorem 4.2.5) reveals that the UMD property of a Banach space X is entirely 'encoded' in the structure of the finite-dimensional subspaces of X. Recall that properties that depend only on the structure of the finite-dimensional subspaces of X are often called super properties. In order to make this into a precise definition we introduce the following terminology.

Definition 4.3.6. *A Banach space X is said to be* finitely representable *in the Banach space Y if for every $\varepsilon > 0$ and every finite-dimensional subspace X_0 of X there exists a finite-dimensional subspace Y_0 of Y and a linear isomorphism $T : X_0 \to Y_0$ such that*

$$\|T\|\|T^{-1}\| \leqslant 1 + \varepsilon.$$

Some of the deepest results in the theory of Banach spaces can be phrased in terms of finite representability. Since these will not be needed, we defer their discussion to the Notes at the end of this chapter.

Armed with the above definition, a property (P) which a Banach space may or may not have is now said to be a *super* property if it is inherited under finite representability, i.e., whenever Y has the property (P) and X is finitely representable in Y, then X has the property (P).

Example 4.3.7. The above definition would not have been stated if the UMD property were not a super property. Indeed it is, for the reasons already indicated: the dyadic UMD property is clearly inherited under finite representability. Further examples of super properties will be given in the next subsection, viz. K-convexity and (martingale) type and cotype.

A Banach space X is called *super-reflexive* if every Banach space that is finitely representable in X is reflexive. Trivially, every super-reflexive Banach space X is reflexive (as X is finitely representable in itself), but the converse is not true: denoting by ℓ^∞_N the space $\mathbb{K}^N$ with the supremum norm, a bit of reflection reveals that Proposition 4.6.7 implies that every Banach space X is finitely representable in the reflexive Banach space $\bigoplus_{n\geqslant 1} \ell^\infty_N$, the direct sum being taken in the ℓ^2-sense. It is known that a Banach space is super-reflexive if and only if it admits an equivalent norm which makes it uniformly convex; we refer the reader to the Notes for more on this.

Now if Y is finitely representable in a UMD space X, then by the above discussion Y has UMD and hence is reflexive by Theorem 4.3.3.

Corollary 4.3.8. *Every UMD space is super-reflexive.*

Similarly, by approximating martingales by simple ones, it can be seen that if X has martingale type $p \in (1, 2]$ or martingale cotype $q \in [2, \infty)$, then X is super-reflexive.

4.3.b Further Banach space properties implied by UMD

Besides being reflexive, UMD spaces enjoy many further good properties, such as K-convexity and non-trivial (martingale) type and cotype. These facts will be proved next. We also present an ingenious example, due to Qiu, which shows that, in the converse direction, neither one of these properties implies the UMD property.

K-convexity

Let us start by defining the notion of K-convexity. This is not the place to treat this topic thoroughly, and the brief discussion presented here can do no justice to its many interesting facets. A fuller treatment of K-convexity will be given in Volume II, where we prove Pisier's characterisation of K-convex Banach spaces as being precisely those with non-trivial type (see Definition 4.3.12).

In what follows we fix a real Rademacher sequence $(r_n)_{n\geqslant 1}$ defined on a probability space Ω.

Definition 4.3.9. *A Banach space X is* K-convex *if, for some/all $p \in (1,\infty)$,*

$$K_{p,X} := \sup_{N\geqslant 1} \|\pi_N\|_{\mathscr{L}(L^p(\Omega;X))}$$

is finite, where

$$\pi_N f := \sum_{n=1}^{N} r_n \mathbb{E}(r_n f), \quad f \in L^p(\Omega;X), \quad N \geqslant 1,$$

are the so-called Rademacher projections*.*

The reason for considering real Rademacher variables in this definition will be apparent in a moment: it allows us to connect K-convexity with properties of the Walsh system (see Lemma 4.3.11).

The p-independence of the above definition is a straightforward consequence of the Kahane–Khintchine inequality (Theorem 3.2.23).

In order to calculate the norm of π_N it suffices to consider functions f which are measurable with respect to $\mathscr{F}_N = \sigma(r_1,\dots,r_N)$. Indeed, with $\mathbb{E}_N = \mathbb{E}(\cdot|\mathscr{F}_N)$, one has

$$\mathbb{E}(r_n f) = \mathbb{E}\mathbb{E}_N(r_n f) = \mathbb{E}(r_n \mathbb{E}_N(f))$$

for all $1 \leqslant n \leqslant N$, so $\pi_N f = \pi_N \mathbb{E}_N(f)$ for all $f \in L^p(\Omega;X)$. From this it follows that $\|\pi_N|_{L^p(\mathscr{F}_N;X)}\| = \|\pi_N\|$.

Proposition 4.3.10. *Every UMD space is K-convex, and $K_{p,X} \leqslant \beta^+_{p,X}$ for every $p \in (1,\infty)$.*

For an integer $N \geqslant 1$, which will be fixed, we consider the probability space $\{-1,1\}^N$ with the probability measure that gives mass 2^{-N} to every point $\omega \in \{-1,1\}^N$. The coordinate mappings $r_n(\omega) := \omega_n$ define a real Rademacher sequence on this space. For any subset $\alpha \subseteq \{1,\dots,N\}$ set

$$w_\alpha := \prod_{n\in\alpha} r_n,$$

with the convention $w_\varnothing := \mathbf{1}$. Note that $w_{\{n\}} = r_n$. The family $(w_\alpha)_\alpha$ is called the *Walsh system* in $\{-1,1\}^N$. It is an orthonormal system of 2^N elements for the Hilbert space $L^2(\{-1,1\}^N)$. Since this space has dimension 2^N, it follows that $(w_\alpha)_\alpha$ is an orthonormal basis.

Lemma 4.3.11. *Let X be a Banach space and let $p \in (1,\infty)$. Then for every function $f : \{-1,1\}^N \to X$ we have*

$$f = \sum_{\alpha\subseteq\{1,\dots,N\}} w_\alpha \mathbb{E}(w_\alpha f).$$

Moreover, the Rademacher projection π_N is given by

$$\pi_N(f) = \sum_{\#\alpha=1} w_\alpha \mathbb{E}(w_\alpha f) = \sum_{n=1}^N r_n \mathbb{E}(r_n f).$$

Proof. The first assertion is immediate from the fact that the Walsh system is an orthonormal basis for $L^2(\{-1,1\}^N)$. For the second assertion it suffices to note that $r_n = w_{\{n\}}$. □

Proof of Proposition 4.3.10. It suffices to prove the theorem for real Banach spaces. Indeed, once we have the theorem for real Banach spaces, then for complex Banach spaces X we turn to $X_{\mathbb{R}}$ and obtain

$$K_{p,X} = K_{p,X_{\mathbb{R}}} \leqslant \beta^+_{p,X_{\mathbb{R}}} \leqslant \beta^+_{p,X}$$

(the last inequality being proved in the same way as the inequality $\beta_{p,X_{\mathbb{R}}} \leqslant \beta_{p,X}$ of Proposition 4.2.10).

So let X be a real Banach space. We will show that the Rademacher projection satisfies $\|\pi_N\|_{L^p(\{-1,1\}^N;X)} \leqslant \beta^+_{p,X}$.

For $1 \leqslant n \leqslant N$ let

$$A_n := \big\{\alpha \subseteq \{1,\dots,N\} : \max\alpha = n\big\}$$

and

$$d_n := \sum_{\alpha\in A_n} w_\alpha \mathbb{E}(w_\alpha f).$$

Then $f = \sum_{n=1}^N d_n$ and $\mathbb{E}(d_n|\sigma(r_1,\dots,r_{n-1})) = 0$, so $(d_n)_{n=1}^N$ is a martingale difference sequence. Moreover, for all $n \in \{1,\dots,N\}$ we have $\mathbb{E}(r_n f) = \mathbb{E}(r_n d_n)$, noting that $\mathbb{E}(r_n w_\alpha) = 0$ for all $\alpha \in A_m$ with $m \neq n$.

Let $(r'_n)_{n=1}^N$ be a real Rademacher sequence on another probability space $(\Omega', \mathscr{F}', \mathbb{P}')$. Then, for all $f \in L^p(\{-1,1\}^N; X)$,

$$
\begin{aligned}
\mathbb{E}\|\pi_N(f)\|^p &= \mathbb{E}\Big\| \sum_{n=1}^N r_n \mathbb{E}(r_n f) \Big\|^p = \mathbb{E}'\Big\| \sum_{n=1}^N r'_n \mathbb{E}(r_n d_n) \Big\|^p \\
&= \mathbb{E}'\Big\| \mathbb{E} \sum_{n=1}^N r'_n r_n d_n \Big\|^p \leqslant \mathbb{E}'\mathbb{E}\Big\| \sum_{n=1}^N r'_n r_n d_n \Big\|^p \\
&\overset{\text{(i)}}{=} \mathbb{E}\mathbb{E}'\Big\| \sum_{n=1}^N r'_n d_n \Big\|^p \overset{\text{(ii)}}{\leqslant} (\beta^+_{p,X})^p \mathbb{E}\Big\| \sum_{n=1}^N d_n \Big\|^p = (\beta^+_{p,X})^p \mathbb{E}\|f\|^p.
\end{aligned}
$$

In (i) we apply Fubini's theorem to interchange the order of integration and use that $(r_n(\omega) r'_n)_{n=1}^N$ is a Rademacher sequence on Ω', and (ii) follows from Proposition 4.2.3. □

Type and cotype

As with K-convexity, the notions of type and cotype are also about the behaviour of Rademacher sequences, and the precise definition reads as follows:

Definition 4.3.12. *Let X be a Banach space and let $(\varepsilon_n)_{n\geqslant 1}$ be a Rademacher sequence.*

(1) *The space X is said to have* type $p \in [1,2]$ *if there exists a constant $\tau \geqslant 0$ such that for all finite sequences $x_1, \dots, x_N$ in X we have*

$$
\Big(\mathbb{E}\Big\| \sum_{n=1}^N \varepsilon_n x_n \Big\|^p\Big)^{1/p} \leqslant \tau \Big(\sum_{n=1}^N \|x_n\|^p \Big)^{1/p}.
$$

(2) *The space X is said to have* cotype $q \in [2,\infty]$ *if there exists a constant $c \geqslant 0$ such that for all finite sequences $x_1, \dots, x_N$ in X we have*

$$
\Big(\sum_{n=1}^N \|x_n\|^q \Big)^{1/q} \leqslant c \Big(\mathbb{E}\Big\| \sum_{n=1}^N \varepsilon_n x_n \Big\|^q\Big)^{1/q},
$$

with the obvious modification for $q = \infty$.

The least admissible constants in (1) and (2) are called the *type p constant* and *cotype q constant* of X and will be denoted by $\tau_{p,X}$ and $c_{q,X}$. From the case $N = 1$ we see that $\tau_{p,X} \geqslant 1$ and $c_{q,X} \geqslant 1$.

By Pisier's theorem quoted in the previous paragraph, K-convex Banach spaces have non-trivial type. In turn, non-trivial type implies finite cotype. As a consequence, UMD spaces enjoy both properties. These results will be proved in the next Volume.

Martingale type and martingale cotype

We recall from Section 3.5.d that a Banach space X has *martingale type* $p \in [1,2]$ if there exists a constant $\tau \geqslant 0$ such that for all finite L^p-martingales $(f_n)_{n=0}^N$,

$$\|f_N\|_{L^p(S;X)} \leqslant \tau\Big(\|f_0\|_{L^p(S;X)}^p + \sum_{n=1}^N \|df_n\|_{L^p(S;X)}^p\Big)^{1/p}.$$

The space X has *martingale cotype* $q \in [2,\infty]$ if there exists a constant $c \geqslant 0$ such that for all finite L^q-martingales $(f_n)_{n=0}^N$,

$$\Big(\|f_0\|_{L^q(S;X)}^q + \sum_{n=1}^N \|df_k\|_{L^q(S;X)}^q\Big)^{1/q} \leqslant c\|f_N\|_{L^q(S;X)}$$

with the obvious modification for $q=\infty$. The least admissible constants are denoted by $\tau_{p,X}^{\rm mart}$ and $c_{q,X}^{\rm mart}$ respectively.

By considering Rademacher differences $df_n = \varepsilon_n x_n$ one sees that martingale type p (martingale cotype q) implies type p (cotype q) and

$$\tau_{p,X} \leqslant \tau_{p,X}^{\rm mart}, \quad c_{q,X} \leqslant c_{q,X}^{\rm mart}.$$

The converse holds for UMD spaces:

Proposition 4.3.13. *Let X be a UMD space and $p \in (1,2]$ and $q \in [2,\infty)$.*

(1) *If X has type p, then X has martingale type p and $\tau_{p,X}^{\rm mart} \leqslant \beta_{p,X}^{-}\tau_{p,X}$.*
(2) *If X has cotype q, then X has martingale cotype q and $c_{q,X}^{\rm mart} \leqslant \beta_{q,X}^{+}c_{q,X}$.*

Here $\beta_{s,X}^{\pm}$ are the constants introduced in Proposition 4.2.3.

Recall that $\beta_{s,X}^{\pm} \leqslant \beta_{s,X}$ for all $s \in (1,\infty)$.

Proof. By Proposition 4.2.3, (the text below) Lemma 4.2.8, and the triangle inequality in $L^p(\Omega \times S;X)$,

$$\begin{aligned}
\|f_N\|_{L^p(S;X)} &= \Big\|f_0 + \sum_{n=1}^N df_n\Big\|_{L^p(S;X)} \\
&\leqslant \beta_{p,X}^{-}\Big(\int_S \Big\|\varepsilon_0 f_0 + \sum_{n=1}^N \varepsilon_n df_n\Big\|_{L^p(\Omega;X)}^p \,\mathrm{d}\mu\Big)^{1/p} \\
&\leqslant \beta_{p,X}^{-}\tau_{p,X}\Big(\int_S \|f_0\|^p + \sum_{n=1}^N \|df_n\|^p \,\mathrm{d}\mu\Big)^{1/p},
\end{aligned}$$

which proves (1). The proof of (2) is similar. □

In the previous paragraph we stated without proof the result that UMD spaces have non-trivial type and finite cotype. In view of Proposition 4.3.13, it follows that UMD spaces also have non-trivial martingale type and finite martingale cotype.

4.3.c Qiu's example

In this section we will present an ingenious example due to Qiu of a non-UMD Banach function space which is reflexive and has non-trivial type and finite cotype. In order to explain the result we fix $p, q \in [1, \infty]$ with $p \neq q$ and define the spaces $(X_k)_{k \geqslant 1}$ inductively by

$$X_1 := \ell_2^p(\ell_2^q), \quad \text{and} \quad X_{k+1} := \ell_2^p(\ell_2^q(X_k)), \quad k \geqslant 1.$$

It will be shown that the UMD constant of X_k satisfies $\beta_{r,X_k} \geqslant c(p,q)^k$ for some constant $c(p,q) > 1$ and all $r \in (1, \infty)$. A suitable direct sum of the spaces X_k will then produce the example.

A projection on $L^p(\Omega; \ell_2^q)$

Proposition 4.3.14. *Let* $\Omega = \{-1, 1\}$ *with uniform probability distribution* $\mathbb{P}$ *and let* $p, q \in [1, \infty]$ *satisfy* $p \neq q$*. Let* P *be the projection on* $L^p(\Omega; \ell_2^q)$ *defined by* $P(f_1, f_2) = (\mathbb{E}f_1, f_2)$*. Then* $\|P\| > 1$*.*

Note that if $p = q$, then $\|P\| = 1$ by convexity.

Proof. The adjoint of P on $(L^p(\Omega; \ell_2^q))^* = L^{p'}(\Omega; \ell_2^{q'})$ is again given by $P^*(f_1, f_2) = (\mathbb{E}f_1, f_2)$, and therefore there is no loss of generality in assuming that $q < p$.

First we consider the case $p < \infty$. Set $f_1(\omega) := 1 + \omega$ $(\omega \in \{-1, 1\})$ and, for an arbitrary but fixed $t > 0$, set

$$f_2(\omega) = \begin{cases} 0, & \text{if } \omega = 1, \\ t^{1/q}, & \text{if } \omega = -1. \end{cases}$$

Then $\mathbb{E}f_1 = 1$ and therefore

$$\|(\mathbb{E}f_1, f_2)\|^p_{L^p(\Omega;\ell_2^q)} = \frac{1}{2}\big(1 + (1+t)^{p/q}\big), \quad \|(f_1, f_2)\|^p_{L^p(\Omega;\ell_2^q)} = \frac{1}{2}(2^p + t^{p/q}).$$

Setting $r = p/q \in (1, \infty)$, it follows that

$$\|P\|^p \geqslant \phi(t) := \frac{1 + (1+t)^r}{2^p + t^r}, \quad t > 0.$$

Taking $t = 2^{p/(r-1)}$, we find that

$$\|P\|^p \geqslant \frac{1 + \big(1 + 2^{p/(r-1)}\big)^r}{2^p + 2^{pr/(r-1)}} > \frac{\big(1 + 2^{p/(r-1)}\big)^r}{2^p + 2^{pr/(r-1)}} = \frac{\big(1 + 2^{p/(r-1)}\big)^{r-1}}{2^p} > 1.$$

Next we consider the case $p = \infty$. Taking limits we obtain

$$\|(\mathbb{E}f_1, f_2)\|_{L^\infty(\Omega;\ell_2^q)} = \lim_{p\to\infty} 2^{-1/p}\big(1 + (1+t)^{p/q}\big)^{1/p} = (1+t)^{1/q},$$

$$\|(f_1, f_2)\|_{L^\infty(\Omega;\ell_2^q)} = \lim_{p\to\infty} 2^{-1/p}(2^p + t^{p/q})^{1/p} = \max(2, t^{1/q}).$$

Taking $t = 2^q$, we find that $\|P\| \geqslant \frac{1}{2}(1 + 2^q)^{1/q} > 1$. □

Let us denote the norm of P on $L^p(\Omega;\ell^q_2)$ by $c(p,q)$. By an optimisation argument using general functions f_1 and f_2, one shows that for $q\in[1,\infty)$,

$$c(\infty,q)=c(1,q')=\Big(\frac{1}{(2^{q/(q-1)}-1)^{q-1}}+1\Big)^{1/q}.$$

Letting $q\to\infty$, this yields $c(\infty,1)=c(1,\infty)=\frac{3}{2}$.

For $p>q$, Hölder's and Jensen's inequalities give the upper bound $c(p,q)\leqslant 2^{\frac{1}{q}-\frac{1}{p}}$, but this seems to be far from optimal. It is an interesting problem to find an exact expression for $c(p,q)$ in general.

A variant of Stein's inequality

A sequence $(x_n)_{n=1}^N$ in a Banach space X will be called 1-*unconditional* if for all scalar sequences $(a_n)_{n=1}^N$ and $(b_n)_{n=1}^N$ satisfying $|a_n|\leqslant|b_n|$ for all $n\in\{1,\dots,N\}$, we have

$$\Big\|\sum_{n=1}^N a_nx_n\Big\|\leqslant\Big\|\sum_{n=1}^N b_nx_n\Big\|.$$

In this situation, if $(\varepsilon_n)_{n\geqslant1}$ is a Rademacher sequence, then for any $p\in[1,\infty]$,

$$\Big\|\sum_{n=1}^N\varepsilon_na_nx_n\Big\|_{L^p(\Omega;X)}=\Big\|\sum_{n=1}^N a_nx_n\Big\|_{L^p(\Omega;X)}. \tag{4.23}$$

For a sequence $x=(x_n)_{n=1}^N$ in X, let $S((x_n)_{n=1}^N)$ denote the least constant $C\in[0,\infty]$ such that whenever $(\Omega,\mathscr{A},\mathbb{P})$ is a probability space with filtration $(\mathscr{F}_n)_{n\geqslant1}$ and $(\phi_n)_{n=1}^N$ is a sequence in $L^\infty(\Omega;X)$, one has

$$\Big\|\sum_{n=1}^N\mathbb{E}(\phi_n|\mathscr{F}_n)x_n\Big\|_{L^1(\Omega;X)}\leqslant C\Big\|\sum_{n=1}^N\phi_nx_n\Big\|_{L^\infty(\Omega;X)},$$

which, in view of (4.23), is a Stein-type inequality.

The supremum of $S((x_n)_{n=1}^N)$ over all finite sequences $(x_n)_{n=1}^N$ in X will be denoted by S_X. Since $L^\infty(\Omega)\hookrightarrow L^p(\Omega)\hookrightarrow L^1(\Omega)$ with continuous embeddings of norm 1, in view of (4.23) it follows from Proposition 4.2.3 and Theorem 4.2.23 that for UMD spaces X we have

$$S_X\leqslant\beta^+_{p,X}\leqslant\beta_{p,X}. \tag{4.24}$$

Thus in order to prove that a given Banach space X does not have the UMD property it suffices to verify that $S_X=\infty$. The main tool for this is the following lemma.

Theorem 4.3.15. *Let X be an M-dimensional Banach space with a 1-unconditional basis $(x_m)_{m=1}^M$. Let Y be another Banach space. Define a norm on $X \otimes Y$ by*

$$\Big\| \sum_{m=1}^{M} x_m \otimes y_m \Big\|_{X\otimes Y} = \Big\| \sum_{m=1}^{M} \|y_m\| \, x_m \Big\|.$$

Then for every finite sequence $(y_n)_{n=1}^N$ in Y we have

$$S\big((x_m \otimes y_n)_{m,n=1}^{M,N}\big) \geqslant S\big((x_m)_{m=1}^M\big) S\big((y_n)_{n=1}^N\big),$$

the left-hand side being evaluated in terms of the norm on $X \otimes Y$ just defined.

Proof. Let $\delta > 0$ be arbitrary and fixed. Choose a probability space $(\Omega, \mathscr{A}, \mathbb{P})$ with filtration $(\mathscr{A}_m)_{m=1}^M$ and a finite sequence $(\phi_m)_{m=1}^M$ in $L^\infty(\Omega)$ such that $\| \sum_{m=1}^M \phi_m x_m \|_{L^\infty(\Omega;X)} \leqslant 1$ and

$$\Big\| \sum_{m=1}^{M} \mathbb{E}(\phi_m|\mathscr{A}_m) x_m \Big\|_{L^1(\Omega;X)} \geqslant S((x_m)_{m=1}^M) - \delta.$$

Similarly, choose a probability space $(T, \mathscr{B}, \nu)$ with filtration $(\mathscr{B}_n)_{n=1}^N$ and a finite sequence $(\psi_n)_{n=1}^N$ in $L^\infty(T)$ such that $\| \sum_{n=1}^N \psi_n y_n \|_{L^\infty(T;X)} \leqslant 1$ and

$$\Big\| \sum_{n=1}^{N} \mathbb{E}(\psi_n|\mathscr{B}_n) y_n \Big\|_{L^1(T;X)} \geqslant S((y_n)_{n=1}^N) - \delta.$$

On the product space $\Omega \times T^M$ we define the product σ-algebras $\mathscr{C}_{m,n}$ by

$$\mathscr{C}_{m,n} = \mathscr{A}_m \times \mathscr{B}^{m-1} \times \mathscr{B}_n \times \{\varnothing, T\}^{M-m},$$

where we use the lexicographical order

$$(m,n) \leqslant (m',n') \iff [m < m' \text{ or } (m = m' \text{ and } n \leqslant n')].$$

In the sequel we write $\mathbf{t} = (t_1, \dots, t_M)$. Let $f_{m,n}(\omega, \mathbf{t}) := \phi_m(\omega)\psi_n(t_m)$ for $(\omega, \mathbf{t}) \in \Omega \times T^M$. Then

$$\mathbb{E}(f_{m,n}|\mathscr{C}_{m,n})(\omega, \mathbf{t}) = \mathbb{E}(\phi_m|\mathscr{A}_m)(\omega)\mathbb{E}(\psi_n|\mathscr{B}_n)(t_m).$$

Define the functions f and $\widetilde{f}$ by

$$f(\omega, \mathbf{t}) = \sum_{m=1}^{M} \sum_{n=1}^{N} f_{m,n}(\omega, \mathbf{t}) x_m \otimes y_n,$$

$$\widetilde{f}(\omega, \mathbf{t}) = \sum_{m=1}^{M} \sum_{n=1}^{N} \mathbb{E}(f_{m,n}|\mathscr{C}_{m,n})(\omega, \mathbf{t}) x_m \otimes y_n.$$

Then

$$\begin{aligned}\|f(\omega,\mathbf{t})\|_{X\otimes Y} &= \Big\|\sum_{m=1}^{M}\phi_m(\omega)x_m\otimes\sum_{n=1}^{N}\psi_n(t_m)y_n\Big\|_{X\otimes Y}\\ &= \Big\|\sum_{m=1}^{M}|\phi_m(\omega)|x_m\Big\|\sum_{n=1}^{N}\psi_n(t_m)y_n\Big\|\Big\|\\ &\overset{(*)}{\leqslant}\Big\|\sum_{m=1}^{M}|\phi_m(\omega)|x_m\Big\| = \Big\|\sum_{m=1}^{M}\phi_m(\omega)x_m\Big\|\leqslant 1,\end{aligned}$$

where $(*)$ uses $\|\sum_{n=1}^N\psi_n y_n\|_{L^\infty(T;X)}\leqslant 1$ and the 1-unconditionality of $(x_m)_{m=1}^M$. We have thus proved that $\|f\|_{L^\infty(\Omega\times T^M;X\otimes Y)}\leqslant 1$.

Similarly,

$$\begin{aligned}&\int_{T^M}\|\widetilde{f}(\omega,\mathbf{t})\|_{X\otimes Y}\,\mathrm{d}\nu^M(\mathbf{t})\\ &=\int_{T^M}\Big\|\sum_{m=1}^{M}\big|\mathbb{E}(\phi_m|\mathscr{A}_m)(\omega)\big|x_m\Big\|\sum_{n=1}^{N}\mathbb{E}(\psi_n|\mathscr{B}_n)(t_m)y_n\Big\|\Big\|\,\mathrm{d}\nu^M(\mathbf{t})\\ &\geqslant\Big\|\int_{T^M}\sum_{m=1}^{M}\big|\mathbb{E}(\phi_m|\mathscr{A}_m)(\omega)\big|x_m\Big\|\sum_{n=1}^{N}\mathbb{E}(\psi_n|\mathscr{B}_n)(t_m)y_n\Big\|\,\mathrm{d}\nu^M(\mathbf{t})\Big\|\\ &=\Big\|\sum_{m=1}^{M}\mathbb{E}(\phi_m|\mathscr{A}_m)(\omega)x_m\Big\|\cdot\Big\|\sum_{n=1}^{N}\mathbb{E}(\psi_n|\mathscr{B}_n)y_n\Big\|_{L^1(T;Y)}.\end{aligned}$$

Integrating over Ω we find that

$$\begin{aligned}&S\big((x_m\otimes y_n)_{m,n=1}^{M,N}\big)\\ &\geqslant\|\widetilde{f}\|_{L^1(\Omega\times T^M;X\otimes Y)}\\ &\geqslant\Big\|\sum_{m=1}^{M}\mathbb{E}(\phi_m|\mathscr{A}_m)x_m\Big\|_{L^1(\Omega;X)}\cdot\Big\|\sum_{n=1}^{N}\mathbb{E}(\psi_n|\mathscr{B}_n)y_n\Big\|_{L^1(T;Y)}\\ &\geqslant\big(S((x_m)_{m=1}^M)-\delta\big)\big(S((y_n)_{n=1}^N)-\delta\big).\end{aligned}$$

Since $\delta>0$ was arbitrary, the required result follows. □

Iterated $\ell_2^p(\ell_2^q)$-spaces

We begin by estimating the constant S_X defined above for $X=\ell_2^p(\ell_2^q)$. Let u_1,u_2 and v_1,v_2 be the standard bases for ℓ_2^p and ℓ_2^q, respectively. Observe that the sequence $(u_m\otimes v_n)_{m,n=1}^{2,2}$ is a 1-unconditional basis of X.

Lemma 4.3.16. *Let $p,q\in[1,\infty]$ and set $X_1=\ell_2^p(\ell_2^q)$. Then*

$$S\big((u_m\otimes v_n)_{m,n=1}^{2,2}\big)\geqslant c(p,q)>1,$$

where $c(p,q)$ is the norm of the projection of Proposition 4.3.14.

Proof. Let $\Omega = \{-1, 1\}$, and define $T : L^\infty(\Omega; X_1) \to L^1(\Omega; X_1)$ by

$$T(\phi \otimes (u_m \otimes v_n))(\omega) = \begin{cases} \mathbb{E}(\phi) u_m \otimes v_n, & \text{if } n = 1, \\ \phi(\omega) u_m \otimes v_n, & \text{if } n = 2, \end{cases}$$

for $\phi \in L^\infty(\Omega)$ and $m = 1, 2$. Let $\phi_1, \phi_2 \in L^\infty(\Omega)$ be fixed and let $f \in L^\infty(\Omega; X_1)$ be given by

$$f(\omega) = \sum_{m=1}^{2} \sum_{n=1}^{2} \phi_n((-1)^{m-1}\omega) u_m \otimes v_n.$$

Applying T we find

$$\begin{aligned} Tf(\omega) &= \mathbb{E}(\phi_1) u_1 \otimes v_1 + \phi_2(\omega) u_1 \otimes v_2 + \mathbb{E}(\phi_1) u_2 \otimes v_1 + \phi_2(-\omega) u_2 \otimes v_2 \\ &= \sum_{m=1}^{2} \sum_{n=1}^{2} \mathbb{E}(\phi_n((-1)^{m-1} \cdot) | \mathscr{A}_{m,n})(\omega) u_m \otimes v_n, \end{aligned}$$

where $\mathscr{A}_{1,1} = \mathscr{A}_{2,1} = \{\varnothing, \Omega\}$ and $\mathscr{A}_{1,2} = \mathscr{A}_{2,2} = \{\varnothing, \{-1\}, \{1\}, \Omega\}$.

The second identity for T implies $S\big((u_m \otimes v_n)_{m,n=1}^{2,2}\big) \geqslant \|T\|$. In order to estimate $\|T\|$, note that for each $\omega \in \Omega$,

$$\begin{aligned} \|f(\omega)\|_{X_1} &= \Big[\big(|\phi_1(\omega)|^q + |\phi_2(\omega)|^q\big)^{p/q} + \big(|\phi_1(-\omega)|^q + |\phi_2(-\omega)|^q\big)^{p/q}\Big]^{1/p} \\ &= 2^{1/p}\Big[\mathbb{E}\big[(|\phi_1|^q + |\phi_2|^q)^{p/q}\big]\Big]^{1/p} = 2^{1/p} \|(\phi_1, \phi_2)\|_{L^p(\Omega; \ell_2^q)}, \\ \|Tf(\omega)\|_{X_1} &= \Big[\big(|\mathbb{E}\phi_1|^q + |\phi_2(\omega)|^q\big)^{p/q} + \big(|\mathbb{E}\phi_1|^q + |\phi_2(-\omega)|^q\big)^{p/q}\Big]^{1/p} \\ &= 2^{1/p}\Big[\mathbb{E}\big[(|\mathbb{E}\phi_1|^q + |\phi_2|^q)^{p/q}\big]\Big]^{1/p} = 2^{1/p} \|(\mathbb{E}\phi_1, \phi_2)\|_{L^p(\Omega; \ell_2^q)}. \end{aligned}$$

This being true for all $\omega \in \Omega$ it follows that

$$\begin{aligned} \|f\|_{L^\infty(\Omega; X_1)} &= 2^{1/p} \|(\phi_1, \phi_2)\|_{L^p(\Omega; \ell_2^q)}, \\ \|Tf\|_{L^1(\Omega; X_1)} &= 2^{1/p} \|(\mathbb{E}\phi_1, \phi_2)\|_{L^p(\Omega; \ell_2^q)}. \end{aligned}$$

Therefore,

$$\|T\| \geqslant \frac{\|Tf\|_{L^1(\Omega; X_1)}}{\|f\|_{L^\infty(\Omega; X_1)}} = \frac{\|(\mathbb{E}\phi_1, \phi_2)\|_{L^p(\Omega; \ell_2^q)}}{\|(\phi_1, \phi_2)\|_{L^p(\Omega; \ell_2^q)}}$$

and from Proposition 4.3.14 we infer that $\|T\| \geqslant c(p, q) > 1$. □

We are now in a position to prove the main result of this section. Recall that $\beta_{r,X}$ is the UMD constant of X and $\beta^+_{r,X}$ is the 'upper' randomised UMD constant of X from Proposition 4.2.3.

Theorem 4.3.17 (Qiu). *Let $r \in (1,\infty)$. Let $p, q \in (1,\infty)$ with $p \neq q$. Let the spaces $(X_k)_{k\geqslant 1}$ be defined by*

$$X_1 = \ell_2^p(\ell_2^q), \quad \textit{and} \quad X_{k+1} = \ell_2^p(\ell_2^q(X_k)) \quad (k \geqslant 1).$$

Then there is a constant $c(p,q) > 1$ such that

$$\beta_{r,X_k} \geqslant \beta^+_{r,X_k} \geqslant S_{X_k} \geqslant c(p,q)^k.$$

Therefore, the direct sum space $X := \ell^p((X_k)_{k\geqslant 1})$ consisting of all sequences $(x_k)_{k\geqslant 1}$ with $x_k \in X_k$ such that

$$\|(x_k)_{k\geqslant 1}\|_X := \Big(\sum_{k\geqslant 1} \|x_k\|^p_{X_k}\Big)^{1/p} < \infty$$

does not have the UMD property. On the other hand, X is reflexive and has type $p \wedge q \wedge 2$ and cotype $p \vee q \vee 2$.

Proof. The estimate $\beta_{r,X_k} \geqslant \beta^+_{r,X_k} \geqslant S_{X_k}$ has already been observed in (4.24). By Lemma 4.3.16 we have $S_{X_1} \geqslant c(p,q)$. Let $(u_m \otimes v_n)_{m,n=1}^{2,2}$ denote the standard basis for X_1. Now by Theorem 4.3.15 we inductively find that

$$S_{X_k} \geqslant S\big((u_m \otimes v_n)_{m,n=1}^{2,2}\big) S_{X_{k-1}} \geqslant c(p,q)c(p,q)^{k-1} = c(p,q)^k.$$

This proves the first assertion.

It is standard to check that the ℓ^p-direct sum of reflexive Banach spaces is reflexive. To prove the assertions concerning type and cotype it is enough to prove that each X_k has type $p \wedge q \wedge 2$ and cotype $p \vee q \vee 2$.

Set $\alpha := p \wedge q \wedge 2$ and $t := p \vee q$. We will check that X has type α with a uniform estimate. In order to do so let $x_1, \dots, x_N \in X_k$. By the triangle inequality, applied repeatedly in $L^{t/p}(\Omega)$ and in $L^{t/q}(\Omega)$, we have

$$\Big\|\sum_{n=1}^N \varepsilon_n x_n\Big\|_{L^t(\Omega;X_k)} \leqslant \Big\|\Big\|\sum_{n=1}^N \varepsilon_n x_n\Big\|_{L^t(\Omega)}\Big\|_{X_k}. \tag{4.25}$$

Applying Khintchine's inequality (see Corollary 3.2.24) to the inner $L^t(\Omega)$-norm yields that

$$\Big\|\sum_{n=1}^N \varepsilon_n x_n\Big\|_{L^t(\Omega)} \leqslant \kappa_{t,2}\Big(\sum_{n=1}^N |x_n|^2\Big)^{1/2} \leqslant \kappa_{t,2}\Big(\sum_{n=1}^N |x_n|^\alpha\Big)^{1/\alpha}.$$

Substituting this into (4.25) and applying the triangle inequality similarly as before, we obtain

$$\Big\|\sum_{n=1}^N \varepsilon_n x_n\Big\|_{L^t(\Omega;X_k)} \leqslant \kappa_{t,2}\Big(\sum_{n=1}^N \|x_n\|^\alpha_{X_k}\Big)^{1/\alpha}.$$

Since the $L^t(\Omega;X_k)$-norm on the left-hand side dominates the $L^\alpha(\Omega;X_k)$-norm, we find that $\tau_{\alpha,X_k} \leqslant \kappa_{t,2}$. Passing to the ℓ^p-direct sum, another triangle inequality argument shows that X has type α.

The assertion concerning the cotype of X is proved in the same way. □

4.4 Decoupling and tangency

Decoupling is a general technique to introduce independence in situations featuring several dependent random variables. The idea is to replace the original objects by new ones in such a way that each of the new objects is 'similar' to its original (for instance, it is equidistributed with its original) and the new objects enjoy a larger degree of independence. A 'decoupling estimate' is then an *a priori* estimate of an expression involving the original objects in terms of the analogous expression involving the new 'decoupled' ones. As we will see, the UMD property of a Banach space X can be used to prove a class of powerful decoupling estimates for sums of X-valued random variables.

4.4.a Elementary decoupling

Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space with a filtration $\mathscr{F} = (\mathscr{F}_n)_{n\geqslant 0}$. With slight abuse of notation, on the product space $(\Omega\times\Omega, \mathscr{A}\times\mathscr{A}, \mathbb{P}\times\mathbb{P})$ we shall consider the filtrations $\mathscr{F} = (\mathscr{F}_n)_{n\geqslant 0}$ and $\widetilde{\mathscr{F}} = (\widetilde{\mathscr{F}}_n)_{n\geqslant 0}$ given by

$$\mathscr{F}_n \times \{\varnothing, \Omega\}, \quad \widetilde{\mathscr{F}}_n = \{\varnothing, \Omega\} \times \mathscr{F}_n.$$

We shall always identify the filtrations $\mathscr{F}$ on Ω and $\Omega\times\Omega$ and think of $\widetilde{\mathscr{F}}$ as an independent copy of $\mathscr{F}$ on $\Omega\times\Omega$. Likewise, given a random variable $\xi : \Omega \to X$, we consider the random variables

$$\xi(\omega, \widetilde{\omega}) := \xi(\omega), \quad \widetilde{\xi}(\omega, \widetilde{\omega}) := \xi(\widetilde{\omega})$$

for $(\omega, \widetilde{\omega}) \in \Omega\times\Omega$. We shall identify the random variables ξ on Ω and $\Omega\times\Omega$ and think of $\widetilde{\xi}$ as an independent copy of ξ on $\Omega\times\Omega$.

Recall that an $\mathscr{L}(X,Y)$-valued sequence $(v_n)_{n\geqslant 1}$ is said to be *strongly predictable* if for the mapping $s \mapsto v_n(s)x$ is strongly $\mathscr{F}_{n-1}$-measurable all for all $n \geqslant 1$ and $x \in X$. In this situation we say that $(v_n)_{n=1}^N$ belongs to $L^\infty_{\rm so} := L^\infty_{\rm so}(S; \mathscr{L}(X,Y))$ if

$$\|v_n\|_{L^\infty_{\rm so}} := \sup_{\|x\|\leqslant 1} \|v_n x\|_{L^\infty(S;Y)} < \infty, \quad \forall n = 1, \dots, N.$$

(see also Section 3.5).

Theorem 4.4.1 (Decoupling). *Let X be a Banach space, Y be a UMD space, and let $p \in (1,\infty)$. Let $(\xi_n)_{n=1}^N$ be an adapted sequence of mean zero random variables in $L^p(\Omega; X)$ and let $(\widetilde{\xi}_n)_{n=1}^N$ be the independent copy of this sequence constructed as above. Let $(v_n)_{n=1}^N$ be a strongly predictable sequence in $L^\infty_{\rm so}(\Omega; \mathscr{L}(X,Y))$. If, for all $n \in \{1, \dots, N\}$, ξ_n is independent of $\mathscr{F}_{n-1}$, then*

$$\frac{1}{\beta_{p,Y}} \Big\| \sum_{n=1}^N v_n \widetilde{\xi}_n \Big\|_{L^p(\Omega\times\Omega;Y)} \leqslant \Big\| \sum_{n=1}^N v_n \xi_n \Big\|_{L^p(\Omega;Y)} \leqslant \beta_{p,Y} \Big\| \sum_{n=1}^N v_n \widetilde{\xi}_n \Big\|_{L^p(\Omega\times\Omega;Y)}. \tag{4.26}$$

Note that we did not impose any conditions on X.

In typical applications, the random variables ξ_n are increments of a (discrete or continuous time) martingale starting at 0 with independent increments.

Remark 4.4.2. The UMD property of Y is a necessary for the conclusion of Theorem 4.4.1 to hold for any given non-trivial Banach space X. Indeed, fixing a one-dimensional subspace of X, we may assume that $X = \mathbb{K} = \mathbb{R}$, and then $\mathscr{L}(X,Y) = \mathscr{L}(\mathbb{R},Y) \eqsim Y$. Choosing $\xi_n = r_n$ and $v_n = \phi_n(r_1,\dots,r_{n-1})$, where $(r_n)_{n=1}^N$ is a real Rademacher sequence, (4.26) says that

$$\begin{aligned}\Big\|\sum_{n=1}^N r_n\phi_n(r_1,\dots,r_{n-1})\Big\|_{L^p(\Omega;Y)} &\eqsim \Big\|\sum_{n=1}^N \widetilde{r}_n\phi_n(r_1,\dots,r_{n-1})\Big\|_{L^p(\Omega\times\Omega;Y)}\\ &= \Big\|\sum_{n=1}^N r_n\widetilde{r}_n\phi_n(r_1,\dots,r_{n-1})\Big\|_{L^p(\Omega\times\Omega;Y)},\end{aligned}$$

where the equality follows from a simple equidistribution argument. But the equivalence of the left and right sides is the randomised version of the (dyadic, via Proposition 3.1.10) UMD property, and therefore equivalent to UMD by Proposition 4.2.3.

Proof of Theorem 4.4.1. As explained above, the random variables $\xi_n : \Omega \to X$ and $\widetilde{\xi}_n : \Omega \to X$ will be interpreted as functions on $\Omega\times\widetilde{\Omega}$ by considering $(\omega,\widetilde{\omega}) \mapsto \xi_n(\omega)$ and $(\omega,\widetilde{\omega}) \mapsto \xi_n(\widetilde{\omega})$, respectively. In the proof we shall work with representatives of ξ_n which are strongly $\mathscr{F}_n$-measurable.

For $n = 1,\dots,N$ define

$$d_{2n-1} := \frac{1}{2}v_n(\xi_n + \widetilde{\xi}_n), \quad d_{2n} := \frac{1}{2}v_n(\xi_n - \widetilde{\xi}_n).$$

We claim that $d = (d_n)_{n=1}^{2N}$ is an L^p-martingale difference sequence with respect to the filtration $\mathscr{G} = (\mathscr{G}_n)_{n=0}^{2N}$, where

$$\mathscr{G}_{2n-1} := \sigma(\mathscr{F}_{n-1}\times\widetilde{\mathscr{F}}_{n-1}, \xi_n + \widetilde{\xi}_n), \quad n \in \{1,\dots,N\}$$

and $\mathscr{G}_{2n} := \mathscr{F}_n \times \widetilde{\mathscr{F}}_n$ for $n \in \{0,\dots,N\}$. Clearly, d is adapted with respect to $\mathscr{G}$. For $n = 1,\dots,N$,

$$\begin{aligned}\mathbb{E}(d_{2n-1}|\mathscr{G}_{2n-2}) &= \frac{1}{2}v_n\mathbb{E}(\xi_n + \widetilde{\xi}_n|\mathscr{G}_{2n-2})\\ &= \frac{1}{2}v_n(\mathbb{E}(\xi_n|\mathscr{F}_{n-1}) + \mathbb{E}(\widetilde{\xi}_n|\widetilde{\mathscr{F}}_{n-1}))\\ &= \frac{1}{2}v_n(\mathbb{E}\xi_n + \mathbb{E}\widetilde{\xi}_n) = 0\end{aligned}$$

by Lemma 3.5.2, the independence and mean zero assumptions. Also, by Lemmas 3.5.2 and 2.6.37, we conclude that

$$\mathbb{E}(d_{2n}|\mathscr{G}_{2n-1}) = \frac{1}{2}v_n\mathbb{E}(\xi_n - \widetilde{\xi}_n|\mathscr{G}_{2n-1}) = 0.$$

Now since

$$\sum_{n=1}^{N} v_n\xi_n = \sum_{j=1}^{2N} d_j \quad \text{and} \quad \sum_{n=1}^{N} v_n\widetilde{\xi}_n = \sum_{j=1}^{2N}(-1)^{j+1}d_j,$$

the result follows from the UMD property applied to the sequences $(d_j)_{j=1}^{2N}$ and $((-1)^{j+1}d_j)_{j=1}^{2N}$. □

Remark 4.4.3. More generally, the result remains true for predictable sequences v on $\Omega \times \Omega$, with essentially the same proof.

4.4.b Tangent sequences

The decoupling result of the previous section has limited applicability for two reasons: it is formulated on probability spaces and the formulation includes an independence assumption. For applications in harmonic analysis, where one often works with dyadic filtrations on $\mathbb{R}^d$, one needs an extension to the σ-finite setting; but here the notion of independence loses its meaning. The notion of tangency solves both problems in an adequate way.

Throughout this subsection $(S, \mathscr{A}, \mu)$ is a measure space endowed with a σ-finite filtration $(\mathscr{F}_n)_{n\geqslant 0}$.

Definition 4.4.4. *Two adapted sequences $(d_n)_{n=1}^N$ and $(e_n)_{n=1}^N$ in $L^0(S;X)$ are called* tangent *to each other if for all $n = 1, \dots, N$ and all $A \in \mathscr{F}_{n-1}$ and Borel sets $B \subseteq X$ we have*

$$\mu(A \cap \{d_n \in B\}) = \mu(A \cap \{e_n \in B\}).$$

Taking $A = S$, in particular this implies that for each $n = 1, \dots, N$ the functions d_n and e_n are identically distribution. Definition 4.4.4 may be rephrased as saying that two sequences are tangent if, for each $n = 1, \dots, N$, *conditionally on $\mathscr{F}_{n-1}$ the functions d_n and e_n are identically distributed.* This terminology is justified by the observation that Definition 4.4.4 is equivalent to the requirement

$$\mathbb{E}(\mathbf{1}_{\{d_n\in B\}}|\mathscr{F}_{n-1}) = \mathbb{E}(\mathbf{1}_{\{e_n\in B\}}|\mathscr{F}_{n-1})$$

in those situations where the existence of the conditional expectations can be guaranteed (e.g., when μ is σ-finite on $\mathscr{F}_0$).

The following lemma provides useful alternative descriptions of tangency:

Lemma 4.4.5. *Let $(d_n)_{n=1}^N$ and $(e_n)_{n=1}^N$ be adapted sequences in $L^0(S;X)$. The following assertions are equivalent:*

(1) *$(d_n)_{n=1}^N$ and $(e_n)_{n=1}^N$ are tangent;*

(2) *for all* $n = 1, \dots, N$, *any measurable space* $(T, \mathscr{B})$, *and any non-negative measurable function* $h : X \times T \to \mathbb{R}_+$ *and* $\mathscr{F}_{n-1}$*-measurable function* $v : S \to T$ *we have*

$$\int_S h(d_n, v)\,\mathrm{d}\mu = \int_S h(e_n, v)\,\mathrm{d}\mu;$$

(3) *for all* $n = 1, \dots, N$, *any measurable space* $(T, \mathscr{B})$, *any Banach space* Y, *any measurable function* $h : X \times T \to Y$ *and* $\mathscr{F}_{n-1}$*-measurable function* $v : S \to T$ *such that* $h(d_n, v), h(e_n, v) \in L^1(S; Y)$ *we have*

$$\int_S h(d_n, v)\,\mathrm{d}\mu = \int_S h(e_n, v)\,\mathrm{d}\mu.$$

Proof. (1)⇒(2): Fix $1 \leqslant n \leqslant N$ and an $\mathscr{F}_{n-1}$-measurable function $v : S \to T$. Also fix a set $F_0 \in \mathscr{F}_0$ of finite measure. Define the finite measures λ_1 and λ_2 on $\mathscr{B}(X) \times \mathscr{B}$ by

$$\lambda_1(D) := \mu(F_0 \cap \{(d_n, v) \in D\}), \quad \lambda_2(D) := \mu(F_0 \cap \{(e_n, v) \in D\}).$$

For all $B \in \mathscr{B}(X)$ and $C \in \mathscr{B}$,

$$\begin{aligned}\lambda_1(B \times C) &= \mu((F_0 \cap \{v \in C\}) \cap \{d_n \in B\}) \\ &= \mu((F_0 \cap \{v \in C\}) \cap \{e_n \in B\}) = \lambda_2(B \times C).\end{aligned}$$

By Dynkin's lemma (Lemma A.1.3) this implies that $\lambda_1 = \lambda_2$. Hence,

$$\mu(F_0 \cap \{(d_n, v) \in D\}) = \mu(F_0 \cap \{(e_n, v) \in D\})$$

for all $D \in \mathscr{B}(X) \times \mathscr{B}$. An exhaustion argument with sets in $\mathscr{F}_0$ of finite measure then implies that

$$\mu\{(d_n, v) \in D\} = \mu\{(e_n, v) \in D\}$$

for all $D \in \mathscr{B}(X) \times \mathscr{B}$.

Pick simple functions $h_j : X \times T \to \mathbb{R}_+$ with $0 \leqslant h_j \uparrow h$ pointwise. By what we just proved and linearity we have

$$\int_S h_j(d_n, v)\,\mathrm{d}\mu = \int_S h_j(e_n, v)\,\mathrm{d}\mu.$$

Letting j tend to infinity on both sides, (2) follows from the monotone convergence theorem.

(2)⇒(3): By the Hahn–Banach theorem it suffices to show that for all $x^* \in X^*$ one has $\int_S \langle h(d_n, v), x^* \rangle\,\mathrm{d}\mu = \int_S \langle h(e_n, v), x^* \rangle\,\mathrm{d}\mu$. But this identity follows from (2) by considering real and imaginary parts, and of these the positive and negative parts.

(3)⇒(2): This is trivial.

(2)⇒(1): Fix $n \geqslant 1$, $B \in \mathscr{B}(X)$ and $A \in \mathscr{F}_{n-1}$. We need to show that

$$\int_S h(d_n, \mathbf{1}_A)\,\mathrm{d}\mu = \int_S h(e_n, \mathbf{1}_A)\,\mathrm{d}\mu,$$

where $h : X \times \mathbb{R}_+ \to \mathbb{R}_+$ is given by $h(x, y) = \mathbf{1}_{\{x\in B\}}\, y$. This is immediate from (2). □

The following proposition asserts that tangency is preserved under some functional operations.

Proposition 4.4.6. *Let X and Y be Banach spaces, with Y separable, let $(T, \mathscr{B})$ be a measurable space, and suppose that for all $n = 1, \dots, N$ a measurable function $h_n : X \times T \to Y$ is given. If $(v_n)_{n=1}^N$ is a T-valued predictable sequence and $(d_n)_{n=1}^N$ and $(e_n)_{n=1}^N$ are X-valued tangent sequences, then $(h_n(d_n, v_n))_{n=1}^N$ and $(h_n(e_n, v_n))_{n=1}^N$ are tangent.*

Proof. First note that the Y-valued functions $h_n(d_n, v_n)$ and $h_n(e_n, v_n)$ are measurable and therefore strongly measurable by the separability of Y and by Theorem 1.1.6.

Fix $n \geqslant 1$, a Borel set $C \subseteq Y$ and a set $A \in \mathscr{F}_{n-1}$. By the definition of tangent sequences we need to show that

$$\mu\big(\{h_n(d_n, v_n) \in C\} \cap A\big) = \mu\big(\{h_n(e_n, v_n) \in C\} \cap A\big).$$

But this follows from Lemma 4.4.5, applied to $h_n^C : X \times T \times S \to \mathbb{R}_+$ given by $h_n^C(x, t, s) = \mathbf{1}_{\{h_n(x,t)\in C\}} \mathbf{1}_A(s)$, the tangent sequences $(d_n)_{n=1}^N$ and $(e_n)_{n=1}^N$ and the $\mathscr{F}_{n-1}$-measurable function $s \mapsto (v_n(s), s)$. □

We continue with some examples of tangency which establish the link with the decoupling theorem of the previous subsection.

Example 4.4.7. Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space with filtration $(\mathscr{F}_n)_{n\geqslant 0}$, and let $\xi = (\xi_n)_{n=1}^N$ and $\widetilde{\xi} = (\widetilde{\xi}_n)_{n=1}^N$ be adapted sequences in $L^0(\Omega; X)$. Suppose that for $n = 1, \dots, N$ the following two assumptions are satisfied:

(i) ξ_n and $\widetilde{\xi}_n$ are identically distributed;
(ii) ξ_n and $\widetilde{\xi}_n$ are independent of $\mathscr{F}_{n-1}$.

Then ξ and $\widetilde{\xi}$ are tangent. Indeed, for $A \in \mathscr{F}_{n-1}$ we have

$$\begin{aligned}\mathbb{P}(A \cap \{\xi_n \in B\}) &= \int_A \mathbf{1}_{\{\xi_n\in B\}}\,\mathrm{d}\mathbb{P}\\ &= \int_A \mathbb{E}(\mathbf{1}_{\{\xi_n\in B\}}|\mathscr{F}_{n-1})\,\mathrm{d}\mathbb{P} = \int_A \mathbb{E}\mathbf{1}_{\{\xi_n\in B\}}\,\mathrm{d}\mathbb{P}\end{aligned}$$

using (ii) and Proposition 2.6.35 in the last step; this proposition also guarantees the existence of the conditional expectation. Similarly,

$$\mathbb{P}(A \cap \{\widetilde{\xi}_n \in B\}) = \int_A \mathbb{E}\mathbf{1}_{\{\widetilde{\xi}_n \in B\}} \, \mathrm{d}\mathbb{P}.$$

Since $\mathbb{E}\mathbf{1}_{\{\xi_n \in B\}} = \mathbb{E}\mathbf{1}_{\{\widetilde{\xi}_n \in B\}}$ by (i), the result follows.

Combining this example with Proposition 4.4.6, we obtain a further example of tangency.

Example 4.4.8. Let $(d_n)_{n=1}^N$ and $(\widetilde{d}_n)_{n=1}^N$ be tangent sequences and let $(v_n)_{n=1}^N$ be $\mathscr{L}(X,Y)$-valued and strongly predictable with respect to the strong operator topology. By strong measurability, there is no loss of generality in assuming Y to be separable.

By Proposition 4.4.6 the sequences $(v_n d_n)_{n=1}^N$ and $(v_n \widetilde{d}_n)_{n=1}^N$ are tangent. Indeed, taking $T = \mathscr{L}(X,Y)$ endowed with the σ-algebra generated by the strong operator topology, the $\mathscr{F}_{n-1}$-measurability of v_n follows from the $\mathscr{F}_{n-1}$-measurability of $s \mapsto v_n(s)x$.

In a similar way, in the setting of Example 4.4.7 we find that $(v_n \xi_n)_{n=1}^N$ and $(v_n \widetilde{\xi}_n)_{n=1}^N$ are tangent.

The main result of this section is the following theorem, which is a generalisation of Theorems 4.2.25 and 4.4.1.

Theorem 4.4.9 (Tangent martingale transform). *Let X be a UMD space and let $p \in (1,\infty)$. Suppose that $(f_n)_{n=0}^N$ and $(g_n)_{n=0}^N$ are X-valued L^p-martingales such that $f_0 = g_0$. Let $v = (v_n)_{n=0}^N$ be a scalar predictable sequence (where we assume v_0 is constant) such that $\|v\|_\infty \leqslant 1$ and let*

$$(v \star g)_n = v_0 g_0 + \sum_{j=1}^n v_j dg_j, \quad n \in \{0, \dots, N\}.$$

If the sequences $(df_n, dg_n)_{n=1}^N$ and $(dg_n, df_n)_{n=1}^N$ are tangent, then

$$\|(v \star g)_N\|_p \leqslant \beta_{p,X} \|f_N\|_p. \tag{4.27}$$

The following remark describes a sufficient condition for the tangency assumption appearing in the formulation of the theorem.

Remark 4.4.10. If, for each $n = 1, \dots, N$, the sequences $(df_n)_{n\geqslant 1}$ and $(dg_n)_{n\geqslant 1}$ are tangent and *conditionally independent given* $\mathscr{F}_{n-1}$, i.e.,

$$\begin{aligned} &\mathbb{P}(\{df_n \in B_1, dg_n \in B_2\} | \mathscr{F}_{n-1}) \\ &\quad = \mathbb{P}(\{df_n \in B_1\} | \mathscr{F}_{n-1}) \mathbb{P}(\{dg_n \in B_2\} | \mathscr{F}_{n-1}) \quad \forall B_1, B_2 \in \mathscr{B}(X), \end{aligned} \tag{4.28}$$

then the sequences $(df_n, dg_n)_{n=1}^N$ and $(dg_n, df_n)_{n=1}^N$ are tangent. In (4.28), $\mathbb{P}(A|\mathscr{F}_{n-1}) := \mathbb{E}(\mathbf{1}_A|\mathscr{F}_{n-1})$ is the conditional probability of A given $\mathscr{F}_{n-1}$. Indeed, this follows from

$$\begin{aligned} \mathbb{P}(\{df_n \in B_1, dg_n \in B_2\} | \mathscr{F}_{n-1}) &= \mathbb{P}(\{df_n \in B_1\} | \mathscr{F}_{n-1}) \mathbb{P}(\{dg_n \in B_2\} | \mathscr{F}_{n-1}) \\ &= \mathbb{P}(\{dg_n \in B_1\} | \mathscr{F}_{n-1}) \mathbb{P}(\{df_n \in B_2\} | \mathscr{F}_{n-1}) \\ &= \mathbb{P}(\{dg_n \in B_1, df_n \in B_2\} | \mathscr{F}_{n-1}). \end{aligned}$$

Let us now turn to the proof of the theorem. Proceeding as the proof of Theorem 4.4.1, we will derive the estimate (4.27) from Theorem 4.2.25 applied to a suitable martingale difference sequence.

Proof of Theorem 4.4.9. Define a filtration $(\mathscr{G}_k)_{k=0}^{2N}$ by

$$\begin{aligned} \mathscr{G}_{2n} &:= \mathscr{F}_n, && 0 \leqslant n \leqslant N, \\ \mathscr{G}_{2n-1} &:= \sigma(\mathscr{F}_{n-1}, df_n + dg_n), && 1 \leqslant n \leqslant N, \end{aligned}$$

and set

$$\begin{aligned} d_{2n} &:= \tfrac{1}{2}(df_n - dg_n), && 1 \leqslant n \leqslant N, \\ d_{2n-1} &:= \tfrac{1}{2}(df_n + dg_n), && 1 \leqslant n \leqslant N. \end{aligned}$$

We claim that $(d_k)_{k=1}^{2N}$ is a martingale difference sequence with respect to $(\mathscr{G}_k)_{k=0}^{2N}$.

Adaptedness is obvious. The identity $\mathbb{E}(d_{2n-1}|\mathscr{G}_{2n-2}) = 0$ follows from the martingale difference property of df_n and dg_n. To check that $\mathbb{E}(d_{2n}|\mathscr{G}_{2n-1}) = 0$, by Dynkin's Lemma A.1.3 and an approximation argument it suffices to show that for all Borel sets $B \in \mathscr{B}(X)$ and $A \in \mathscr{F}_{n-1}$ with $\mu(A) < \infty$ we have

$$T := \int_{A\cap\{df_n+dg_n\in B\}} d_{2n}\,\mathrm{d}\mu = 0.$$

Since (df_n, dg_n) and (dg_n, df_n) are tangent and $A \in \mathscr{F}_{n-1}$, Lemma 4.4.5 applied with $h : X \times X \times S \to X$ given by $h(x,y,s) := (x-y)\mathbf{1}_{\{x+y\in B\}}\mathbf{1}_A(s)$ implies that

$$T = \frac{1}{2}\int_{A\cap\{df_n+dg_n\in B\}} df_n - dg_n\,\mathrm{d}\mu = \frac{1}{2}\int_{A\cap\{dg_n+df_n\in B\}} dg_n - df_n\,\mathrm{d}\mu = -T.$$

Hence $T = 0$ as required.

Let $\theta_{2k-1} := v_k$ and $\theta_{2k} := -v_k$ for $k \in \{1,\dots,N\}$. Since $(d_k)_{k=1}^{2N}$ is a martingale difference sequence and $(\theta_k)_{k=0}^{2N}$ is predictable, it follows from Theorem 4.2.25 that

$$\|(v \star g)_N\|_p = \Big\|v_0 f_0 + \sum_{k=1}^{2N}\theta_k d_k\Big\|_p \leqslant \beta_{p,X}\Big\|f_0 + \sum_{k=1}^{2N} d_k\Big\|_p = \beta_{p,X}\|f_N\|_p.$$

□

The following consequence of Theorem 4.4.9 can be thought of as a generalisation of Theorem 4.4.1 to the σ-finite setting, in which the independence assumption is replaced by a tangency assumption for martingale differences (cf. Examples 4.4.7 and 4.4.8). This theorem and some of its variations will be discussed in the Notes at the end of this chapter. In particular, it can be shown that it is enough to assume that $(df_n)_{n=1}^N$ and $(dg_n)_{n=1}^N$ are tangent at the cost of an extra multiplicative UMD constant $\beta_{p,X}$; see Notes before (4.50).

Theorem 4.4.11 (Hitczenko–McConnell). *Let X be a UMD space and let $p \in (1,\infty)$. Suppose that $(f_n)_{n=0}^N$ and $(g_n)_{n=0}^N$ are X-valued L^p-martingales such that $f_0 = g_0$. If the sequences $(df_n, dg_n)_{n=1}^N$ and $(dg_n, df_n)_{n=1}^N$ are tangent, then*

$$\frac{1}{\beta_{p,X}}\|g_N\|_p \leqslant \|f_N\|_p \leqslant \beta_{p,X}\|g_N\|_p. \tag{4.29}$$

Proof. The second estimate in (4.29) follows from Theorem 4.4.9 with $v_k = 1$. The first estimate in (4.29) follows by interchanging f and g. □

Theorem 4.4.1 is a special case of Theorem 4.4.11:

Example 4.4.12. As in Example 4.4.7, let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space with filtration $(\mathscr{F}_n)_{n\geqslant 0}$, and let $\xi = (\xi_n)_{n=1}^N$ and $\widetilde{\xi} = (\widetilde{\xi}_n)_{n=1}^N$ be adapted sequences in $L^0(\Omega; X)$. If ξ_n and $\widetilde{\xi}_n$ are independent for each $n = 1, \dots, N$, the sequences $(\xi_n, \widetilde{\xi}_n)_{n=1}^N$ and $(\widetilde{\xi}_n, \xi_n)_{n=1}^N$ satisfy the assumptions (i) and (ii) of Example 4.4.7 and therefore they are tangent. If $(v_n)_{n=1}^N$ is $\mathscr{L}(X,Y)$-valued and strongly predictable, Example 4.4.8 implies that the sequences $(v_n\xi_n, v_n\widetilde{\xi}_n)_{n=1}^N$ and $(v_n\widetilde{\xi}_n, v_n\xi_n)_{n=1}^N$ are tangent as well. Therefore, Theorem 4.4.1 is a special case of Theorem 4.4.11.

In the following example we show that for fairly general adapted sequences $(d_n)_{n=1}^N$ one can construct a sequence $(e_n)_{n=1}^n$ such that (d_n, e_n) and (e_n, d_n) are tangent. It turns out that many tangent sequences can be modelled in this way and this is the main reason we include it here.

Example 4.4.13. Let $(S, \mathscr{A}, \mu)$ be a finite measure space and consider its $2N$-fold product space: $(S^{2N}, \mathscr{A}^{2N}, \mu^{2N})$. For each $1 \leqslant n \leqslant N$ let $h_n \in L^p(S^n; X)$ be such that for almost all $\mathbf{s}_{n-1} := (s_1, \dots, s_{n-1}) \in S^{n-1}$

$$\int_S h_n(\mathbf{s}_{n-1}, s_n)\,\mathrm{d}\mu(s_n) = 0.$$

Let $\mathscr{F}_n := (\mathscr{A}^n \times \{S, \varnothing\}^{N-n}) \times \mathscr{A}^n \times \{S, \varnothing\}^{N-n}$ and define $d_n : S^{2N} \to X$ and $e_n : S^{2N} \to X$ by

$$d_n(\mathbf{s}_N, \mathbf{t}_N) := h_n(\mathbf{s}_{n-1}, s_n) \quad \text{and} \quad e_n(\mathbf{s}_N, \mathbf{t}_N) := h_n(\mathbf{s}_{n-1}, t_n).$$

The sequences $(d_n)_{n=1}^N$ and $(e_n)_{n=1}^N$ are martingale difference sequences with respect to the filtration $(\mathscr{F}_n)_{n=0}^N$. Moreover, we claim that $(d_n, e_n)_{n=1}^N$ and $(e_n, d_n)_{n=1}^N$ are tangent. In many situations $(d_n)_{n=1}^N$ can be modelled in the above form, and the sequence $(e_n)_{n=1}^N$ is constructed from it and is often simpler to handle.

To prove the claim, note that for $1 \leqslant n \leqslant N$, $A \in \mathscr{F}_{n-1}$ and Borel sets $B_1, B_2 \in \mathscr{B}(X)$ we find that for almost all $\mathbf{s}_{n-1} = (s_1, \dots, s_{n-1}) \in S^{n-1}$,

$$\mu \times \mu(\{d_n(\mathbf{s}_{n-1}, \cdot) \in B_1\} \cap \{e_n(\mathbf{s}_{n-1}, \cdot) \in B_2\})$$

$$= \int_S \int_S \mathbf{1}_{\{h_n(\mathbf{s}_{n-1},s_n)\in B_1\}} \mathbf{1}_{\{h_n(\mathbf{s}_{n-1},t_n)\in B_2\}} \, \mathrm{d}\mu(s_n)\, \mathrm{d}\mu(t_n)$$
$$= \mu \times \mu(\{e_n(\mathbf{s}_{n-1},\cdot) \in B_1\} \cap \{d_n(\mathbf{s}_{n-1},\cdot) \in B_2\}).$$

Integration over $\mathbf{s}_{n-1} \in A$ gives the desired result.

The validity of the inequality of Theorem 4.4.11 already implies the UMD property, and it even suffices to consider Paley–Walsh martingale differences for this purpose. This is the content of the next proposition.

Proposition 4.4.14. *Let X be a Banach space and $p \in (1,\infty)$. If there exists a constant $C \geqslant 0$ with the property that for all X-valued Paley–Walsh martingale difference sequences $(d_n)_{n=1}^N$ and $(e_n)_{n=1}^N$ such that $(d_n, e_n)_{n=1}^N$ and $(e_n, d_n)_{n=1}^N$ are tangent, then*

$$\Big\| \sum_{n=1}^N d_n \Big\|_p \leqslant C \Big\| \sum_{n=1}^N e_n \Big\|_p,$$

then X is a UMD space and $\beta_{p,X}^{\mathbb{R}} \leqslant C$.

Proof. By Theorem 4.2.5, it suffices to check the UMD property for Paley–Walsh martingale difference sequences $(d_n)_{n=1}^N$ defined on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$.

By Proposition 3.1.10 we may assume that d_n is of the form

$$d_n = r_n \phi_n(r_1, \dots, r_{n-1}),$$

where $(r_n)_{n=1}^N$ is a real Rademacher sequence on the underlying probability space Ω. For fixed signs $\epsilon_1, \dots \epsilon_N$ in $\{-1, 1\}$ let $e_n := \epsilon_n d_n$. It is easily checked that $(r_n, \epsilon_n r_n)_{n=1}^N$ and $(\epsilon_n r_n, r_n)_{n=1}^N$ are tangent. Hence, by Example 4.4.8, also $(d_n, e_n)_{n=1}^N$ and $(e_n, d_n)_{n=1}^N$ are tangent. Using the assumption of the proposition, we obtain

$$\Big\| \sum_{n=1}^N d_n \Big\|_p \leqslant C \Big\| \sum_{n=1}^N e_n \Big\|_p = C \Big\| \sum_{n=1}^N \epsilon_n d_n \Big\|_p.$$

□

We end this section with an application of Theorem 4.4.9; some applications mentioned in the Notes. To state the result we first introduce some notation. Consider a σ-finite measure space $(S, \mathscr{F}, \mu)$. For each $n \in \mathbb{Z}$ let $\mathscr{A}_n$ be a partition of S into at most countably many atoms of finite measure. We assume that $\mathscr{A}_n \subseteq \mathscr{A}_{n+1}$ and $\mathscr{A}_n \cap \mathscr{A}_{n+1} = \varnothing$ for all $n \in \mathbb{Z}$, i.e., every set $A \in \mathscr{A}_n$ can be written as a union of at least two sets of finite non-zero measure from the refined partition $\mathscr{A}_{n+1}$. Let $\mathscr{A} := \bigcup_{n\in\mathbb{Z}} \mathscr{A}_n$. For a set $A \in \mathscr{A}_n$ we let

$$\mathrm{child}_{\mathscr{A}}(A) := \{A' \in \mathscr{A}_{n+1} : A' \subseteq A\}.$$

For each $A \in \mathscr{A}$, set $T_A := A$ and define a σ-algebra $\mathscr{G}_A$ and a measure ν_A on T_A by setting

$$\mathscr{G}_A := \sigma(\mathrm{child}_{\mathscr{A}}(A)) \quad \text{and} \quad \nu_A := \frac{\mu|_A}{\mu(A)}.$$

For each $A \in \mathscr{A}$, $(T_A, \mathscr{G}_A, \nu_A)$ is a probability space. We shall consider the product space

$$(T, \mathscr{G}, \nu) := \prod_{A \in \mathscr{A}} (T_A, \mathscr{G}_A, \nu_A).$$

A function $\phi : S \to X$ is called an *atom supported in* $A \in \mathscr{A}$ if

- $\mathrm{supp}(\phi) \subseteq A$;
- ϕ is constant on each $A' \in \mathrm{child}_{\mathscr{A}}(A)$;
- $\int_S \phi \, \mathrm{d}\mu = 0$.

Corollary 4.4.15. *Let X be a UMD space and $p \in (1, \infty)$. For each $A \in \mathscr{A}$ let $\phi_A : S \to X$ be an atom supported in A and let $f = \sum_{A \in \mathscr{A}} \phi_A$ be unconditionally convergent in $L^p(S; X)$. Then for any sequence $(\epsilon_A)_{A \in \mathscr{A}}$ of signs in $\{z \in \mathbb{K} : |z| = 1\}$, we have*

$$\begin{aligned} &\beta_{p,X}^{-1} \|f\|_{L^p(S;X)} \\ &\quad \leqslant \Big(\int_S \int_T \Big\| \sum_{A \in \mathscr{A}} \epsilon_A \mathbf{1}_A(s) \phi_A(t) \Big\|^p \, \mathrm{d}\nu(t) \, \mathrm{d}\mu(s) \Big)^{1/p} \leqslant \beta_{p,X} \|f\|_{L^p(S;X)}, \end{aligned}$$

where the expression in the middle also converges unconditionally in $L^p(S \times T; X)$.

By applying this estimate to $\epsilon_A = \varepsilon_A(\omega)$, where $(\varepsilon_A)_{A \in \mathscr{A}}$ is a Rademacher sequence on a probability space $(\Omega, \mathbb{P})$, and taking $L^p(\Omega)$-norms on both sides we find that

$$\|f\|_{L^p(S;X)} \eqsim_{p,X} \Big(\int_S \int_T \mathbb{E} \Big\| \sum_{A \in \mathscr{A}} \varepsilon_A \mathbf{1}_A(s) \phi_A(t) \Big\|^p \, \mathrm{d}\nu(t) \, \mathrm{d}\mu(s) \Big)^{1/p}.$$

The usefulness of this estimate partly comes from the fact that for fixed $s \in S$, the elements $(\varepsilon_A \mathbf{1}_A(s) \phi_A)_{A \in \mathscr{A}}$ are independent symmetric random variables. This additional structure is often simpler to handle than f itself.

Proof. By an approximation argument it suffices to consider the case where $\phi_A \neq 0$ for only finitely many sets $A \in \mathscr{A}$. Thus we may just as well assume that $\mathscr{A}$ is finite, say $\mathscr{A} = \{A_1, \dots, A_N\}$ with either $A_n \subseteq A_m$ or $A_n \cap A_m = \varnothing$ for all $1 \leqslant n < m \leqslant N$. Define a filtration $(\mathscr{F}_n)_{n=0}^N$ by

$$\mathscr{F}_n := \sigma(\{\mathrm{child}_{\mathscr{A}}(A_1), \dots, \mathrm{child}_{\mathscr{A}}(A_n)\}), \quad \text{for } n \in \{0, \dots, N\}.$$

Observe that $(\phi_{A_n})_{n=1}^N$ defines a martingale difference sequence with respect to this filtration. The required measurability is obvious, and fixing

$n \in \{1, \ldots, N\}$ it follows that for each $m < n$ there are two possibilities: either $A_n \subseteq A_m$ or $A_n \cap A_m = \varnothing$. In the latter case, $\int_{A_m} \phi_{A_n} \, \mathrm{d}\mu = 0$ by the support condition of ϕ_{A_n}. In the former case the conditions on ϕ_{A_n} yield

$$\int_{A_m} \phi_{A_n} \, \mathrm{d}\mu = \int_{A_n} \phi_{A_n} \, \mathrm{d}\mu = 0.$$

By the additional support condition on the ϕ_A's we can redefine $(T, \mathscr{G}, \nu)$ as the product of the probability spaces $(T_{A_n}, \mathscr{G}_{A_n}, \nu_{A_n})_{n=1}^N$. On T we consider the filtration $(\widetilde{\mathscr{G}}_n)_{n=0}^N$ defined by

$$\widetilde{\mathscr{G}}_n := \mathscr{G}_{A_1} \times \ldots \mathscr{G}_{A_n} \times \{\varnothing, A_{n+1}\} \times \ldots \times \{\varnothing, A_N\}.$$

On $(S \times T, \mathscr{F}_N \times \widetilde{\mathscr{G}}_N, \mu \times \nu)$, for $n \in \{1, \ldots, N\}$ we define

$$df_n(s, \mathbf{t}) := \phi_{A_n}(s) \quad \text{and} \quad dg_n(s, \mathbf{t}) := \mathbf{1}_{A_n}(s) \phi_{A_n}(t_n),$$

where $\mathbf{t} = (t_1, \ldots, t_N)$. By the previous observation, these sequences are martingale difference sequences with respect to the filtration $(\mathscr{F}_n \times \widetilde{\mathscr{G}}_n)_{n=0}^N$.

We claim that $(df_n, dg_n)_{n=1}^N$ and $(dg_n, df_n)_{n=1}^N$ are tangent. As soon as we have proved this, the required result follows from Theorem 4.4.9 with $v_n := \epsilon_{A_n}$.

To check the tangency, let $B_1, B_2 \subseteq X$ be Borel sets, let $A \in \mathscr{F}_{n-1}$ and $C \in \widetilde{\mathscr{G}}_{n-1}$ both be atoms. Then we can write

$$C = A'_1 \times \ldots \times A'_{n-1} \times A_n \times \ldots \times A_N,$$

where $A'_j \in \mathrm{child}_{\mathscr{A}}(A_j)$ for each $j = 1, \ldots, n-1$. First consider the case $A_n \subseteq A$. Then, with $K := \prod_{j=1}^{n-1} \mu(A'_j)/\mu(A_j)$,

$$\begin{aligned}
&(\mu \times \nu)\big(\{df_n \in B_1, dg_n \in B_2\} \cap (A \times C)\big) \\
&= (\mu \times \nu_{A_n})\big(\{(s, t_n) \in A \times A_n : \ \phi_{A_n}(s) \in B_1, \ \mathbf{1}_{A_n}(s)\phi_{A_n}(t_n) \in B_2\}\big) \cdot K \\
&= (\mu \times \nu_{A_n})\big(\{(s, t_n) \in A_n \times A_n : \ \phi_{A_n}(s) \in B_1, \ \phi_{A_n}(t_n) \in B_2\}\big) \cdot K \\
&\qquad + (\mu \times \nu_{A_n})\big(\{(s, t_n) \in (A \setminus A_n) \times A_n : \ 0 \in B_1, \ 0 \in B_2\}\big) \cdot K \\
&= \mu(\{\phi_{A_n} \in B_1\} \cap A_n) \cdot \mu(\{\phi_{A_n} \in B_2\} \cap A_n) \cdot \frac{K}{\mu(A_n)} \\
&\qquad + \mu(A \setminus A_n) \mathbf{1}_{B_1 \times B_2}(0,0) \cdot K \\
&= (\mu \times \nu)\big(\{dg_n \in B_1, df_n \in B_2\} \cap (A \times C)\big),
\end{aligned}$$

where the last equality follows by reversing the previous identities with B_1 and B_2 interchanged, noting that the penultimate expression is symmetric in B_1 and B_2. If $A_n \not\subseteq A$, then $A_n \cap A = \varnothing$ and in the same way as before one sees that

$$(\mu \times \nu)\big(\{df_n \in B_1, dg_n \in B_2\} \cap (A \times C)\big)$$

$$= \mu(A)\mathbf{1}_{B_1\times B_2}(0,0)\cdot K$$
$$= (\mu\times\nu)\big(\{dg_n \in B_1, df_n \in B_2\}\cap (A\times C)\big).$$

This completes the proof of the tangency. □

4.5 Burkholder functions and sharp UMD constants

4.5.a Concave functions

We shall now introduce several notions of concavity for scalar-valued functions defined on a Banach space X or on the product $X\times X$. These will be important later on when we turn to Burkholder's characterisation of UMD spaces X in terms of the existence of suitable concave functions on $X\times X$.

Definition 4.5.1. *A function $u : X \to \mathbb{R}$ is called* concave *if for all $x_0, x_1 \in X$ and $0 < \lambda < 1$ we have*

$$u((1-\lambda)x_0 + \lambda x_1) \geqslant (1-\lambda)u(x_0) + \lambda u(x_1).$$

A function $u : X \to \mathbb{R}$ is called midpoint concave *if the above holds with $\lambda = \frac{1}{2}$.*

Lemma 4.5.2. *Let $u : X \to \mathbb{R}$ be locally bounded from below. Then u is concave if and only if it is midpoint concave.*

Proof. Suppose, for a contradiction, that $u : X \to \mathbb{R}$ is midpoint concave and locally bounded from below, but fails to be concave. Reparametrising the line-segment for which the concavity condition fails by the unit interval, this amounts to the assumption that $u(\lambda) < (1-\lambda)u(0) + \lambda u(1)$ for a midpoint concave function $u : [0,1] \to \mathbb{R}$ which is bounded from below (we can drop the adjective "locally" since $[0,1]$ is compact). By adding an affine function, which does not affect the concavity properties and boundedness properties of u, we may further assume that $u(0) = u(1) = 0$ and $u(\lambda) < 0$.

By iterating the midpoint concavity assumption, it readily follows that we have $u(t) \geqslant (1-t)u(0) + tu(1) = 0$ at every dyadic rational $t = k2^{-j}$; $j \in \mathbb{N}$, $k = 0, 1, \dots, 2^j$. Fix such a dyadic rational t close enough to λ, so that also $\lambda_1 := \lambda - (t-\lambda) = 2\lambda - t$ belongs to $(0,1)$. Now $\lambda = \frac{1}{2}(\lambda_1 + t)$, and hence midpoint concavity implies that $u(\lambda) \geqslant \frac{1}{2}(u(\lambda_1)+u(t)) \geqslant \frac{1}{2}u(\lambda_1)$. Thus $u(\lambda_1) \leqslant 2u(\lambda)$.

Repeating the argument with λ_1 in place of λ, we find another point $\lambda_2 \in (0,1)$ with $u(\lambda_2) \leqslant 2u(\lambda_1) \leqslant 4u(\lambda)$, and then by induction a sequence of points $\lambda_n \in (0,1)$ such that $u(\lambda_n) \leqslant 2^n u(\lambda)$. But this tends to $-\infty$ as $n \to \infty$, contradicting the boundedness from below. □

Definition 4.5.3. *A function $u : X \times X \to \mathbb{R}$ is said to be* (midpoint) zigzag-concave *if for all $x, y \in X$ and signs ϵ in $\mathbb{K}$ the function*

$$z \mapsto u(x + z, y + \epsilon z)$$

is (midpoint) concave on X.

In the following variant of Jensen's inequality we do not assume any continuity properties of u.

Lemma 4.5.4. *A function $u : X \times X \to \mathbb{R}$ is zigzag-concave if and only if for every simple random variable $\xi : \Omega \to X$ with $\mathbb{E}\xi = 0$ and all $x, y \in X$ and signs $\epsilon \in \mathbb{K}$ one has*

$$\mathbb{E}u(x + \xi, y + \epsilon\xi) \leqslant u(x, y).$$

Proof. If u is zigzag-concave, then for any simple random variable $\xi = \sum_{n=1}^{N} \mathbf{1}_{A_n} z_n$ with $\sum_{n=1}^{N} \mu(A_n) = 1$ and $\mathbb{E}\xi = 0$ we have

$$\begin{aligned}
\mathbb{E}u(x + \xi, y + \epsilon\xi) &= \sum_{n=1}^{N} \mu(A_n) u(x + z_n, y + \epsilon z_n) \\
&\leqslant u\Big(x + \sum_{n=1}^{N} \mu(A_n) z_n, y + \epsilon \sum_{n=1}^{N} \mu(A_n) z_n\Big) \\
&= u(x + \mathbb{E}\xi, y + \epsilon\mathbb{E}\xi) = u(x, y).
\end{aligned}$$

To prove the converse pick $x, y, z_1, z_2 \in X$ and $\lambda \in (0, 1)$, and set $z := (1-\lambda)z_1 + \lambda z_2$. Let the random variable ξ take the value $z_1 - z$ with probability $1 - \lambda$ and the value $z_2 - z$ with probability λ. Then $\mathbb{E}\xi = 0$ and hence

$$\begin{aligned}
(1 - \lambda)u(x + z_1, y + \epsilon z_1) + \lambda u(x + z_2, y + \epsilon z_2) &= \mathbb{E}u(x + z + \xi, y + \epsilon z + \epsilon\xi) \\
&\leqslant u(x + z, y + \epsilon z).
\end{aligned}$$

□

Lemma 4.5.5. *Let $u : X \times X \to \mathbb{R}$ be zigzag-concave and satisfy $u(0, 0) = 0$. If $f := (f_n)_{n\geqslant 0}$ is a μ-simple martingale on $(S, \mathscr{A}, \mu)$ with a finitely non-zero difference sequence and $g = \epsilon \star f$ is its transform by a sequence $\epsilon := (\epsilon_n)_{n\geqslant 1}$ of signs in $\mathbb{K}$, then for all $n \geqslant 1$ we have*

$$\int_S u(f_n, g_n)\,\mathrm{d}\mu \leqslant \int_S u(f_{n-1}, g_{n-1})\,\mathrm{d}\mu.$$

Proof. Fix $n \geqslant 1$. Since f is μ-simple, without loss of generality we may assume $\mu(S) < \infty$. Moreover, replacing $\mathscr{F}_k$ by the σ-algebra generated by f_0 and the differences df_j for $j \leqslant k$, we may assume $\mathscr{F}_k$ consists of finitely many sets. Let $\mathscr{F}_k^*$ be the family of atoms of $\mathscr{F}_k$. For a non-empty set $A \in \mathscr{A}$ let

$\langle f_k \rangle_A$ denote the average of f_k on A. Since f_{n-1} and g_{n-1} are constant on each $B \in \mathscr{F}_{n-1}^*$ we find that

$$\begin{aligned}\int_S u(f_n, g_n)\,\mathrm{d}\mu &= \sum_{B\in\mathscr{F}_{n-1}^*} \int_B u(\langle f_{n-1}\rangle_B + df_n, \langle g_{n-1}\rangle_B + \epsilon_n df_n)\,\mathrm{d}\mu \\ &\leqslant \sum_{B\in\mathscr{F}_{n-1}^*} \mu(B) u(\langle f_{n-1}\rangle_B, \langle g_{n-1}\rangle_B) = \int_S u(f_{n-1}, g_{n-1})\,\mathrm{d}\mu,\end{aligned}$$

where in the estimate we applied Lemma 4.5.4 with the mean zero random variable df_n and probability measure $\mu|_B/\mu(B)$. □

4.5.b Burkholder's theorem

The main result of this section is the following theorem, due to Burkholder, which states that a Banach space X is a UMD space if and only if X is a UMD space with respect to the unit interval $[0,1)$, if and only if there exists a suitable function on $X \times X$ which satisfies the concavity property of Definition 4.5.3.

Theorem 4.5.6 (Burkholder). *For a Banach space X the following assertions are equivalent:*

(1) *X is a UMD space;*
(2) *X has the dyadic UMD property with respect to Paley–Walsh martingale differences on any fixed probability space which supports a real Rademacher sequence $(r_n)_{n\geqslant 1}$ and with filtration generated by $(r_n)_{n\geqslant 1}$;*
(3) *for some (equivalently, all) $p \in (1,\infty)$ there exist a constant $\beta \in [1,\infty)$ and a zigzag-concave function $u : X \times X \to \mathbb{R}$ which satisfies*

$$u(x,y) \geqslant \|y\|^p - \beta^p \|x\|^p, \quad x, y \in X. \tag{4.30}$$

Moreover, the infimum of all $\beta \in [1,\infty)$ for which such a function u exists coincides with the UMD constant $\beta_{p,X}$.

The equivalence of (1) and (3) is an instance of the general phenomenon that the validity of certain martingale inequalities (such as the inequality defining the UMD property) can often be characterised by the existence of a special function u, usually defined on X or $X \times X$, for which a certain special inequality holds (such as (4.30)). In such situations the function u will be referred to as a *Burkholder function* for the relevant martingale inequality.

To prepare for the proof of Theorem 4.5.6 a couple of elementary observations will be useful. Firstly, if there exists a zigzag-concave function u satisfying (4.30), then there also exists a zigzag-concave function $\widetilde{u}$ which, in addition to (4.30), satisfies

$$\widetilde{u}(\alpha x, \alpha y) = |\alpha|^p \widetilde{u}(x,y) \tag{4.31}$$

for all $\alpha \in \mathbb{K}$ and $x, y \in X$. Namely, define $u_\alpha(x,y) := |\alpha|^{-p} u(\alpha x, \alpha y)$ for $\alpha \neq 0$. This function is zigzag-concave and satisfies (4.30). Then

$$\widetilde{u} := \inf_{\alpha \neq 0} u_\alpha$$

has the same properties; here we appeal to the fact that the pointwise infimum of zigzag-concave functions is again zigzag-concave. Now $\widetilde{u}$ has the desired homogeneity property by a simple change of variable. But then also

$$\widetilde{u}(0 \cdot x, 0 \cdot y) = \widetilde{u}(0,0) = \widetilde{u}(\alpha \cdot 0, \alpha \cdot 0) = |\alpha|^p \widetilde{u}(0,0) \to 0 = |0|^p \widetilde{u}(x,y)$$

as $\alpha \to 0$, and the claim follows for all $\alpha \in \mathbb{R}$.

Secondly, we may impose that

$$u(x,y) = u(\epsilon_1 x, \epsilon_2 y) \tag{4.32}$$

for all signs $\epsilon_1, \epsilon_2 \in \mathbb{K}$. If this condition is not already satisfied, then the new function $\inf_{|\epsilon_1|=|\epsilon_2|=1} u(\epsilon_1 x, \epsilon_2 y)$ will satisfy it; this function is zigzag-concave and satisfies (4.30).

Having made these reductions, as a consequence we also have

$$u(\epsilon_1 z, \epsilon_2 z) = u(z,z) \leqslant 0 \tag{4.33}$$

for all $z \in X$ and all signs ϵ_1, ϵ_2. Indeed, the first identity holds by (4.32). To see that both are non-positive note that by the zigzag-concavity

$$u(z,z) = \frac{1}{2} u(0+z, 0+z) + \frac{1}{2} u(0-z, 0-z) \leqslant u(0,0) = 0. \tag{4.34}$$

In the remainder of this section we shall write

$$F_\beta^p(x,y) := \|y\|^p - \beta^p \|x\|^p \tag{4.35}$$

for the function on the right-hand side of (4.30).

Proof of Theorem 4.5.6, (3)⇒(1). Suppose that u is zigzag-concave and $u \geqslant F_\beta^p$ for some fixed $p \in (1,\infty)$ and $\beta \in [1,\infty)$. By approximation (see Lemma 3.6.17) and Lemma 4.2.8 it suffices to check the estimate

$$\int_S \|g_n\|^p - \beta^p \|f_n\|^p \,\mathrm{d}\mu = \int_S F_\beta^p(f_n, g_n) \,\mathrm{d}\mu \leqslant 0$$

for all μ-simple martingales f and their transforms $g = \epsilon \star f$, where $\epsilon = (\epsilon_n)_{n \geqslant 1}$ is a sequence of signs in $\mathbb{K}$.

By the assumption, Lemma 4.5.5, and (4.33),

$$\int_S F_\beta^p(f_n, g_n) \,\mathrm{d}\mu \leqslant \int_S u(f_n, g_n) \,\mathrm{d}\mu$$

$$\leqslant \int_S u(f_{n-1}, g_{n-1})\,\mathrm{d}\mu$$
$$\leqslant \ldots \leqslant \int_S u(f_0, g_0)\,\mathrm{d}\mu = \int_S u(f_0, f_0)\,\mathrm{d}\mu \leqslant 0.$$

This concludes the proof of the implication (3)⇒(1), with bound $\beta_{p,X} \leqslant \beta$. □

We complete the proof of Theorem 4.5.6 by constructing the zigzag-concave function from the dyadic UMD property on any given probability space supporting a real Rademacher sequence and with filtration generated by this sequence. The result provides an alternative proof of Theorem 4.2.5.

Completion of the proof of Theorem 4.5.6. Fix an arbitrary $p \in (1,\infty)$. We will show that (3) holds with $\beta = \beta^{\Delta}_{p,X}$, the dyadic UMD constant implicit in the formulation of (2). This will conclude the proof of the theorem, for the implication (1)⇒(2) is trivial, with $\beta^{\Delta}_{p,X} \leqslant \beta_{p,X}$, and we have already proved the implication (3)⇒(1) with inequality $\beta_{p,X} \leqslant \beta$.

So let us assume that (2) holds. More precisely, let Ω be a fixed probability space on which a real Rademacher sequence $(r_n)_{n\geqslant 1}$ is defined, and assume that X has the UMD property (with optimal constant denoted $\beta^{\Delta}_{p,X} =: \beta$) with respect to X-valued martingales adapted to the filtration $(\mathscr{F}_n)_{n\geqslant 0}$ generated by $(r_n)_{n\geqslant 1}$, i.e., $\mathscr{F}_0 := \{\varnothing, \Omega\}$ and $\mathscr{F}_n := \sigma(r_1, \ldots, r_n)$, $n \geqslant 1$.

For elements $x, y \in X$ we shall write

$$(f, g) \in \mathbb{S}(x, y)$$

to abbreviate the assertion that $f = (f_n)_{n=0}^{\infty}$ and $g = (g_n)_{n=0}^{\infty}$ are martingales on Ω with respect to the filtration $(\mathscr{F}_n)_{n\geqslant 0}$ with the following properties:

(i) $f_0 \equiv x$ and $g_0 \equiv y$;
(ii) f and g have finitely non-zero difference sequences;
(iii) $(dg_n)_{n\geqslant 1} = (\epsilon_n df_n)_{n\geqslant 1}$ for some sequence $(\epsilon_n)_{n=1}^{\infty}$ of signs in $\mathbb{K}$.

We further denote

$$f_\infty := \lim_{N\to\infty} f_N;$$

this limit exists since f_N is eventually constant.

Consider the function $U : X \times X \to \mathbb{R} \cup \{\infty\}$,

$$U(x, y) := \sup \Big\{ \mathbb{E} F^p_\beta(f_\infty, g_\infty) : (f, g) \in \mathbb{S}(x, y) \Big\}.$$

We will check that U enjoys the following properties:

(a) $U(x, y) < \infty$ for all $x, y \in X$;
(b) $U(\epsilon x, \widetilde{\epsilon} y) = U(x, y)$ for all signs $\epsilon, \widetilde{\epsilon}$ in $\mathbb{K}$;
(c) $U(x, y) \geqslant F^p_\beta(x, y)$;
(d) U is zigzag-concave in the sense of Definition 4.5.3.

The first condition is quite essential: it would be easy to fulfil the other three conditions by choosing $U \equiv \infty$.

Step 1 – We begin by showing that U enjoys properties (b) and (c). Clearly, $(f,g) \in \mathbb{S}(x,y)$ if and only if $(\epsilon f, \widetilde{\epsilon} g) \in \mathbb{S}(\epsilon x, \widetilde{\epsilon} y)$. Therefore, from $F_\beta^p(\epsilon f_\infty, \widetilde{\epsilon} g_\infty) = F_\beta^p(f_\infty, g_\infty)$ we infer that $U(\epsilon x, \widetilde{\epsilon} y) = U(x,y)$. Also, from $(x,y) \in \mathbb{S}(x,y)$ one obtains

$$U(x,y) \geqslant \mathbb{E} F_\beta^p(x,y) = F_\beta^p(x,y).$$

Step 2 – Next we prove that U is zigzag-concave, as required in (d). By Step 1, U is locally bounded from below. Therefore, by Lemma 4.5.2 it suffices to prove that U is midpoint zigzag-concave in the sense of Definition 4.5.3. We will allow that U takes the value ∞ in this part of the proof.

Given $x, y, z_1, z_2 \in X$ and a sign $\epsilon_0 \in \mathbb{K}$, with $z_0 := \frac{1}{2}(z_1 + z_2)$ and $x_i := x + z_i$, $y_i := y + \epsilon_0 z_i$ for $i = 0, 1, 2$ we need to show that

$$U(x_0, y_0) = U(x + z_0, y + \epsilon_0 z_0) \geqslant \frac{1}{2} \sum_{i=1}^{2} U(x + z_i, y + \epsilon_0 z_i) = \frac{1}{2} \sum_{i=1}^{2} U(x_i, y_i).$$

We fix $m_i < U(x_i, y_i)$ and claim that $U(x_0, y_0) > \frac{1}{2}(m_1 + m_2)$.

For $i = 1, 2$, pick $(f^i, g^i) \in \mathbb{S}(x_i, y_i)$ such that

$$\mathbb{E} F_\beta^p(f_\infty^i, g_\infty^i) > m_i.$$

We need to find a pair of martingales $(f,g) \in \mathbb{S}(x_0, y_0)$ such that

$$\mathbb{E} F_\beta^p(f_\infty, g_\infty) > \frac{1}{2}(m_1 + m_2).$$

Since f^i and g^i are dyadic martingales, they take the form

$$f_n^i = \phi_n^i(r_1, r_2, \dots, r_n), \qquad g_n^i = \psi_n^i(r_1, r_2, \dots, r_n)$$

for suitable functions $\phi_n^i, \psi_n^i : \{-1,1\}^n \to X$. We then define the auxiliary non-adapted sequences

$$\widetilde{f}_n^1 := \phi_n^1(r_2, r_4, \dots, r_{2n}), \qquad \widetilde{f}_n^2 := \phi_n^2(r_3, r_5, \dots, r_{2n+1}),$$

with an analogous definition of $\widetilde{g}_n^i$, $i = 1, 2$. Then let

$$\begin{aligned}
f_0 &:= x_0 = \frac{1}{2}(x_1 + x_2), \\
f_1 &:= x_0 + \frac{r_1}{2}(x_1 - x_2) = \frac{1 + r_1}{2} x_1 + \frac{1 - r_1}{2} x_2, \\
f_{2n} &:= \frac{1 + r_1}{2} \widetilde{f}_n^1 + \frac{1 - r_1}{2} \widetilde{f}_{n-1}^2, \qquad n \geqslant 1, \\
f_{2n+1} &:= \frac{1 + r_1}{2} \widetilde{f}_n^1 + \frac{1 - r_1}{2} \widetilde{f}_n^2, \qquad n \geqslant 1.
\end{aligned}$$

It is immediate that $(f_n)_{n\geqslant 1}$ is adapted. Moreover,

$$df_{2n} = \frac{1+r_1}{2} d\widetilde{f}_n^1, \qquad df_{2n+1} = \frac{1-r_1}{2} d\widetilde{f}_n^2, \qquad n \geqslant 1,$$

which are seen to be martingale differences from the fact that df_n^i are. And clearly $df_1 = \frac{1}{2} r_1(x_1 - x_2)$ is a martingale difference as well.

We define g_n in an analogous fashion, replacing x_i by y_i and $\widetilde{f}_n^i$ by $\widetilde{g}_n^i$ in the definition of f_n. Since $dg_n^i = \epsilon_n^i df_n^i$ for some signs ϵ_n^i, it follows that

$$dg_{2n} = \frac{1+r_1}{2} d\widetilde{g}_n^1 = \frac{1+r_1}{2} \epsilon_n^1 d\widetilde{f}_n^1 = \epsilon_n^1 df_{2n}, \qquad n \geqslant 1$$

and similarly $dg_{2n+1} = \epsilon_n^2 df_{2n+1}$ for $n \geqslant 1$. Moreover, recalling that $x_i = x + z_i$ and $y_i = y + \epsilon_0 z_i$, we find that

$$y_1 - y_2 = \epsilon_0(z_1 - z_2) = \epsilon_0(x_1 - x_2),$$

and hence

$$dg_1 = \frac{r_1}{2}(y_1 - y_2) = \frac{\epsilon_0 r_1}{2}(x_1 - x_2) = \epsilon_0 df_1.$$

So altogether, $dg_n = \epsilon_n' df_n$, where $\epsilon_1' = \epsilon_0$ and $\epsilon_{2n}' = \epsilon_n^1$, $\epsilon_{2n+1}' = \epsilon_n^2$ for $n \geqslant 1$.

We conclude that $(f, g) \in \mathbb{S}(x_0, y_0)$. Moreover,

$$\begin{aligned}(f_\infty, g_\infty) &= \frac{1+r_1}{2}(\widetilde{f}_\infty^1, \widetilde{g}_\infty^1) + \frac{1-r_1}{2}(\widetilde{f}_\infty^2, \widetilde{g}_\infty^2) \\ &= \mathbf{1}_{\{r_1=1\}}(\widetilde{f}_\infty^1, \widetilde{g}_\infty^1) + \mathbf{1}_{\{r_1=-1\}}(\widetilde{f}_\infty^2, \widetilde{g}_\infty^2).\end{aligned}$$

Using the fact that r_1 is independent of $(\widetilde{f}_\infty^i, \widetilde{g}_\infty^i)$, which is equidistributed with (f_∞^i, g_∞^i), we deduce that

$$\begin{aligned}U(x_0, y_0) &\geqslant \mathbb{E} F_\beta^p(f_\infty, g_\infty) \\ &= \mathbb{E}[\mathbf{1}_{\{r_1=1\}} F_\beta^p(\widetilde{f}_\infty^1, \widetilde{g}_\infty^1)] + \mathbb{E}[\mathbf{1}_{\{r_1=-1\}} F_\beta^p(\widetilde{f}_\infty^2, \widetilde{g}_\infty^2)] \\ &= \mathbb{E}[\mathbf{1}_{\{r_1=1\}}]\mathbb{E}[F_\beta^p(\widetilde{f}_\infty^1, \widetilde{g}_\infty^1)] + \mathbb{E}[\mathbf{1}_{\{r_1=-1\}}]\mathbb{E}[F_\beta^p(\widetilde{f}_\infty^2, \widetilde{g}_\infty^2)] \\ &= \frac{1}{2}\Big(\mathbb{E}[F_\beta^p(f_\infty^1, g_\infty^1)] + \mathbb{E}[F_\beta^p(f_\infty^2, g_\infty^2)]\Big) > \frac{1}{2}(m_1 + m_2).\end{aligned}$$

Since $m_i < U(x_i, y_i)$ were arbitrary, this concludes the proof of the zigzag-concavity of U, as required by property (d).

Step 3 – Up to this point, the UMD property has not been used in the consideration of the properties of U. It enters in the proof of the finiteness of U. The central step in proving finiteness is to show that

$$U(0, 0) \leqslant 0.$$

Let $f, g \in \mathbb{S}(0, 0)$. Then $g = \epsilon \star f$ for some sequence $\epsilon = (\epsilon_n)_{n\geqslant 1}$ of signs in $\mathbb{K}$. Choose $N \in \mathbb{N}$ such that $df_n = 0$ for all $n > N$. By the dyadic UMD property assumed in (2),

$$\mathbb{E}F^p_\beta(f_\infty, g_\infty) = \mathbb{E}\big(\|g_N\|^p - \beta^p\|f_N\|^p\big) \leqslant 0$$

recalling that we took $\beta = \beta^{\Delta}_{p,X}$. Taking the supremum over all $(f,g) \in \mathbb{S}(0,0)$, the required result follows.

The proof of Step 3 is finished by simple concavity considerations. We begin by noting that for all $x \in X$ we have $U(x,x) \leqslant 0$. Indeed, using the properties of U, this can be seen as in (4.34). Next we check that $U(x,y) < \infty$ by using the zigzag-concavity. Writing $a = \frac{x+y}{2}$, $b_1 = \frac{x-y}{2}$, $b_2 = -b_1$ and $\epsilon = -1$ we find

$$U(x,y) + U(y,x) = U(a+b_1, a+\epsilon b_1) + U(a+b_2, a+\epsilon b_2) \leqslant 2U(a,a) \leqslant 0.$$

Therefore, both terms on the left hand side must be finite. □

4.5.c Optimal constants for the real line

In this section we shall determine the exact value of $\beta_{p,\mathbb{R}}$.

Theorem 4.5.7 (Burkholder). *For all $p \in (1,\infty)$,*

$$\beta_{p,\mathbb{R}} = p^* - 1,$$

where $p^* := \max\{p, p'\}$.

For the Haar system defined in Subsection 4.2.b, this implies:

Corollary 4.5.8. *The system of Haar functions $(h_I)_{I\in\mathscr{D}^0([0,1))}$ is an unconditional basis of $L^p(0,1;\mathbb{R})$ with unconditionality constant $p^* - 1$.*

Proof. This follows at once from Corollary 4.5.15 and Theorem 4.5.7. □

The complex versions of Theorem 4.5.7 and Corollary 4.5.16 are also true, but require a non-trivial extension of the argument, which we present in Subsection 4.5.d.

Search for the auxiliary function

By Theorem 4.5.6, one should find a function $u : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ such that

(i) for all $\epsilon \in \{-1,1\}$ and $x,y \in \mathbb{R}$, $z \mapsto u(x+z, y+\epsilon z)$ is concave;
(ii) for all $x,y \in \mathbb{R}$, $u(x,y) \geqslant F(x,y) = |y|^p - \beta^p|x|^p$.

Since we are presently working over the reals, (i) is precisely the assertion that u is zigzag-concave (cf. Definition 4.5.3). In (ii), $\beta > 0$ is arbitrary but fixed for the moment; later on we will set $\beta = p^* - 1$. By (4.31), one may additionally demand that

(iii) $u(\epsilon x, \widetilde{\epsilon} y) = u(x,y)$ for all $\epsilon, \widetilde{\epsilon} \in \{-1,1\}$ and $x,y \in \mathbb{R}$;
(iv) $u(\alpha x, \alpha y) = |\alpha|^p u(x,y)$ for all $\alpha \in \mathbb{R}$ and $x,y \in \mathbb{R}$.

In order to find a function with the properties (i)-(iv), we will reduce the search to finding a function $w : \mathbb{R} \to \mathbb{R}$ with certain properties. A possible function u will be eventually given in (4.38). An independent proof of the properties of u (and other properties) will be given in Lemmas 4.5.19, 4.5.20, 4.5.21 and 4.5.22.

If u has the desired properties, then for $x \neq -y$ we have, putting $t := |x+y|$ and $s := \frac{y-x}{x+y}$,

$$\begin{aligned} u(x,y) &= u\Big(\frac{|x+y|}{x+y}x, \frac{|x+y|}{x+y}y\Big) \\ &= u\Big(\frac{t(1-s)}{2}, \frac{t(1+s)}{2}\Big) = t^p u\Big(\frac{1-s}{2}, \frac{1+s}{2}\Big), \end{aligned}$$

where we applied (iii) and (iv). This motivates us to define the function $w : \mathbb{R} \to \mathbb{R}$ by

$$w(s) := u\Big(\frac{1-s}{2}, \frac{1+s}{2}\Big).$$

Throughout the rest of this subsection we also let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(s) := F\Big(\frac{1-s}{2}, \frac{1+s}{2}\Big) = \Big|\frac{1+s}{2}\Big|^p - \beta^p \Big|\frac{1-s}{2}\Big|^p,$$

where $\beta > 0$ is arbitrary for the moment and will be fixed later on. Now the properties (i)-(iv) for u imply the following properties for w:

(a) w is concave;
(b) for all $s \in \mathbb{R}$, $w(s) \geqslant f(s)$;
(c) for all $s \neq 0$, $w(1/s) = |s|^{-p} w(s)$.

Property (a) follows from (i) and (b) follows from (ii). The identity in (c) follows from (iii) and (iv):

$$w(1/s) = u\Big(\frac{s-1}{2s}, \frac{s+1}{2s}\Big) = |s|^{-p} u\Big(\frac{s-1}{2}, \frac{s+1}{2}\Big) = |s|^{-p} w(s).$$

Conversely, if w is a function verifying the properties (a)-(c), one may define

$$u(x,y) := \begin{cases} |y-x|^p w\Big(\dfrac{x+y}{y-x}\Big), & \text{if } x \neq y, \\ |2x|^p w(0), & \text{if } x = y. \end{cases}$$

Let us prove that then u again satisfies the properties (i)-(iv), so that searching for w is completely equivalent to searching for u.

First we show that the symmetry properties (iii) and (iv) hold: $u(\epsilon x, \widetilde{\epsilon} y) = u(x,y)$ and $u(\alpha x, \alpha y) = |\alpha|^p u(x,y)$.

Proof. To prove the property (iii), since u is even it suffices to consider $\epsilon = -1$ and $\widetilde{\epsilon} = 1$. If $x \neq y$ and $x \neq -y$, then (c) implies

$$u(-x, y) = |x + y|^p w\Big(\frac{y - x}{x + y}\Big) = |y - x|^p w\Big(\frac{x + y}{y - x}\Big) = u(x, y).$$

If $x = y$ or $x = -y$, then

$$u(-x, y) = |2x|^p w(0) = u(x, y).$$

Turning to the proof of (iv), if $\alpha \neq 0$ and $x \neq y$, then

$$u(\alpha x, \alpha y) = |\alpha y - \alpha x|^p w\Big(\frac{\alpha x + \alpha y}{\alpha y - \alpha x}\Big) = |\alpha|^p |y - x|^p w\Big(\frac{x + y}{y - x}\Big) = |\alpha|^p u(x, y).$$

If $x = y$, then

$$u(\alpha x, \alpha y) = |\alpha|^p |2x|^p w(0) = |\alpha|^p u(x, y).$$

Finally if $\alpha = 0$ then $u(0, 0) = |2 \cdot 0|^p w(0) = 0$, so $u(0 \cdot x, 0 \cdot y) = 0 = |0|^p u(x, y)$. □

Another basic observation is the inequality $w(0) \leqslant 0$.

Proof. Assume, contrary to the claim, that $w(0) > 0$. Since a concave function is continuous, we have $w(t) \geqslant \delta > 0$ for all $|t| \leqslant \varepsilon$. Thus

$$w(t^{-1}) = |t|^{-p} w(t) \geqslant \delta |t|^{-p} \to \infty \quad \text{as } t \to 0.$$

In particular, for all M there is an N, such that $w(x) \geqslant M$, when $|x| \geqslant N$. But by concavity we then have $w(x) \geqslant M$ also on $(-N, N)$. Thus $w \geqslant M$ everywhere, and by the arbitrariness of M it follows that $w \equiv +\infty$. This contradicts the real-valuedness of w. □

From the previous observation it follows that for $x = y$, the function $z \mapsto u(x + z, y + z) = |2(x + z)|^p w(0)$ is concave. For $x \neq y$, this follows directly from the concavity of w. By the symmetry properties of u, the concavity of $z \mapsto u(x + z, y - z) = u(x + z, -y + z)$ follows as well. We conclude that u enjoys property (i), i.e., u is zigzag-concave.

To conclude we prove that u has property (ii). If $x \neq -y$, then with $s = \frac{y - x}{x + y}$ we find

$$u(x, y) = u(-x, y) = |x + y|^p w(s) \geqslant |x + y|^p F\Big(\frac{1 - s}{2}, \frac{1 + s}{2}\Big) = F(x, y),$$

where we used property (i) for u, the inequality $w \geqslant f$, and the homogeneity properties of the function F. Finally,

$$u(x, -x) = u(x, x) = |2x|^p w(0) \geqslant |2x|^p F(\tfrac{1}{2}, \tfrac{1}{2}) = F(x, -x).$$

The search for w: miscellanea

Whether or not it is possible to find $w \geqslant f$ with the required properties depends on the function f, and then on the constant β in its definition. Let us note that it is necessary that $\beta \geqslant 1$. Indeed, if $\beta < 1$, then $w(x) \geqslant f(x) \to \infty$ as $x \to \pm\infty$, and this is impossible for a concave function w, as we saw before.

The function f, and then the problem of finding a dominating w, is a little different depending on whether $p \in (1,2)$, $p = 2$, or $p \in (2,\infty)$. Let us notice that the case $p = 2$, unlike the other two, is trivial. (This is not surprising, since Burkholder's inequality itself is also easy in this case.) This is because

$$f(x) = \Big(\frac{x+1}{2}\Big)^2 - \beta^2\Big(\frac{x-1}{2}\Big)^2 = \frac{1}{4}\Big((1-\beta^2)x^2 + 2(1+\beta^2)x + (1-\beta^2)\Big)$$

is already concave for all $\beta \geqslant 1$. With the minimal choice $\beta = 1$, it follows that $f(x) = x$, and one can take $w(x) = x$, $u(x,y) = y^2 - x^2$.

In the sequel, we will concentrate on the case $p \in (2,\infty)$, leaving the treatment of $p \in (1,2)$ as an exercise. Note that, just for the proof of Burkholder's inequality, one would not need to repeat all the concave function constructions, but one could derive the inequality for $p \in (1,2)$ from the case $p \in (2,\infty)$ by a duality argument.

The shape of the function f

Let us compute the derivatives

$$\begin{aligned} f(x) &= \Big|\frac{x+1}{2}\Big|^p - \beta^p\Big|\frac{x-1}{2}\Big|^p, \\ f'(x) &= \frac{p}{2}\Big(\operatorname{sgn}(x+1)\Big|\frac{x+1}{2}\Big|^{p-1} - \beta^p \operatorname{sgn}(x-1)\Big|\frac{x-1}{2}\Big|^{p-1}\Big), \\ f''(x) &= \frac{p(p-1)}{4}\Big(\Big|\frac{x+1}{2}\Big|^{p-2} - \beta^p\Big|\frac{x-1}{2}\Big|^{p-2}\Big). \end{aligned}$$

For $k = 0, 2$, it follows that

$$\begin{aligned} f^{(k)}(x) = 0 \ &\Leftrightarrow \ \Big|\frac{x+1}{2}\Big|^{p-k} = \beta^p\Big|\frac{x-1}{2}\Big|^{p-k} \\ &\Leftrightarrow \ |x+1| = \beta_k|x-1|, \quad \beta_k := \beta^{p/(p-k)}. \end{aligned}$$

For $k = 1$, one has the same condition as above, and in addition $\operatorname{sgn}(x+1) = \operatorname{sgn}(x-1)$, i.e., $|x| > 1$. Solving for x,

$$\begin{aligned} |x+1| = \beta_k|x-1| \ &\Leftrightarrow \ \begin{cases} x+1 = \beta_k(1-x) \\ x+1 = \beta_k(x-1) \end{cases} \\ &\Leftrightarrow \ \begin{cases} x = x_k \\ x = 1/x_k \end{cases} \quad x_k := \frac{\beta_k - 1}{\beta_k + 1} \in [0,1). \end{aligned}$$

Since $\beta \geqslant 1$ and $p > 2$, hence $1 = p/(p-0) < p/(p-1) < p/(p-2)$, it follows that $1 \leqslant \beta_0 \leqslant \beta_1 \leqslant \beta_2$ and then $0 \leqslant x_0 \leqslant x_1 \leqslant x_2 < 1 < 1/x_2 \leqslant 1/x_1 \leqslant 1/x_0$, and all inequalities are strict if $\beta > 1$. It will be shown shortly that this is the case.

So $f^{(k)}$ has zeroes at x_k and $1/x_k$ for $k = 0, 2$ and only at $1/x_1$ for $k = 1$. A routine investigation of the signs reveals that $f^{(k)}$ is positive on $(x_k, 1/x_k)$ and negative on $\complement[x_k, 1/x_k]$ for $k = 0, 2$, while f' is positive on $(-\infty, 1/x_1)$ and negative on $(1/x_1, \infty)$. The fact that $f'' > 0$ on $(x_2, 1/x_2)$ shows that f itself does not qualify for w, since it fails to be concave on that interval.

The next lemma proves that $\beta \geqslant p - 1$:

Lemma 4.5.9. *Let $p \in [2, \infty)$. If there exists a concave $w \geqslant f$ with $w(x) = |x|^p w(1/x)$, then $\beta \geqslant p - 1$.*

Proof. Note that $x_0 = \frac{\beta-1}{\beta+1}$, so $0 = F(\frac{2}{\beta+1}, \frac{2\beta}{\beta+1}) = f(x_0) \leqslant w(x_0)$ and $2^p = f(1) \leqslant w(1)$.

For $x \in (0, 1)$, consider the difference quotient

$$\frac{w(1) - w(x)}{1 - x} = \frac{w(1) - x^p w(1/x)}{1 - x} = w(1)\frac{1 - x^p}{1 - x} + x^{p-1}\frac{w(1) - w(1/x)}{1/x - 1}.$$

A concave function has one-sided derivatives $D_{\pm}w(x) = \lim_{y \to x\pm}(w(y) - w(x))/(y - x)$ at every point. So in particular, taking the limit $x \nearrow 1$ above, it follows that

$$D_{-}w(1) = pw(1) - D_{+}w(1).$$

By the condition $w(x_0) \geqslant 0$ and concavity, for all $x \in (x_0, 1)$ we have

$$\begin{aligned} \frac{w(1)}{1 - x_0} \geqslant \frac{w(1) - w(x_0)}{1 - x_0} &\geqslant \frac{w(1) - w(x)}{1 - x} \\ &\geqslant D_{-}w(1) \geqslant \frac{1}{2}\big(D_{-}w(1) + D_{+}w(1)\big) = \frac{p}{2}w(1). \end{aligned}$$

Dividing by $w(1) > 0$, it follows that $p/2 \leqslant 1/(1 - x_0) = (\beta + 1)/2$, and this completes the proof. $\square$

The form of w on $[x_0, 1]$

From now on, we fix $\beta := p - 1 = p^* - 1$ and try to find a dominating function w. With this choice of β, the left and the right sides of $(*)$ in the proof of Lemma 4.5.9 are equal, and hence we obtain a chain of equalities for $x \in [x_0, 1]$. This gives various pieces of useful information.

First, we must have $w(x_0) = 0$ and $D_{-}w(1) = D_{+}w(1)$. Hence the derivative $w'(1)$ exists and equals $w'(1) = p/2 \cdot w(1)$. Finally, from the equality $(w(1) - w(x))/(1 - x) = p/2 \cdot w(1)$, it follows that

$$w(x) = \alpha_p\Big(\frac{p}{2}x - \big(\frac{p}{2} - 1\big)\Big), \quad x \in [x_0, 1] = [1 - \frac{2}{p}, 1], \qquad \alpha_p := w(1).$$

So w is an affine function on this interval, and it remains to determine the constant α_p. The following Lemma 4.5.10 (noting that $w(x_0) = 0 = f(x_0)$) implies that we must have

$$\begin{aligned}\frac{p}{2}\alpha_p = w'(x_0) = f'(x_0) &= \frac{p}{2}\Big(\big(\frac{x_0+1}{2}\big)^{p-1} + \beta^p\big(\frac{1-x_0}{2}\big)^{p-1}\Big)\\ &= \frac{p}{2}\Big(\big(\frac{p-1}{p}\big)^{p-1} + (p-1)^p\big(\frac{1}{p}\big)^{p-1}\Big)\\ &= \frac{p}{2}\Big(\frac{p-1}{p}\Big)^{p-1}\big(1+(p-1)\big),\end{aligned}$$

and hence

$$\alpha_p = p\Big(\frac{p-1}{p}\Big)^{p-1}.$$

The lemma still has to be proved:

Lemma 4.5.10. *Let w be concave, $w \geqslant f$ and $w(x_0) = f(x_0)$ at a point x_0 at which f is differentiable. Then w is also differentiable at x_0 and $w'(x_0) = f'(x_0)$.*

Proof. Let $h > 0$. Then

$$\begin{aligned}0 &\leqslant w(x_0 \pm h) - f(x_0 \pm h)\\ &= [w(x_0) \pm hD_\pm w(x_0) + o(h)] - [f(x_0) \pm hf'(x_0) + o(h)]\\ &= \pm h[D_\pm w(x_0) - f'(x_0)] + o(h).\end{aligned}$$

For the + case, $D_+w(x_0) < f'(x_0)$ would lead to a contradiction, since then the right-hand side in the above would be negative for small $h > 0$. Similarly, $D_-w(x_0) > f'(x_0)$ would lead to a contradiction in the $-$ case. Hence

$$D_-w(x_0) \leqslant f'(x_0) \leqslant D_+w(x_0) \leqslant D_-w(x_0),$$

where the last inequality is due to concavity. Thus all the expressions are equal. □

Checking that the candidate for w works

We have seen that if a function w with the required properties exists, it has to be $w(x) = \alpha_p\big(p/2 \cdot x - (p/2-1)\big)$ on the interval $[x_0, 1]$. We check next that this function satisfies the condition $w \geqslant f$ on this interval.

The second derivative f'' is negative on (x_0, x_2) and positive on $(x_2, 1)$, thus f' is decreasing on the first interval and increasing on the second. Hence, on the whole interval,

$$\begin{aligned}f' \leqslant \max\{f'(x_0), f'(1)\} &= \max\{w'(x_0), p/2\}\\ &= \max\{\alpha_p, 1\}\cdot p/2 = \alpha_p \cdot p/2 \equiv w'.\end{aligned}$$

The estimate $\alpha_p \geqslant 1$, which was used above, is left as an exercise.

So $(f-w)' \leqslant 0$ and hence $f - w$ is decreasing on $[x_0, 1]$, thus $f - w \leqslant (f-w)(x_0) = 0$ on $[x_0, 1]$. This completes the check.

First choice: minimal function w

With the function w determined on $[x_0, 1]$, it is also determined on $[1, 1/x_0]$ due to the requirement that $w(x) = |x|^p w(1/x)$. This gives

$$\begin{aligned} w(x) &= \alpha_p\Big(\frac{p}{2}x^{p-1} - \big(\frac{p}{2} - 1\big)x^p\Big), \qquad x \in [1, 1/x_0] = [1, \frac{p}{p-2}] \\ w'(x) &= \alpha_p\Big(\frac{p}{2}(p-1)x^{p-2} - \frac{p-2}{2}px^{p-1}\Big), \\ w''(x) &= \frac{1}{2}\alpha_p p(p-1)(p-2)x^{p-3}(1-x) \leqslant 0, \end{aligned}$$

since $p > 2$ and $x \geqslant 1$. So the function is concave on $[1, 1/x_0]$.

To check that w, which is concave on $[x_0, 1]$ and $[1, 1/x_0]$, is actually concave on the whole interval $[x_0, 1/x_0]$, one has to compare the one-sided derivatives at 1. But by the above,

$$D_+w(1) = \alpha_p\Big(\frac{p}{2}(p-1) - \frac{p}{2}(p-2)\Big) = \alpha_p\frac{p}{2} = D_-w(1),$$

so this is fine.

Moreover, since $w(x) \geqslant f(x)$ on $[x_0, 1]$, it follows at once that

$$w(x) = x^p w(x^{-1}) \geqslant x^p f(x^{-1}) = f(x)$$

also for $x \in [1, 1/x_0]$. In particular, $w(1/x_0) = f(1/x_0) = 0$. Moreover,

$$-x^{-2}w'(x^{-1}) = D[w(x^{-1})] = D[x^{-p}w(x)] = -px^{-p-1}w(x) + x^{-p}w'(x).$$

Since the same computation holds with f in place of w, and since $w^{(k)}(x_0) = f^{(k)}(x_0)$ for $k = 0, 1$, it follows that also $w'(1/x_0) = f'(1/x_0)$.

With the function w defined on $[x_0, 1/x_0]$ as above, it is seen that the following function provides the required concave dominant of f on all $\mathbb{R}$:

$$\widetilde{w}(x) := \begin{cases} w(x), & x \in [x_0, 1/x_0] = [\frac{p-2}{p}, \frac{p}{p-2}] \\ f(x), & \text{otherwise.} \end{cases}$$

Indeed, we have seen that it is concave on the intervals $(-\infty, x_0]$, $[x_0, 1/x_0]$ and $[1/x_0, \infty)$ and the one-sided derivatives agree at x_0 and $1/x_0$, so it is actually concave on $\mathbb{R}$. The function $\widetilde{w}$ also dominates f, and it satisfies $\widetilde{w}(x) = |x|^p\widetilde{w}(1/x)$, since both f and w do. This completes the search for the function w, and hence the proof of Burkholder's inequality in the case $p \in (2, \infty)$.

The function $\widetilde{w}$ is actually the smallest possible solution of the problem. Indeed, on the interval $[x_0, 1/x_0]$ the function was uniquely determined, and on the rest of $\mathbb{R}$, where there could be some freedom, it clearly cannot be smaller, since otherwise it would not dominate f.

Second choice for w

The function w is not uniquely determined outside $[x_0, 1/x_0]$, so if we do not care about having the minimal solution, there is some freedom of choice. Consider the function

$$w(x) := \alpha_p\Big(\Big|\frac{x+1}{2}\Big| - (p-1)\Big|\frac{x-1}{2}\Big|\Big)\max\{1, |x|\}^{p-1} \tag{4.36}$$

$$= \alpha_p \cdot \begin{cases} -\dfrac{p-2}{2}|x|^p - \dfrac{p}{2}|x|^{p-1}, & x \in (-\infty, -1], \\ \dfrac{p}{2}x - \dfrac{p-2}{2}, & x \in [-1, 1], \\ -\dfrac{p-2}{2}x^p + \dfrac{p}{2}x^{p-1}, & x \in [1, \infty), \end{cases} \tag{4.37}$$

which coincides with the earlier definition on the interval $[x_0, 1/x_0]$. It also satisfies the property $w(x) = |x|^p w(1/x)$, as one readily checks.

Clearly w is concave on $[-1, 1]$, and the proof above for the concavity on $[1, 1/x_0]$ actually works for $[1, \infty)$, since the function has the same expression on this whole half-line. For $x \in (-\infty, -1)$,

$$w''(x) = -\frac{1}{2}\alpha_p p(p-1)(p-2)\big(|x|^{p-2} + |x|^{p-3}\big) \leqslant 0,$$

so it is also concave there. The identity $D_- w(1) = D_+ w(1)$ has already been checked above, and

$$\begin{aligned} D_- w(-1) &= \alpha_p \frac{1}{2}\Big((p-2)p + p(p-1)\Big) \\ &= \alpha_p \frac{p}{2}(2p-3) > \alpha_p \frac{p}{2}(4-3) = D_+ w(-1). \end{aligned}$$

So the derivative has a discontinuity, but the inequality is in the right direction for concavity.

It remains to be shown that $w \geqslant f$ everywhere. This was already done for $[x_0, 1/x_0]$. As for $[-1, x_0]$, we have $f'' \leqslant 0$, and hence f' is decreasing, thus $f' \geqslant f'(x_0) = w'(x_0) \equiv w'$. Then $f - w$ is increasing, thus $f - w \leqslant (f-w)(x_0) = 0$, on $[-1, x_0]$. Finally, the mapping $x \mapsto x^{-1}$ transforms $(-\infty, 1]$ onto $[-1, 0)$ and $[1/x_0, \infty)$ onto $(0, x_0]$ so the domination on these remaining intervals follows as in $(*)$ of the previous paragraph.

The zigzag-concave function u

With simple algebra one checks that

$$\max\{1, |x|\} = \Big|\frac{x+1}{2}\Big| + \Big|\frac{x-1}{2}\Big|,$$

and hence the function w from (4.36) equals

$$w(x) = \alpha_p\Big(\Big|\frac{x+1}{2}\Big| - \beta_p\Big|\frac{x-1}{2}\Big|\Big)\Big(\Big|\frac{x+1}{2}\Big| + \Big|\frac{x-1}{2}\Big|\Big)^{p-1},$$

where $\beta_p = p-1$ and $\alpha_p = p\big(\frac{p-1}{p}\big)^{p-1}$. From here one also gets the required zigzag-concave function or Burkholder function u, the search for which was reduced to the search for w:

$$u(x,y) = |y-x|^p w\Big(\frac{x+y}{y-x}\Big) = \alpha_p\big(|y| - \beta_p|x|\big)\big(|y|+|x|\big)^{p-1}. \tag{4.38}$$

Note that these formulae also work for $p=2$, since then $\alpha_2 = \beta_2 = 1$ and

$$w(x) = \Big|\frac{x+1}{2}\Big|^2 - \Big|\frac{x-1}{2}\Big|^2 = x, \qquad u(x,y) = y^2 - x^2.$$

As it turns out, they also work for $p \in (1,2)$, provided that one chooses the constants α_p and β_p appropriately. The definitions

$$\alpha_p := p\Big(1 - \frac{1}{p^*}\Big)^{p-1}, \qquad \beta_p := p^* - 1$$

are good for all $p \in (1,\infty)$, where $p^* = \max\{p, p'\}$.

4.5.d Differential subordination

In Theorem 4.5.7 we computed the UMD constants of the real line to be $\beta_{p,\mathbb{R}} = p^* - 1$. As we shall see in this section, the constant $p^* - 1$ features in several martingale inequalities in Hilbert spaces. It will follow from these that $\beta_{p,\mathbb{C}} = p^* - 1$ and, in fact, $\beta_{p,H} = p^* - 1$ for any real or complex Hilbert space H.

Definition 4.5.11. *Let X and Y be Banach spaces. A Y-valued martingale g is* differentially subordinate *to an X-valued martingale f if almost everywhere we have that $g_{-\infty}$ and $f_{-\infty}$ exist and*

$$\|g_{-\infty}\|_Y \leqslant \|f_{-\infty}\|_X, \qquad \|dg_n\|_Y \leqslant \|dg_n\|_X \quad \forall n \in \mathbb{Z}.$$

Recall that if f and g are L^p-martingales with $1 \leqslant p < \infty$, the limits $f_{-\infty} = \lim_{n\to\infty} f_n$ and $g_{-\infty} = \lim_{n\to\infty} g_n$ exist, both in L^p and almost everywhere (see Corollary 3.3.7).

Example 4.5.12 (Martingale transforms). Let $(v_n)_{n\in\mathbb{Z}\cup\{-\infty\}}$ be an $\mathscr{L}(X,Y)$-valued strongly predictable sequence satisfying $\|v_n\|_\infty \leqslant 1$ for $n \in \mathbb{Z}\cup\{-\infty\}$. If f is a martingale with finitely non-zero difference sequence $(df_n)_{n\in\mathbb{Z}}$, the associated martingale transform

$$g := T_v f := v_{-\infty} f_{-\infty} + \sum_{n\in\mathbb{Z}} v_n df_n$$

is differentially subordinate to f.

Not every differentially subordinated martingale is given by a martingale transform:

Example 4.5.13. Let $S = \{s_1, s_2, s_3\}$ with $\mu(\{s_j\}) = 1/3$ for every j. Let $\mathscr{F}_n$ be the trivial σ-algebra for $n \leqslant 0$ and $\mathscr{F}_n$ the power set of S for $n \geqslant 1$. Let f and g be the martingales with $f_{-\infty} = g_{-\infty} = 0$, $df_n = dg_n = 0$ for $n \neq 1$, and only one non-zero difference given by

$$df_1 = \mathbf{1}_{\{s_1\}} + \mathbf{1}_{\{s_2\}} - 2 \cdot \mathbf{1}_{\{s_3\}} \quad \text{and} \quad dg_1 = \mathbf{1}_{\{s_1\}} - \mathbf{1}_{\{s_2\}}.$$

Then $|dg_1| \leqslant |df_1|$. On the other hand, if v_1 is $\mathscr{F}_0$-measurable, then it is constant. But clearly there does not exist a constant v_1 such that $dg_1 = v_1 df_1$.

If X and Y are Hilbert spaces and a martingale $g \in L^2(S;Y)$ is differentially subordinate to a martingale $f \in L^2(S;X)$ with finitely non-zero difference sequence, then by the orthogonality of martingale differences we have

$$\|g\|_2^2 = \|g_{-\infty}\|_2^2 + \sum_{n\in\mathbb{Z}} \|dg_n\|_2^2 \leqslant \|f_{-\infty}\|_2^2 + \sum_{n\in\mathbb{Z}} \|df_n\|_2^2 = \|f\|_2^2.$$

The main result of this section says that this extends to pth moments in the following way.

Theorem 4.5.14 (Burkholder). *Let X and Y be Hilbert spaces, let $p \in (1,\infty)$, let $f := (f_n)_{n\in\mathbb{Z}}$ be an L^p-martingale with values in X and let $g := (g_n)_{n\in\mathbb{Z}}$ be an L^p-martingale with values in Y. If g is differentially subordinate with respect to f, then*

$$\|g_n\|_{L^p(S;Y)} \leqslant (p^* - 1)\|f_n\|_{L^p(S;X)}, \quad n \in \mathbb{Z}.$$

The constant $p^* - 1$ is the best possible.

If, in addition, f is a bounded L^p-martingale, then also g is L^p-bounded. Then, by Theorem 3.3.16, both f_∞ and g_∞ exist, with convergence both in L^p and in the almost everywhere sense, and $\|g_\infty\|_p \leqslant (p^* - 1)\|f_\infty\|_{L^p(S;X)}$.

In combination with Example 4.5.12 this leads to the following result for martingale transforms in Hilbert spaces.

Corollary 4.5.15. *Let X and Y be (real or complex) Hilbert spaces, and $v = (v_n)_{n\in\mathbb{Z}\cup\{-\infty\}}$ be a strongly predictable $\mathscr{L}(X,Y)$-valued sequence with $\|v\|_\infty \leqslant 1$. Then the associated martingale transform satisfies*

$$\|T_v\|_{\mathscr{L}(L^p(S;X),L^p(S;Y))} \leqslant p^* - 1.$$

This constant is best possible. In particular, the UMD constants of any (real or complex) Hilbert space H are given by

$$\beta_{p,H} = \beta_{p,\mathbb{C}} = \beta_{p,\mathbb{R}} = p^* - 1.$$

Proof. If $f = (f_n)_{n\in\mathbb{Z}}$ is an L^p-martingale, then its transform by T_v (under the stated assumptions on v) is an L^p-martingale differentially subordinate to f by Example 4.5.12. Thus the claimed norm bound for T_v is a direct consequence of Theorem 4.5.14.

Specialising to scalar transforming sequences taking values in the unit sphere of $\mathbb{K}$, it follows in particular that $\beta_{p,H} \leqslant p^* - 1$ for any Hilbert space H. In the converse direction, from Proposition 4.2.10 we infer that $\beta_{p,H} \geqslant \beta^{\mathbb{R}}_{p,H} \geqslant \beta_{p,\mathbb{R}}$. The identity $\beta_{p,\mathbb{R}} = p^* - 1$ has been proved in Theorem 4.5.7. □

The next result refers to the system of Haar functions defined in Subsection 4.2.b and complements the real-valued result of Corollary 4.5.8.

Corollary 4.5.16. *The system of Haar functions $(h_I)_{I\in\mathscr{D}^0([0,1))}$ is an unconditional basis of $L^p(0,1;\mathbb{C})$ with unconditionality constant p^*-1.*

Proof. This follows at once from Corollary 4.5.15 and Theorem 4.2.13. □

The following simple example shows that neither Theorem 4.5.14 nor Corollary 4.5.15 hold for $X = Y = \ell^p$ with $p \neq 2$.

Example 4.5.17. Let $X = \ell^p$ with $p \in [1,\infty)$, with standard unit bases $(e_k)_{k\geqslant 1}$, and let $(\varepsilon_n)_{n\geqslant 0}$ be a Rademacher sequence. Let f be the ℓ^p-valued martingale whose differences are given by

$$df_n := \varepsilon_n e_1, \quad n \geqslant 0$$

and $f_0 := 0$, and let g be the martingale with martingale differences

$$dg_n := \varepsilon_n e_n$$

and $g_0 := 0$. Clearly, $\|df_n\|_p = \|dg_n\|_p = 1$ for all $n \geqslant 1$, and in particular f and g are differentially subordinated to each other. Now by Khintchine's inequality (3.34),

$$\|f_N\|_{L^p(\Omega;X)} = \Big\| \sum_{n=1}^N \varepsilon_n \Big\|_{L^p(\Omega)} \geqslant c_p\sqrt{N}.$$

On the other hand, pointwise in Ω, $\|g_N\|_X = N^{1/p}$. Clearly, the estimates $\|f_N\|_{L^p(\Omega;X)} \leqslant C\|g_N\|_{L^p(\Omega;X)}$ and $\|g_N\|_{L^p(\Omega;X)} \leqslant C\|f_N\|_{L^p(\Omega;X)}$ for all $N \geqslant 1$, can only hold if $p = 2$.

Note that the previous example is given by an operator-valued martingale transform. Actually, in Proposition 3.7.2 we have seen that validity of the L^p-boundedness of all martingale transforms for Banach space X and Y can be used to characterize martingale cotype 2 of X and martingale type 2 of Y.

Our calculation of the UMD constant $\beta_{p,\mathbb{R}}$ in Theorem 4.5.7 depended on the clever use of a peculiar function suited for that purpose. Inspired by the particular form of this function it is possible to make an educated guess for another function, defined on $X \times Y$, which can be used to prove Theorem 4.5.14.

Recall from Section 4.5.c that $F(x,y) \leqslant u(x,y)$, where the function $F = F_p : \mathbb{R}^2 \to \mathbb{R}$ and the Burkholder function $u = u_p : \mathbb{R}^2 \to \mathbb{R}$ are as in (4.35) and (4.38), i.e.,

$$F(x,y) := |y|^p - \beta_p^p |x|^p$$

and

$$u(x,y) = \alpha_p\big(|y| - \beta_p|x|\big)\big(|y| + |x|\big)^{p-1},$$

where $\alpha_p := p(1-1/p^*)^{p-1}$ and $\beta_p := p^* - 1$. For completeness we give a self-contained proof of this domination property below.

Lemma 4.5.18. *For all $p \in (1,\infty)$ and $x, y \in \mathbb{R}$ we have the estimate*

$$|y|^p - \beta_p^p|x|^p \leqslant \alpha_p(|y| - \beta_p|x|)(|x| + |y|)^{p-1}.$$

Proof. The case $p = 2$ reduces to the identity $|y|^2 - |x|^2 = (|y| - |x|)(|y| + |x|)$, so we henceforth consider $p \neq 2$.

By homogeneity it suffices to consider the case $|x| + |y| = 1$. Writing $|x| = s$ and $|y| = 1 - s$, we need to show the positivity for $s \in [0,1]$ of the function

$$\phi(s) := \alpha_p((1-s) - \beta_p s) - (1-s)^p + \beta_p^p s^p.$$

Noting that $1 + \beta_p = p^*$, we have

$$\begin{aligned}\phi'(s) &= -\alpha_p p^* + p(1-s)^{p-1} + p\beta_p^p s^{p-1},\\ \phi''(s) &= p(p-1)[-(1-s)^{p-2} + \beta_p^p s^{p-2}],\end{aligned}$$

and one checks that

$$\phi\Big(\frac{1}{p^*}\Big) = \phi'\Big(\frac{1}{p^*}\Big) = 0, \qquad \phi''\Big(\frac{1}{p^*}\Big) = p(p-1)\Big(\frac{\beta_p}{p^*}\Big)^{p-2}(\beta_p^2 - 1) > 0.$$

It is also immediate that ϕ'' is monotone on $(0,1)$ with a unique zero, which we denote by s_p. From this point on, it is convenient to make a case study.

Case $p < 2$:

Now ϕ'' is decreasing on $(0,1)$, hence positive on $(0, s_p)$, and thus $1/p^* < s_p$. Then ϕ' is increasing on $(0, s_p)$ with zero at $1/p^*$, hence $\phi' < 0$ on $(0, 1/p^*)$ and $\phi' > 0$ on $(1/p^*, s_p)$. Accordingly, ϕ is decreasing on $(0, 1/p^*)$ and increasing on $(1/p^*, s_p)$ with zero at $1/p^*$, hence non-negative on $[0, s_p]$. Since $\phi'' < 0$ on $(s_p, 1)$ and hence ϕ is concave on this interval, it remains to check that $\phi(1) \geqslant 0$, which is seen from

$$\phi(1) = \alpha_p(1 - p^*) + \beta_p^p = \beta_p^p[-p^{2-p}(p-1)^{p-1} + 1]$$

and observing that

$$p^{2-p}(p-1)^{p-1} \leqslant (2-p)p + (p-1)(p-1) = 1$$

by the arithmetic–geometric mean inequality with weights $(2-p) + (p-1) = 1$.

Case $p > 2$:

Now ϕ'' is increasing on $(0, 1)$, hence positive on $(s_p, 1)$, and thus $s_p < 1/p^*$. Then ϕ' is decreasing on $(0, s_p)$ and increasing on $(s_p, 1)$, hence $\phi' < 0$ on $(s_p, 1/p^*)$ and $\phi' > 0$ on $(1/p^*, 1)$. Accordingly, ϕ is decreasing on $(s_p, 1/p^*)$ and increasing on $(1/p^*, 1)$ with zero at $1/p^*$, hence non-negative on $[s_p, 1]$. Since $\phi'' < 0$ on $(0, s_p)$ and hence ϕ is concave on this interval, it remains to check that $\phi(0) \geqslant 0$. Since $\phi(0) = \alpha_p - 1 = p^{2-p}(p-1)^{p-1} - 1$, this follows again from the arithmetic–geometric mean inequality

$$p^{(p-2)/(p-1)} \leqslant \frac{p-2}{p-1}p + \frac{1}{p-1}1 = p - 1$$

with weights $(p-2)/(p-1) + 1/(p-1) = 1$. □

The aim will be to show that $\int_S V_p(f, g)\,\mathrm{d}\mu \leqslant 0$ for all f and g as in the theorem, where $V_p : X \times Y \to \mathbb{R}$ is given by

$$V_p(x, y) = \|y\|_Y^p - \beta_p^p \|x\|_X^p.$$

From the above we see that the Burkholder function $U_p : X \times Y \to \mathbb{R}$ given by

$$U_p(x, y) = \alpha_p(\|y\|_Y - \beta_p\|x\|_X)(\|x\|_X + \|y\|_Y)^{p-1}. \tag{4.39}$$

satisfies $V_p(x, y) \leqslant U_p(x, y)$ for all $x \in X$ and $y \in Y$. Thus it is enough to show

$$\int_S U_p(f, g)\,\mathrm{d}\mu \leqslant 0. \tag{4.40}$$

This task is facilitated by the following:

Lemma 4.5.19 (Osękowski's integral representation). *For every $p \in (1, \infty) \setminus \{2\}$, the function U_p defined in (4.39) has an integral representation*

$$U_p(x, y) = c_p \int_0^\infty t^{p-1} u\Big(\frac{x}{t}, \frac{y}{t}\Big)\,\mathrm{d}t, \quad x \in X,\ y \in Y. \tag{4.41}$$

where $u : X \times Y \to \mathbb{R}$ is a piecewise polynomial function of $\|x\|_X$ and $\|y\|_Y$, and has the following properties with functions

$$A : X \times Y \to X, \qquad B : X \times Y \to Y,$$

that are piecewise polynomial in x, y, $x/\|x\|_X$ and $y/\|y\|_Y$: For all $x, v \in X$ and $y, w \in Y$ with $\|w\|_Y \leqslant \|v\|_X$:

$$u(x + v, y + w) \leqslant u(x, y) + (A(x, y)|v) + (B(x, y)|w), \tag{4.42}$$

$$u(x, y) \leqslant 0 \quad \text{if} \quad \|y\|_Y \leqslant \|x\|_X, \tag{4.43}$$

$$\int_0^\infty t^{p-1}\Big|u\Big(\frac{x}{t}, \frac{y}{t}\Big)\Big|\,\mathrm{d}t \leqslant C_p(\|x\|_X^p + \|y\|_Y^p). \tag{4.44}$$

We have purposely left the subscript p from the functions u, A and B. The proof will show that a fixed choice of these functions works for all $p \in (1,2)$, and another fixed choice for all $p \in (2,\infty)$. The property (4.42) is an important concavity property which is stronger than the zigzag-concavity from Definition 4.5.3. It implies, in particular, the following:

Lemma 4.5.20. *The estimate* (4.42) *implies that the function*

$$t \in \mathbb{R} \mapsto u(x+tv, y+tw)$$

is concave for all $x, v \in X$ and $y, w \in Y$ with $\|w\|_Y \leqslant \|v\|_X$.

Proof. Let $t = \lambda_0 t_0 + \lambda_1 t_1$ with $\lambda_i \geqslant 0$ such that $\lambda_0 + \lambda_1 = 1$. Then

$$\begin{aligned} u(x+t_i v, y+t_i w) &= u(x+tv+(t_i-t)v, y+tw+(t_i-t)w) \\ &\leqslant u(x+tv, y+tw) + (A|(t_i-t)v) + (B|(t_i-t)w) \end{aligned}$$

by the assumption with $A = A(x+tv, y+tw)$ and likewise for B. Multiplying by λ_i and summing over $i = 0, 1$, we have

$$\sum_{i=0}^{1} \lambda_i u(x+t_i v, y+t_i w) \leqslant u(x+tv, y+tw),$$

since the terms with A and B cancel out using $\sum_{i=0}^{1} \lambda_i (t_i - t) = 0$. □

Proof of Theorem 4.5.14, assuming Lemma 4.5.19. Forgetting the complex structure and replacing the inner product by $\Re(x|y)$, it is enough to consider the case of real Hilbert spaces. Moreover, replacing f and g by $(f_k)_{k=m}^n$ and $(g_k)_{k=m}^n$ for some fixed $m \leqslant n$ it suffices to consider the case where f and g are finitely non-zero.

In order to prove the estimate (4.40) one may use the representation (4.41) and Fubini's theorem (which is allowed by (4.44)) to reduce our considerations to the estimate $\int_S u(f_n, g_n)\,\mathrm{d}\mu \leqslant 0$. Now the latter estimate is surprisingly simple to prove, given that u has the properties listed in Lemma 4.5.19. For each $k \in \mathbb{Z}$ writing $\mathbb{E}_k$ for the conditional expectation with respect to $\mathscr{F}_k$, by applying (4.42) with $v = df_{k+1}$ and $w = dg_{k+1}$ (here we use the differential subordination to have $\|w\|_Y \leqslant \|v\|_X$ in (4.42)) we find that for any $k \in \mathbb{Z}$

$$\begin{aligned} \mathbb{E}_k u(f_{k+1}, g_{k+1}) - u(f_k, g_k) &\leqslant \mathbb{E}_k(A(f_k,g_k)|df_{k+1}) + \mathbb{E}_k(B(f_k,g_k)|dg_{k+1}) \\ &= (A(f_k,g_k)|\mathbb{E}_k df_{k+1}) + (B(f_k,g_k)|\mathbb{E}_k dg_{k+1}) \\ &= 0 + 0, \end{aligned}$$

where Proposition 2.6.31 was used to pull out the terms $A(f_k, g_k)$ and $B(f_k, g_k)$ from the conditional expectation. Since f and g have only finitely many differences which are non-zero, we can iterate this argument to find that

$$\int_S u(f_n, g_n)\,\mathrm{d}\mu = \int_S \mathbb{E}_{n-1} u(f_n, g_n)\,\mathrm{d}\mu$$

$$\leqslant \int_S u(f_{n-1}, g_{n-1})\,\mathrm{d}\mu \leqslant \ldots \leqslant \int_S u(f_{-\infty}, g_{-\infty})\,\mathrm{d}\mu.$$

By (4.43) and $\|g_{-\infty}\|_Y \leqslant \|f_{-\infty}\|_X$ we find $u(f_{-\infty}, g_{-\infty}) \leqslant 0$ and this completes the proof of (4.40). □

It remains to prove Lemma 4.5.19. This is done separately for the cases $p \in (1,2)$ and $p \in (2,\infty)$. We abbreviate $\|x\| = \|x\|_X$ and $\|y\| = \|y\|_Y$ in the sequel.

Lemma 4.5.21 (Case $p \in (1,2)$). *For $p \in (1,2)$, the functions*

$$u(x,y) = \begin{cases} \|y\|^2 - \|x\|^2, & \text{if } \|x\| + \|y\| < 1; \\ 1 - 2\|x\|, & \text{if } \|x\| + \|y\| \geqslant 1, \end{cases}$$

and

$$A(x,y) = \begin{cases} -2x, & \text{if } \|x\| + \|y\| < 1; \\ -\frac{2x}{\|x\|}, & \text{if } \|x\| + \|y\| \geqslant 1, \end{cases} \qquad \text{where } \frac{0}{\|0\|} := 0,$$

$$B(x,y) = \begin{cases} 2y, & \text{if } \|x\| + \|y\| < 1; \\ 0, & \text{if } \|x\| + \|y\| \geqslant 1, \end{cases}$$

satisfy the assertions of Lemma 4.5.19 with $c_p = \frac{1}{2}p^{3-p}(2-p)$.

Proof. It is straightforward to verify the integral representation (4.41) and the two estimates (4.43) and (4.44), so we concentrate on the concavity property (4.42). We first note that

$$u(x,y) \leqslant 1 - 2\|x\|, \qquad \forall x \in X,\ y \in Y. \tag{4.45}$$

Indeed, this is clear for $\|x\| + \|y\| \geqslant 1$, and for $\|x\| + \|y\| < 1$ this follows from $u(x,y) = \|y\|^2 - \|x\|^2 \leqslant (1 - \|x\|)^2 - \|x\|^2 = 1 - 2\|x\|$.

Now let $x, v \in X$ and $y, w \in Y$, and assume that $\|w\| \leqslant \|v\|$. The proof of (4.42) is split into several cases according to the piecewise definition of u, A and B.

Case $\|x\| + \|y\| < 1$ and $\|x + v\| + \|y + w\| < 1$:

In this case (recall that we reduced to the case of real Hilbert spaces), we find

$$\begin{aligned} u(x+v, y+w) &= \|y\|^2 - \|x\|^2 + 2(y|w) - 2(x|v) + \|w\|^2 - \|v\|^2 \\ &\leqslant u(x,y) + (A(x,y)|v) + (B(x,y)|w). \end{aligned}$$

Case $\|x\|+\|y\|<1$ *and* $\|x+v\|+\|y+w\|\geqslant 1$*:*

We need to show that

$$1-2\|x+v\| \leqslant \|y\|^2-\|x\|^2+2(y|w)-2(x|v). \tag{4.46}$$

We write the right side as

$$\begin{aligned}
&\|y\|^2-\|x\|^2+2(y|w)-2(x|v)\\
&=\|y\|^2-\|x\|^2+(\|y+w\|^2-\|y\|^2-\|w\|^2)-(\|x+v\|^2-\|x\|^2-\|v\|^2)\\
&=\|y+w\|^2-\|x+v\|^2+\|v\|^2-\|w\|^2\\
&\geqslant \|y+w\|^2-\|x+v\|^2
\end{aligned}$$

by $\|w\|\leqslant\|v\|$. What remains is the product of $\|y+w\|+\|x+v\|\geqslant 1$ and

$$\|y+w\|-\|x+v\|=(\|y+w\|+\|x+w\|)-2\|x+v\|\geqslant 1-2\|x+v\|,$$

which proves (4.46).

Case $\|x\|+\|y\|\geqslant 1$*:*

Now by (4.45) and the Cauchy–Schwarz inequality we find that

$$\begin{aligned}
u(x+v,y+w)&\leqslant 1-2\|x+v\|\\
&\leqslant 1-2\Big(x+v\Big|\frac{x}{\|x\|}\Big)\\
&=1-2\|x\|-2\Big(\frac{x}{\|x\|}\Big|v\Big)\\
&=u(x,y)+(A(x,y)|v)+(B(x,y)|w).
\end{aligned}$$

□

Lemma 4.5.22 (Case $p\in(2,\infty)$). *For* $p\in(2,\infty)$*, the functions*

$$u_p(x,y)=\begin{cases}0, & \text{if } \|x\|+\|y\|\leqslant 1;\\ (\|y\|-1)^2-\|x\|^2, & \text{if } \|x\|+\|y\|>1,\end{cases}$$

and

$$A(x,y)=\begin{cases}0, & \text{if } \|x\|+\|y\|\leqslant 1;\\ -2x, & \text{if } \|x\|+\|y\|>1.\end{cases}$$

$$B(x,y)=\begin{cases}0, & \text{if } \|x\|+\|y\|\leqslant 1;\\ 2y-\frac{2y}{\|y\|}, & \text{if } \|x\|+\|y\|>1.\end{cases}$$

satisfy the assertions of Lemma 4.5.19 with $c_p=\frac12 p^{3-p}(p-1)^p(p-2)$.

Proof. The integral representation (4.41) and the estimate (4.44) are straightforward. To check (4.43) for $\|y\| \leqslant \|x\|$ note that this is clear for $\|x\|+\|y\| \leqslant 1$ and for $\|x\| + \|y\| > 1$ we estimate

$$\begin{aligned} u_p(x, y) &= 1 - 2\|y\| - (\|x\| + \|y\|)(\|x\| - \|y\|) \\ &\leqslant 1 - 2\|y\| - (\|x\| - \|y\|) = 1 - \|y\| - \|x\| < 0. \end{aligned}$$

For the proof of (4.42), we first check that

$$u_p(x, y) \leqslant (\|y\| - 1)^2 - \|x\|^2 \quad x \in X,\ y \in Y. \tag{4.47}$$

Indeed, for $\|x\| + \|y\| > 1$ this is trivial. For $\|x\| + \|y\| \leqslant 1$, it suffices to note that (4.47) if and only if $(1-\|y\|)^2 - \|x\|^2 \geqslant 0$ and this holds as $1-\|y\| \geqslant \|x\|$.

Now let $x, v \in X$ and $y, w \in Y$ be such that $\|w\| \leqslant \|v\|$. We consider the following cases:

Case $\|x\| + \|y\| > 1$*:*

By (4.47), we have

$$\begin{aligned} &u(x + v, y + w) \leqslant (\|y + w\| - 1)^2 - \|x + v\|^2 \\ &= (\|y\| - 1)^2 - \|x\|^2 - 2(x|v) + 2(y|w) + 2\|y\| - 2\|y + w\| + \|w\|^2 - \|v\|^2 \\ &= u(x, y) + (A(x, y)|v) + (B(x, y)|w) + 2T(y, w) + \|w\|^2 - \|v\|^2 \\ &\leqslant u(x, y) + (A(x, y)|v) + (B(x, y)|w), \end{aligned}$$

where we used $\|w\| \leqslant \|v\|$ and Cauchy–Schwarz inequality to find that

$$T(y, w) := \frac{(y|w)}{\|y\|} + \|y\| - \|y + w\| = \frac{(y|y + w)}{\|y\|} - \|y + w\| \leqslant 0.$$

Case $\|x\| + \|y\| \leqslant 1$*:*

If also $\|x+v\|+\|y+w\| \leqslant 1$, the estimate (4.42) is trivial as both sides vanish, so let $\|x + v\| + \|y + w\| > 1$. Then (4.42) reduces to showing

$$(\|y + w\| - 1)^2 - \|x + v\|^2 \leqslant 0.$$

If $\|y + w\| \leqslant 1$ this can be seen with a similar argument as below (4.47). If $\|y + w\| > 1$, then

$$\|w\| \geqslant \|y + w\| - \|y\| > 1 - \|y\| \geqslant \|x\|.$$

Therefore, combining this with $\|w\| \leqslant \|v\|$ and the Cauchy–Schwarz inequality, we find

$$\|y + w\| - 1 \leqslant \|y\| + \|w\| - 1 \leqslant \|w\| - \|x\| \leqslant \|v\| - \|x\| \leqslant \|x + v\|,$$

which completes the proof. □

Clearly the combination of Lemmas 4.5.21 and 4.5.22 completes the proof of Osękowski's Lemma 4.5.19, and thereby the proof of Theorem 4.5.14.

4.6 Notes

For many years, standard references for the theory of UMD spaces have been the survey papers of Rubio de Francia [1986] and Burkholder [2001]. An extensive recent account is given in the monograph of Pisier [2016]. A lot of material on UMD spaces, with the point of view of operator ideals, is also found in Pietsch and Wenzel [1998].

The class of UMD Banach spaces was introduced, essentially as a side remark, by Maurey [1975, Remarque 4] and Pisier [1975b, Remarque 2] under the name (I) (for "inconditionnalité de différences de martingales"): the exposition by Maurey [1975] concerned the classical unconditionality estimates for real-valued martingales, and Pisier had pointed out that the same argument proves the p-independence of the UMD property for an arbitrary Banach space X. The goal in Pisier [1975b] was the construction of a super-reflexive Banach space for which the Rademacher type is strictly larger than the martingale type (or, equivalently, the optimal modulus of smoothness of an equivalent norm), and he remarked in the end that such a space cannot be UMD.

It is mostly through the fundamental contributions of Burkholder that the central importance of the class of UMD spaces was recognised. In the fundamental paper Burkholder [1981a], a geometric characterisation of UMD spaces was obtained, and the connection with harmonic analysis was established a little later in Burkholder [1983], where it was shown that the periodic Hilbert transform acts boundedly on $L^p(\mathbb{T};X)$ for $1 < p < \infty$ if X is a UMD space. The converse result, that the UMD property is also necessary for the boundedness of the vector-valued Hilbert transform, was established in Bourgain [1983]. (Both these results will be treated in Chapter 5.) This deep connection between probability theory and harmonic analysis in Banach spaces turned out to have a wealth of implications. Influenced by these works, the theory of UMD spaces underwent a rapid development during the 1990's. Important connections with areas as diverse as parabolic partial differential equations, operator theory, stochastic integration, and Banach space theory were established, many of which will be explored in these volumes.

Section 4.1

A general introduction to unconditional convergence, Schauder bases, and Schauder decompositions in Banach spaces can be found in Lindenstrauss and Tzafriri [1977]. Theorem 4.1.1 is a classical result due to Burkholder [1966]. Proofs of the classical facts that $L^1(0,1)$ and $C[0,1]$ do not have an unconditional Schauder basis can be found in Albiac and Kalton [2006, Proposition 3.5.4 and Theorem 6.3.3]. The standard bases $(e_n \otimes e_m)_{n,m\in\mathbb{N}}$ for $\mathscr{C}^p(\ell^2)$, $1 \leqslant p < \infty$, and $\mathscr{K}(\ell^2)$ also fail to be unconditional (see Kwapień and Pełczyński [1970, Theorem 2.2]). On the other hand, for finite subsets $F \subseteq \mathbb{N}$ consider the orthogonal projection $p_F : \ell^2 \to \ell^2$ onto $\operatorname{span}\{e_n : n \in F\}$. The row and column projections R_F and C_F on $\mathscr{C}^p(\ell^2)$ given by

$$R_F(u) = p_F u, \quad C_F(u) = up_F$$

satisfy $\|R_F\| \leqslant 1$ and $\|C_F\| \leqslant 1$ and therefore the Schauder decompositions $(R_{\{n\}})_{n\geqslant 1}$ and $(C_{\{n\}})_{n\geqslant 1}$ are unconditional in $\mathscr{C}^p(\ell^2)$ for any $p \in [1,\infty)$ and in $\mathscr{K}(\ell^2)$. Further positive and negative results on (random) unconditionality can be found in Billard, Kwapień, Pelczyński, and Samuel [1986] and Pisier [1978]. Connections between the unconditionality of Schauder decompositions and the UMD property are studied in De Pagter and Witvliet [1999], and connections with harmonic analysis and R-boundedness are discussed in Clément, De Pagter, Sukochev, and Witvliet [2000] and Witvliet [2000].

Section 4.2

The results of this section are by now classical. A special feature of our presentation is the simultaneous treatment of real and complex signs in the defining inequality for the UMD property. Furthermore, continuing the philosophy of the previous chapter, we have stated our results in the σ-finite context whenever this is possible in order to facilitate applications in harmonic analysis.

The result that the dyadic analogue of the UMD property already implies UMD (Theorem 4.2.5) was sketched in Maurey [1975]. The approximation argument which underlies this result has already been given in Theorem 3.6.1. The alternative proof given in Theorem 4.5.6 is due to Burkholder [1986]. The p-independence of the UMD property (Theorem 4.2.7) was proved in Maurey [1975] and, with a different proof, in Burkholder [1981a]. Our approach based on the general extrapolation result for martingale transforms of Theorem 3.5.4 leads to the explicit estimate

$$\beta_{p,X} \leqslant 100\Big(\frac{p}{q} + \frac{p'}{q'}\Big)\beta_{q,X}, \quad 1 < p, q < \infty,$$

which seems to be new, although the number 100 is not sharp. Proposition 4.2.10 compares the complex and real UMD constants. For Hilbert spaces these constants coincide (as follows from Proposition 4.2.22), but it seems to be unknown whether or not this holds in general.

Theorem 4.2.13 will play a central role in the next chapter; it gives the equivalence of UMD with the unconditionality of the Haar system, and even the coincidence of the relevant constants. The scalar case of this equivalence can be found in Burkholder [1984, Theorems 14.1 and 15.1], where the unconditionality constant is shown to be $p^* - 1$ (see Corollary 4.5.16). The Haar basis is an example of a *monotone basis*, i.e., a Schauder basis $(x_n)_{n\geqslant 1}$ with the property that for all scalars $(a_j)_{j\geqslant 1}$ the following estimate holds:

$$\Big\| \sum_{j=1}^m a_j x_j \Big\|_X \leqslant \Big\| \sum_{j=1}^n a_j x_j \Big\|_X \qquad \forall m \leqslant n.$$

The following result was proved in Burkholder [1984, Theorem 15.2].

Theorem 4.6.1 (Burkholder). *Every monotone basis in $L^p(0,1)$ has unconditionality constant $p^* - 1$.*

Aldous [1979] used the UMD property (in his notation, property (I_p)) as a tool to study general unconditional bases in $L^p(0,1;X)$. Indeed, he proved that if $L^p(0,1;X)$ has an unconditional basis, then every martingale difference sequence in $L^p(0,1;X)$ is also unconditional, and therefore X is a UMD space. In the converse direction, Bukhvalov [1987] showed that if X has an unconditional basis and X is UMD, then $L^p(0,1;X)$ has an unconditional basis. In the same paper, some extensions of these results to $E(X)$ are obtained for suitable Banach function spaces E. Further results along this line can be found in Sukochev and Ferleger [1995]. Versions of such results for the Hardy space $H^1(0,1;X)$ are given in Müller and Schechtman [1991].

Further examples of UMD spaces

The examples and constructions in Subsection 4.2.c are standard. Further examples of UMD spaces can be obtained by applying interpolation and extrapolation methods. These include:

- Reflexive Lorentz spaces (by real interpolation of $L^{p_0}(S)$ and $L^{p_1}(S)$);
- Reflexive Lorentz–Zygmund spaces (see Cobos [1986]);
- Reflexive Besov, Triebel–Lizorkin, and Hardy–Sobolev spaces (see Cobos and Fernandez [1988]);
- Reflexive Orlicz spaces (see Fernandez and Garcia [1991] and Liu [1989]).

The following characterisation for vector-valued Banach function spaces is due to Rubio de Francia [1986].

Theorem 4.6.2 (Rubio de Francia). *Let E be a Banach function space and X be a Banach space. Then $E(X)$ is a UMD space if and only if both E and X are UMD spaces.*

It was shown in Bourgain [1986b] that the Schatten classes $\mathscr{C}^p$, $p \in (1,\infty)$, are UMD spaces (see Proposition 5.4.2 in the next chapter). This result was extended in Berkson, Gillespie, and Muhly [1986, 1987] and Pisier and Xu [2003] (in increasing order of generality) to the non-commutative spaces $L^p(\mathscr{M})$, $p \in (1,\infty)$, over an arbitrary von Neumann algebra $\mathscr{M}$, and several results on these spaces have been deduced with the help of their UMD property. In addition to this it is possible to prove a *non-commutative* version of the UMD inequality in $L^p(\mathscr{M})$ in terms of non-commutative martingales. More details can be found in Pisier and Xu [1997].

UMD Banach function spaces and the Hardy–Littlewood maximal function

The main ingredient in the proof of Theorem 4.6.2 presented in Rubio de Francia [1986] is a characterisation of UMD Banach function spaces in terms of the Hardy–Littlewood maximal function which we shall explain next.

Let E be a Banach function space over a measure space (S,Σ,μ). For $f \in L^1_{\rm loc}(\mathbb{R}^d;E)$ set

$$\widetilde{M}f(x)(s) := \sup_{B\ni x} \frac{1}{|B|}\int_B |f(y)(s)|\,\mathrm{d}y,$$

where the supremum is taken over all Euclidean balls with rational radius and centre with rational coordinates.

A Banach function space E is said to have the *Hardy–Littlewood property* if the sub-additive mapping $\widetilde{M}$ is bounded on $L^p(\mathbb{R}^d;E)$ for all $p\in(1,\infty)$. For an investigation of this property we refer the reader to García-Cuerva, Macías, and Torrea [1993]. In particular it is proved in this paper that the boundedness of $\widetilde{M}$ on $L^p(\mathbb{R}^d;E)$ for some $p\in(1,\infty)$ and dimension d implies the boundedness for all $p\in(1,\infty)$ and all dimensions d. Furthermore, they show that the boundedness of the corresponding variant of $\widetilde{M}$ on $L^p(\mathbb{T};E)$ is also equivalent.

Example 4.6.3. The space $E=\ell^q$ with $q\in(1,\infty]$ has the Hardy–Littlewood property (see Theorem 3.2.28).

The importance of the Hardy–Littlewood property comes from the following deep result which was proved in Bourgain [1984b] and Rubio de Francia [1986].

Theorem 4.6.4. *Let E be a Banach function space. The following assertions are equivalent:*

(1) *the space E has UMD;*
(2) *the spaces E and E^* both have the Hardy–Littlewood property.*

UMD is not a three-space property

A property (P) that a Banach space may or may not have is called a *three-space property* if, whenever X is a Banach space and $Y\subset X$ is a closed subspace, X has (P) if and only if Y and X/Y have (P). The following was proved in Kalton [1992]:

Theorem 4.6.5 (Kalton). *There exist a non-UMD Banach space X, with a closed subspace Y isomorphic to a Hilbert space, such that X/Y is isomorphic to a Hilbert space.*

The randomised UMD inequalities

In Proposition 4.2.3 it was observed that if $(f_n)_{n\geqslant1}$ is an L^p-martingale with values in a UMD space X and $(\varepsilon_n)_{n\geqslant1}$ is a Rademacher sequence on a probability space Ω, then

$$\frac{1}{\beta^-_{p,X}}\Big\|\sum_{n=1}^N df_n\Big\|_{L^p(S;X)} \leqslant \Big\|\sum_{n=1}^N \varepsilon_n df_n\Big\|_{L^p(S\times\Omega;X)} \leqslant \beta^+_{p,X}\,\Big\|\sum_{n=1}^N df_n\Big\|_{L^p(S;X)}.$$

The validity of each of these *randomised UMD inequalities* for all L^p-martingales with values in a given Banach space X (with constants depending only on p and X) can be isolated as two Banach space properties that X may or may not have. Motivated by earlier results in Garling [1986], these properties have been isolated and studied in Garling [1990] under the names *lower* and *upper estimates for random martingale transforms*, abbreviated as LERMT and UERMT. In the following discussion, we shall follow the terminology of Van Neerven and Veraar [2006] and Veraar [2007] and refer to these properties as UMD$^-$ and UMD$^+$. As a rule of thumb, any result for UMD spaces involving only one of the constants $\beta^{\pm}_{p,X}$ only requires the corresponding one-sided version of the UMD property. For instance, Stein's inequality in Theorem 4.2.23 holds for every UMD$^+$ space and Proposition 4.3.10 actually shows that every UMD$^+$ space is K-convex.

It was shown in Garling [1990] that both properties are independent of $p \in (1,\infty)$. For UMD$^+$ this also follows from the p-independence of the L^p-boundedness of a suitable operator-valued martingale transform (see Remark 4.2.24 for a similar situation).

It follows from Theorem 4.3.3 and Corollary 4.3.8 that every UMD$^+$ space is super-reflexive. This result is due to Garling [1990], where it was also shown that every L^1 space, and more generally every Banach space which is finitely representable in ℓ^1, is a UMD$^-$ space. Analogues for non-commutative L^1 have been obtained in Pisier and Xu [1997]; in particular their results imply that the Schatten class $\mathscr{C}^1$ is a UMD$^-$ space. Some characterisations of UMD$^-$ lattices were given in Zhang and Zhang [2012].

It is an open problem whether every UMD$^+$ space is UMD. As a possible hint towards a negative answer, Geiss [1999, Corollary 5 and the subsequent discussion] has also shown that there cannot exist any linear bound relating $\beta^+_{p,X}$ and $\beta_{p,X}$. Borrowing some terminology from the next subsection, Geiss [1999] also constructs an example of a UMD$^+$ *operator* which is not a UMD *operator*.

The part of Proposition 4.2.19 concerning $\beta^-_{p,\ell^\infty_{2^n}}$ is due to Garling [1990]. Further information on the constants β_{2,ℓ^1_n} and β_{2,ℓ^∞_n} and their randomised variants can be found in Pietsch and Wenzel [1998] and Wenzel [2005a].

It seems to be unknown whether UMD$^-$ and UMD$^+$ are implied by the corresponding notions when one restricts the defining inequalities to Paley–Walsh martingales. It is known, however, that the constant $\beta^-_{p,X}$ and its dyadic analogue, which one could denote by $\beta^{-,\Delta}_{p,X}$, are not the same in general. As explained in Cox and Veraar [2011], already in the case $X = \mathbb{R}$ this follows from Burkholder [1991, Theorem 3.3] and Hitczenko [1994, Theorem 1.1].

UMD operators

Let X and Y be Banach spaces and let $p \in (1,\infty)$. A bounded operator $T : X \to Y$ is said to be a *UMD operator* if there is a constant $C = C_{p,T}$ such that for all finite L^p-martingales $(f_n)_{n=0}^N$ with values in X and all scalars

$|\epsilon_n| = 1$, $n = 1, \dots, N$, we have

$$\Big\| \sum_{n=1}^{N} \epsilon_n T df_n \Big\|_{L^p(S;Y)} \leqslant C \Big\| \sum_{n=1}^{N} df_n \Big\|_{L^p(S;X)}.$$

As usual, it is understood that the quantification extends over all measure spaces supporting martingales. The least admissible constant C in this inequality will be denoted by $\beta_{p,T}$. Clearly, X is a UMD space if and only if the identity operator I_X is a UMD operator. In the same way as in Theorem 4.2.5 one shows that it suffices to consider Paley–Walsh martingales in this definition. As in Theorem 4.2.7 one shows that the property of being a UMD operator is independent of $p \in (1, \infty)$. This justifies the terminology "UMD operator" without reference to the index p. The class of UMD operators and the related class of Hilbert transform operators were introduced in the papers of Bourgain and Davis [1986] and Trautman [1986], where it was shown that they can be equivalently defined in terms of martingale transforms.

The class of all UMD operators is an operator ideal in the sense of $\mathfrak{Pietsch}$. For a systematic discussion of the UMD property from this point of view (and many other Banach space properties that can be treated in the same manner) we refer the reader to Pietsch and Wenzel [1998].

In Wenzel [2004] it is shown that the summation operator $\Sigma_N : \ell^1_N \to \ell^\infty_N$,

$$\Sigma_N\big((c_n)_{n=1}^N\big) := \Big(\sum_{k=1}^{n} c_k\Big)_{n=1}^{N},$$

satisfies

$$\sqrt{n} \lesssim \beta_{2,\Sigma_{2^n}} \lesssim n.$$

It is an open problem whether in fact $\beta_{2,\Sigma_{2^n}} \eqsim \sqrt{n}$. As is shown in Wenzel [2004], a positive answer to this question would have the interesting consequence that the second part of Problem O.7 has a negative solution. Some numerical evidence in favour of $\beta_{2,\Sigma_{2^n}} \eqsim \sqrt{n}$ is presented in Wenzel [2004].

Martingale inequalities derived from the UMD property

Theorem 4.2.23 is stated as a lemma in Bourgain [1984a, 1986b]; it is the vector-valued extension of an inequality of Stein [1970b, Theorem 8]. The proof presented here is from Figiel and Wojtaszczyk [2001, Lemma 34] and gives the estimate with constant $\beta^+_{p,X}$. An alternative proof can be found in Clément, De Pagter, Sukochev, and Witvliet [2000, Proposition 3.8].

The boundedness result for the martingale transform T_v in Theorem 4.2.25 with estimate $\|T_v\| \leqslant \beta_{p,X}$ was obtained in Burkholder [1981a, Theorem 2.2] and, with a simplified proof, in Burkholder [1984, Lemma 2.1]. The proof presented here uses the reduction argument of Proposition 3.6.16 which allows a reduction to the setting where the predictable sequence becomes constant. An advantage is that it also works for predictable complex-valued sequences.

Versions of the martingale transform inequalities for weighted L^p spaces of UMD-valued functions are given by Tozoni [1995, 1996].

The following weak type characterisation of the UMD property is due to Osękowski [2012a]: X is a UMD space if and only if for some (equivalently, all) $p \in (1,\infty)$ there exists a constant C such that the weak type estimate

$$\|g\|_{p,\infty} \leqslant C\|f\|_{p,\infty}$$

holds for all martingale transforms g of X-valued martingales f by a deterministic sequence with values in $\{-1,1\}$ (equivalently, by a predictable sequence with values in $[-1,1]$).

Unlike in the case of transforms T_v with scalar-valued v, in the case of operator-valued coefficients we do not obtain bounded operators (see Proposition 3.7.2). A characterisation of the boundedness in the operator-valued case was obtained in Girardi and Weis [2005]. Operator-valued martingale transforms in rearrangement invariant Banach function spaces were studied in Jiao, Wu, and Popa [2013].

Beyond unconditionality of the Haar system

The unconditionality of the Haar system (Theorem 4.2.13) means in particular, imposing some order on the countable family (D), that the series

$$\sum_{I\in\mathscr{D}} x_{\tau(I)}h_{\tau(I)} = \sum_{I\in\mathscr{D}} x_I h_I$$

converges in $L^p(\mathbb{R};X)$ independently of the permutation τ of $\mathscr{D}$. A related but more delicate matter is to study the series

$$\sum_{I\in\mathscr{D}} x_I h_{\tau(I)},$$

where a permutation is applied only to the Haar function, and not the coefficient x_I. Defining a linear operator on $L^2(\mathbb{R})$ uniquely by the property that $T_\tau : h_I \mapsto h_{\tau(I)}$, the latter series is, in fact, the action of $T_\tau \otimes I_X$ on the original series.

The investigation of such transformations was started by Figiel [1988] in the important particular case that $\tau(I) = I + m\ell(I)$ is a translation of the interval $I \in \mathscr{D}$ by a fixed multiple of its side-length. His estimates had profound applications in the analysis of vector-valued singular integrals, and will be discussed from this point of view in the Notes of Chapter 5. For more general permutations τ, the mapping properties (both boundedness and bounded invertibility) of $T_\tau \otimes I_X$ on $L^p(\mathbb{R};X)$ have been explored by Geiss and Müller [2009].

Walsh and Vilenkin systems

While Theorem 4.2.13 on the unconditionality of the Haar system is a fairly immediate reformulation of the UMD property, the following related characterisation may appear more surprising at first, in that it refers to a mere Schauder decomposition, without unconditionality:

Theorem 4.6.6. *Let X be a Banach space and $p \in (1,\infty)$, and let Ω be a probability space generated by a real Rademacher sequence $(r_k)_{k=1}^{\infty}$. Then X is a UMD space if and only if the Walsh projections*

$$P_S : f \mapsto w_S \langle w_S, f\rangle, \qquad w_S := \prod_{k\in S} r_k,$$

where S ranges over finite subsets of $\mathbb{N}$, form a Schauder decomposition of $L^p(\Omega;X)$ when enumerated according to

$$n(S) := \sum_{k\in S} 2^k.$$

Already in the case $X = \mathbb{R}$ the Walsh system does not lead to an unconditional Schauder decomposition unless $p = 2$ (see Example 4.1.12).

Sketch of proof. The density of finite linear combinations of $P_S f$ follows easily after observing that $(w_S)_{S\subseteq\{1,\dots,k\}}$ is an orthonormal basis of the 2^k-dimensional space $L^2(\sigma(r_1,\dots,r_k))$. The Schauder decomposition property is then equivalent to the uniform boundedness of the partial sum projections

$$\overline{P}_T := \sum_{S:n(S)\leqslant n(T)} P_S.$$

Let $E_k := \mathbb{E}(\ |\sigma(r_1,\dots,r_k))$, $D_k := E_k - E_{k-1}$, and

$$D_S := \sum_{k\in S} D_k.$$

The UMD property, expressed in terms of Paley–Walsh martingales, is equivalent to the uniform boundedness of the projections D_S, and in particular

$$\|D_S\|_{\mathscr{L}(L^p(\Omega;X))} \leqslant \beta_{p,X}.$$

The theorem then follows from the key identity

$$\overline{P}_T f = w_T D_T(w_T f), \tag{4.48}$$

since clearly the multiplication by w_T is an isometry of $L^p(\Omega;X)$. Equation (4.48), in turn, can be verified by testing on each $f = w_S$, and using the straightforward identities $w_T w_S = w_{T\Delta S}$ (where Δ is the symmetric difference of sets), $D_k w_R = \delta_{k,\max R} w_R$ and thus $D_T w_R = \mathbf{1}_T(\max R) w_R$, and noting that $\max(S\Delta T) \in T$ if and only if $\max(T\setminus S) \geqslant \max(S\setminus T)$ if and only if $n(T) \geqslant n(S)$. $\square$

Theorem 4.6.6 is contained in Tozoni [1995], where several further equivalent conditions, some of them involving weighted $L^p(\Omega;X)$ spaces, are given. It also admits an extension to the setting of *Vilenkin systems*. These are indexed by finitely non-zero sequences

$$\alpha = (\alpha_k)_{k=1}^{\infty} \in \prod_{k=1}^{\infty} \mathbb{Z}_{p_k}, \qquad p_k \in \mathbb{N}+2,$$

enumerated according to

$$n(\alpha) := \sum_{k=1}^{\infty} \alpha_k \prod_{j=1}^{k} p_j.$$

The *Vilenkin function* corresponding to the index α is then

$$w_\alpha := \prod_{k=1}^{\infty} \vartheta_k^{\alpha_k},$$

where $(\vartheta_k)_{k=1}^{\infty}$ is a sequence of independent random variables, and ϑ_k is uniformly distributed over the p_kth roots of unity on the complex unit circle. With the obvious identification of each sequence $\alpha \in \{0,1\}^{\mathbb{Z}_+}$ with the set of its non-zero coordinates, the Walsh system corresponds to the case $p_k \equiv 2$.

Theorem 4.6.6 remains true for all *bounded Vilenkin systems*, which means that the sequence $(p_k)_{k=1}^{\infty}$ is bounded; see Clément, De Pagter, Sukochev, and Witvliet [2000]. The proof is based on similar but somewhat more elaborate ideas as in the sketch for the Walsh system given above; in particular, a version of the key identity (4.48) is also true in this generality. The result was reproved in the special case of Banach function spaces by Weisz [2007b]. The extension to unbounded Vilenkin systems remains open even in this case.

The pointwise (almost everywhere) convergence of the Walsh and Vilenkin expansions of a function has also been studied. For scalar-valued functions in $L^p(\Omega)$, the almost everywhere convergence of the Walsh series is a classical result of Billard [1966/1967], which appeared shortly after the celebrated theorem of Carleson [1966] concerning the corresponding question for Fourier series. (We return to this in the Notes of Chapter 5.) Billard's theorem has been extended to $L^p(\Omega;X)$ under the following assumptions:

- if X is a UMD space with an unconditional basis, by Weisz [2007a], and
- if $X = [X_0, X_1]_\theta$ is a complex interpolation space between a Hilbert space X_0 and a UMD space X_1, by Hytönen and Lacey [2012].

By a theorem of Rubio de Francia [1986], the latter class of spaces contains the former. On the other hand, Weisz [2007a] deals with general bounded Vilenkin systems, while Hytönen and Lacey [2012] consider the Walsh system only. The extension of these results to general UMD spaces remains open.

Section 4.3

Helly's criterion (see Proposition 4.3.1) goes back to Helly [1921]. James's Theorem 4.3.2 is from James [1963/1964]. The notions of finite representability and super-reflexivity were introduced by James; the main references are James [1972a,b]. In the latter, it is shown that uniformly convex spaces are super-reflexive. The converse result that every super-reflexive space has an equivalent uniformly convex norm is due to Enflo [1972]. A comprehensive treatment of all aspects of (super-)reflexivity is provided in Van Dulst [1978]. Theorem 4.3.3 and Corollary 4.3.8 are due to Maurey [1975] and Aldous [1979]. Proposition 4.3.4 can be found in Pisier [1975a]. Most of the results of Subsection 4.3.b are standard, but some will only be proved in Volume II.

The example of Subsection 4.3.c is based on Qiu [2012]. It simplifies an older and more complicated example of a reflexive Banach function space without the UMD property of Bourgain [1984c]. Some additional geometric properties were built into Bourgain's example in Garling [1990].

More on finite representability

The simplest non-trivial example of finite representability is contained in the following proposition.

Proposition 4.6.7. *Every Banach space X is finitely representable in c_0.*

Proof. Let X_0 be a finite-dimensional subspace of X. Let $0 < \varepsilon < 1$ be arbitrary and fixed, and choose an ε-net $\{x_1^*, \dots, x_N^*\}$ in the unit ball of X_0^*, which is compact since X_0 is finite-dimensional. Consider the linear mapping $T : X_0 \to c_0$ defined by

$$Tx := (\langle x, x_1^*\rangle, \dots, \langle x, x_N^*\rangle, 0, 0, \dots).$$

We check that T is an isomorphism onto its range and that $\|T\|\|T^{-1}\| \leqslant 1+\varepsilon$.

Obviously we have $\|T\| \leqslant 1$. Let $x \in X_0$ be arbitrary and choose $x^* \in X_0^*$ of norm one such that $|\langle x, x^*\rangle| = \|x\|$. Choose $1 \leqslant n_0 \leqslant N$ such that $\|x_{n_0}^* - x^*\| \leqslant \varepsilon$. Then $|\langle x, x_{n_0}^*\rangle| \geqslant |\langle x, x^*\rangle| - \varepsilon\|x\| = (1-\varepsilon)|\langle x, x^*\rangle|$ and

$$(1-\varepsilon)\|x\| = (1-\varepsilon)|\langle x, x^*\rangle| \leqslant |\langle x, x_{n_0}^*\rangle| \leqslant \|Tx\|.$$

Thus T is an isomorphism onto its range whose inverse satisfies $\|T^{-1}\| \leqslant 1/(1-\varepsilon)$. $\square$

As a consequence of this result, if a Banach space Y does *not* have a certain super property (P), then c_0 also does not have (P) and more generally, any Banach space X in which c_0 can be finitely represented does not have (P). By the Maurey–Pisier theorem stated below, the class of such spaces X coincides with Banach spaces not having finite cotype.

Without proof we mention three celebrated results on finite representability; their proofs may be found in Albiac and Kalton [2006].

Theorem 4.6.8 (Principle of local reflexivity). *The bi-dual X^{**} of a Banach space X is always finitely representable in X.*

Theorem 4.6.9 (Dvoretzky–Rogers). *The space ℓ^2 is finitely representable in every infinite-dimensional Banach space X.*

Theorem 4.6.10 (Maurey–Pisier). *Let X be a Banach space.*

(1) *The space ℓ^1 is finitely representable in X if and only if X has no non-trivial type.*
(2) *The space c_0 is finitely representable in X if and only if X has no finite cotype.*

The Maurey–Pisier theorem has many important consequences for the behaviour of Banach space-valued random sums and will be discussed in detail in Volume II.

Section 4.4

The first paper to use decoupling techniques for Banach space-valued martingale difference sequences is Burkholder [1983]. In this paper the boundedness of the Hilbert transform on $L^p(\mathbb{R}^d; X)$ for UMD spaces X was proved for $1 < p < \infty$. In Garling [1986], a version of the decoupling inequality of Theorem 4.4.1 for stochastic integrals was obtained and used to give a new proof of Burkholder's result. Around the same time, the concept of tangent processes was introduced in Jacod [1984] and further developed in Kwapień and Woyczyński [1986, 1989, 1991]. Accounts of the history of the subject are given in Peña and Giné [1999] and Kwapień and Woyczyński [2002]. We also recommend the overview of decoupling properties and their mutual relations presented in Cox and Geiss [2016].

The first L^p-estimates for tangent martingale differences were obtained in Zinn [1985] and Hitczenko [1988b]. Their results can be viewed as forerunners of the decoupling result for tangent sequences contained in Theorems 4.4.9 and 4.4.11. Theorem 4.4.9 can be viewed as a common generalisation of Theorems 4.2.25 and 4.4.1, as it combines a martingale transform with tangent martingale differences. Using Theorem 3.5.4, the corresponding weak type $(1, 1)$ can be obtained by the same argument. Versions of Theorem 4.4.11 were proved in Hitczenko [1988a] and McConnell [1989] by using the Burkholder function constructed in Theorem 4.5.6. Their results will be described in more detail shortly. An alternative approach, based on a concrete representation of (tangent) martingale differences, was given in Montgomery-Smith [1998]. It uses the UMD property directly. Our proof of Theorem 4.4.11 uses the same type of argument, but avoids the representation theorem.

The main assumption in both Theorems 4.4.9 and 4.4.11 is the tangency of the pairs $(df_n, dg_n)_{n=1}^N$ and $(dg_n, df_n)_{n=1}^N$. The stronger assumption involving the conditional independence of $(df_n)_{n=1}^N$ and $(dg_n)_{n=1}^N$ discussed in Remark

4.4.10 goes back to McConnell [1989]. In this paper it was shown that if X is a UMD space X and $p \in (1,\infty)$, then for all pairs of X-valued tangent martingale difference sequences $(df_n)_{n=1}^N$ and $(dg_n)_{n=1}^N$ that are conditionally independent given $\mathscr{F}_{n-1}$ one has

$$\Big\| \sum_{n=1}^N dg_n \Big\|_{L^p(\Omega;X)} \leqslant \beta_{p,X} \Big\| \sum_{n=1}^N df_n \Big\|_{L^p(\Omega;X)}.$$

Another condition implying the tangency of certain sequences of pairs of random variables goes back to Hitczenko [1988a]. In order to explain it we need a result of Kwapień and Woyczyński [1991] to the effect that if $(d_n)_{n\geqslant 1}$ is an adapted sequence of random variables, there exists a so-called *decoupled tangent sequence* $(c_n)_{n\geqslant 1}$ (on a possibly enlarged filtered probability space). By this we mean that $(c_n)_{n\geqslant 1}$ is tangent to $(d_n)_{n\geqslant 1}$ and moreover satisfies the so-called *(CI) condition*: there exists a σ-algebra $\mathscr{G}$ containing $\mathscr{F}_\infty$ such that:

(i) for all $B \in \mathscr{B}(X)$ and all $n \geqslant 1$ we have

$$\mathbb{P}(\{c_n \in B\}|\mathscr{F}_{n-1}) = \mathbb{P}(\{c_n \in B\}|\mathscr{G});$$

(ii) for all $N \in \mathbb{N}$ and $\forall B_1, \dots, B_N \in \mathscr{B}(X)$ we have

$$\mathbb{P}\Big(\bigcap_{n=1}^N \{c_n \in B_n\} \Big| \mathscr{G} \Big) = \prod_{n=1}^N \mathbb{P}(\{c_n \in B_n\}|\mathscr{G}).$$

The sequence $(c_n)_{n\geqslant 1}$ is unique in the following sense: any other tangent sequence $(c_n')_{n\geqslant 1}$ satisfying the (CI) condition has the same distribution as $(c_n)_{n\geqslant 1}$. Now, given two tangent martingale difference sequences $(df_n)_{n\geqslant 1}$ and $(dg_n)_{n\geqslant 1}$, let $(c_n)_{n\geqslant 1}$ be the decoupled tangent sequence to $(df_n)_{n\geqslant 1}$ (and then also to $(dg_n)_{n\geqslant 1}$). We claim that for all $N \geqslant 1$

$$(df_n, c_n)_{n=1}^N \text{ and } (c_n, df_n)_{n=1}^N \text{ are tangent} \tag{4.49}$$

and similarly $(dg_n, c_n)_{n=1}^N$ and $(c_n, dg_n)_{n=1}^N$ are tangent. Admitting this claim for the moment, by applying Theorem 4.4.11 twice we obtain

$$\beta_{p,X}^{-1} \Big\| \sum_{n=1}^N dg_n \Big\|_{L^p(\Omega;X)} \leqslant \Big\| \sum_{n=1}^N c_n \Big\|_{L^p(\Omega;X)} \leqslant \beta_{p,X} \Big\| \sum_{n=1}^N df_n \Big\|_{L^p(\Omega;X)}.$$

This leads to the following result of Hitczenko [1988a]: If X is a UMD space and $p \in (1,\infty)$, then for all X-valued tangent martingale difference sequences $(df_n)_{n=1}^N$ and $(dg_n)_{n=1}^N$ one has

$$\Big\| \sum_{n=1}^N dg_n \Big\|_{L^p(\Omega;X)} \leqslant \beta_{p,X}^2 \Big\| \sum_{n=1}^N f_n \Big\|_{L^p(\Omega;X)}. \tag{4.50}$$

To prove the claim (4.49), note that if $B_1, B_2 \in \mathscr{B}(X)$, then by standard properties of the conditional expectations, the fact that $\mathscr{F}_{n-1}$ is contained in $\mathscr{G}$, and the tangency assumption, we have

$$\begin{aligned}
\mathbb{P}(\{df_n \in B_1, c_n \in B_2\}|\mathscr{F}_{n-1}) &= \mathbb{E}(\mathbb{P}(\{df_n \in B_1, c_n \in B_2\}|\mathscr{G})|\mathscr{F}_{n-1}) \\
&= \mathbb{E}(\mathbf{1}_{df_n \in B_1}\mathbb{P}(\{c_n \in B_2\}|\mathscr{G})|\mathscr{F}_{n-1}) \\
&= \mathbb{E}(\mathbf{1}_{df_n \in B_1}\mathbb{P}(\{c_n \in B_2\}|\mathscr{F}_{n-1})|\mathscr{F}_{n-1}) \\
&= \mathbb{E}(\mathbf{1}_{df_n \in B_1}\mathbb{P}(\{df_n \in B_2\}|\mathscr{F}_{n-1})|\mathscr{F}_{n-1}) \\
&= \mathbb{P}(\{df_n \in B_1, df_n \in B_2\}|\mathscr{F}_{n-1}),
\end{aligned}$$

which proves the tangency of $(df_n, c_n)_{n=0}^N$ and $(df_n, df_n)_{n=0}^N$. In a similar way one shows the tangency of (c_n, df_n) and (df_n, df_n). The same holds with df_n replaced by dg_n.

Let $(d_n)_{n\geqslant 1}$ be an adapted *scalar-valued* sequence of random variables (not necessarily a martingale difference sequence) and let $(c_n)_{n\geqslant 1}$ be its decoupled tangent sequence. It was shown in Hitczenko [1994] that there exists a constant K such that for all $p \in [1,\infty)$ we have

$$\Big\| \sum_{n=1}^N d_n \Big\|_p \leqslant K \Big\| \sum_{n=1}^N c_n \Big\|_p. \tag{4.51}$$

The point here is that K is independent of p, so this result is genuinely stronger than the ones discussed above. In the vector-valued situation, the validity of the right-hand side estimate in (4.51) with a p-dependent constant K_p has been studied as a Banach space property, the so-called *decoupling property*, in Cox and Veraar [2007]. In Cox and Veraar [2011] it is proved that if a Banach space has the decoupling property for some exponent $p \in [1,\infty)$, then it has this property for all exponents $p \in [1,\infty)$; this fact justifies the terminology "decoupling property" without reference to the exponent p. In Cox and Veraar [2007] it is shown that all UMD spaces and L^1 spaces enjoy the decoupling property. It is furthermore shown that a Banach space has the Paley–Walsh version of this property if and only if it has the Paley–Walsh version of the UMD$^-$ property. In Cox and Veraar [2011] it is shown that Hilbert spaces enjoy the decoupling property, with constants $K_p = K$ independent of p, thus generalising Hitczenko's inequality (4.51) to Hilbert spaces.

In a slightly different formulation, Corollary 4.4.15 appears in Hytönen [2014]. Yet another formulation was given in Hänninen and Hytönen [2016], with a proof that uses the UMD property in a direct way. In both papers the result is applied to the study of singular integrals.

Section 4.5

Theorem 4.5.6 was first obtained in Burkholder [1986] where it was stated in terms of the existence of a bi-concave function $U : X \times X \to \mathbb{R}$ satisfying

$$U(x,y) \geqslant \Big\| \frac{x+y}{2} \Big\|^p - \beta^p \Big\| \frac{x-y}{2} \Big\|^p.$$

Here, 'bi-concavity' means concavity in both variables separately. In our presentation we have 'rotated' Burkholder's function U by $45°$ (thus $U(x,y) = u(\frac{x-y}{2}, \frac{x+y}{2})$, where u is as in Theorem 4.5.6). The advantage of this rotation is that it allows a treatment of complex signs as well. It also turns out to be more appropriate for proving L^p-estimates for differentially subordinated martingales in Section 4.5.d.

Theorem 4.5.7 was first obtained by Burkholder [1984]. Several variations of the proof were published in subsequent papers. We have borrowed some arguments from the summer school lectures Burkholder [1991], where an explicit formula for an optimal (in a certain sense) function u is obtained in the scalar case. This formula can be used to construct an explicit function u for $X = \ell^p$ as well.

Theorem 4.5.14 was obtained in Burkholder [1984]. Our proof follows the presentation in Osękowski [2012b].

ζ-convexity

A Banach space X said to be *ζ-convex* if there exists a bi-convex function $\zeta : X \times X \to \mathbb{R}$ such that

$$\zeta(0,0) > 0 \tag{4.52}$$

and

$$\zeta(x,y) \leqslant \|x+y\| \quad \text{whenever} \quad \|x\| = \|y\| = 1. \tag{4.53}$$

Here, 'bi-convexity' means that both $x \mapsto \zeta(x,y)$ and $y \mapsto \zeta(x,y)$ are convex functions. The equivalence of the UMD property with ζ-convexity is due to Burkholder [1981b]; the formulation given here is from Burkholder [1986].

Theorem 4.6.11 (Burkholder). *A Banach space is ζ-convex if and only if it is a UMD space. More precisely, the following assertions hold.*

(1) *If X is ζ-convex, and if $\zeta : X \times X \to \mathbb{R}$ is bi-convex and satisfies (4.52) and (4.53), then X is UMD and for all $1 < p < \infty$ we have*

$$\beta_{p,X} \leqslant \frac{72}{\zeta(0,0)} \frac{(p+1)^2}{p-1}.$$

(2) *If X is UMD, then there exists a bi-convex function $\zeta : X \times X \to \mathbb{R}$ that satisfies (4.52) and (4.53) such that, for all $p \in (1,\infty)$,*

$$\beta_{p,X} \geqslant \frac{1}{\zeta(0,0)}.$$

That UMD implies ζ-convexity can be proved alternatively by using the function u from Section 4.5 to construct a bi-convex function ζ satisfying $\zeta(0,0) > 0$ and (4.53) as follows. Let

$$\zeta(x,y) := \frac{1 - u\big(\frac{x+y}{2}, \frac{x-y}{2}\big)}{p\beta^p}.$$

The zigzag-concavity of u implies that ζ is bi-convex, and since $u(0,0) = 0$ (this was derived below 4.5.6) we find that $\zeta(0,0) \geqslant \frac{1}{p\beta^p}$ (this bound on ζ is worse than the one in Theorem 4.6.11). To obtain the estimate (4.53) for $\|x\| = \|y\| = 1$, note that by (4.30)

$$p\beta^p\zeta(x,y) \leqslant 1 - \Big\|\frac{x-y}{2}\Big\|^p + \beta^p\Big\|\frac{x+y}{2}\Big\|^p \leqslant 1 - (1-t)^p + \beta^p t^p =: f(t),$$

where $t := \|\frac{x+y}{2}\| \in [0,1]$; we used $\|\frac{x-y}{2}\| \leqslant \|x\| - \|\frac{x+y}{2}\| = 1 - t$. Thus in order to obtain (4.53) it remains to verify that $f(t) \leqslant 2p\beta^p t$. This is equivalent to verifying that $g(t) \leqslant 0$, where $g : [0,1] \to \mathbb{R}$ is given by $g(t) = 1 - (1-t)^p - \beta^p pt - \beta^p(pt - t^p)$. This indeed holds since $pt \geqslant t^p$, $\beta \geqslant 1$, and $1 - (1-t)^p - pt \leqslant 0$.

Theorem 4.6.11 was complemented in Burkholder [1986] by the following characterisation of Hilbert spaces.

Theorem 4.6.12. *A Banach space X is isometric to a Hilbert space if and only if there exists a bi-convex function $\zeta : X \times X \to \mathbb{R}$ such that*

$$\zeta(0,0) = 1$$

and (4.53) holds.

For Hilbert spaces X, a bi-convex function satisfying $\zeta(x,y) \leqslant \|x+y\|$ for $\|x\| = \|y\| = 1$ is given by

$$\zeta(x,y) = 1 + \Re(x|y).$$

Another rare instance where an explicit ζ-function can be given is in the case $X = \ell^1_N$ Osękowski [2016].

The paper Burkholder [1981b] contains the further interesting result that a Banach space X is isometric to a Hilbert space if and only if

$$\lambda\mathbb{P}(g^\star > \lambda) \leqslant 2\|f\|_1 \quad \forall\lambda > 0$$

for all martingale transforms g of X-valued L^1-martingales f by predictable real-valued sequences v such that $\|v\|_\infty \leqslant 1$. To see this result in the right perspective, note that UMD spaces X are characterised by the validity of the inequality

$$\lambda\mathbb{P}(g^\star > \lambda) \lesssim_X \|f\|_1 \quad \forall\lambda > 0.$$

Indeed, this follows by combining Theorems 3.5.4 and 4.2.25 and the definition of the UMD property.

Further properties and problems related to Burkholder's function

A general reference for Burkholder functions and related topics is Osękowski [2012b].

By integrating the concavity estimate (4.42) of Osękowski's Lemma 4.5.19, we arrive at a similar conclusion

$$U_p(x+v, y+w) \leqslant U_p(x,y) + (A_p(x,y)|v) + (B_p(x,y)|w) \quad \forall \, \|w\| \leqslant \|v\| \tag{4.54}$$

for the original Burkholder function U_p, where the reader may evaluate explicit expressions for the constants A_p and B_p by a direct computation.

The function $U_p : \mathbb{C} \times \mathbb{C} \to \mathbb{R}$, can be alternatively interpreted as a function of four real variables, or, of 2×2 real matrices. Under such an interpretation, (4.54) implies a property known as *rank-one concavity*:

Definition 4.6.13. *A function $F : \mathbb{R}^{n \times n} \to \mathbb{R}$ is called* rank-one concave *if*

$$\mathbb{R} \ni t \mapsto F(A + tH)$$

is concave for every $A \in \mathbb{R}^{n \times n}$ and every rank-one matrix $H \in \mathbb{R}^{n \times n}$.

To be specific about the interpretation of the action of U_p on $\mathbb{R}^{2 \times 2}$, let us consider the $\mathbb{R}$-linear one-to-one transformation

$$\Gamma : \mathbb{R}^{2 \times 2} \to \mathbb{C} \times \mathbb{C}, A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (x, y) = \begin{pmatrix} \frac{1}{2}(a+d) + i\frac{1}{2}(c-b) \\ \frac{1}{2}(a-d) + i\frac{1}{2}(c+b) \end{pmatrix}. \tag{4.55}$$

For a rank-one matrix

$$H = \mathbf{h} \otimes \mathbf{k} = \begin{pmatrix} h_1 k_1 & h_1 k_2 \\ h_2 k_1 & h_2 k_2 \end{pmatrix},$$

we have

$$\Gamma(H) = (v, w), \qquad \begin{cases} v = \frac{1}{2}(h_1 k_1 + h_2 k_2) + i\frac{1}{2}(h_2 k_1 - h_1 k_2), \\ w = \frac{1}{2}(h_1 k_1 - h_2 k_2) + i\frac{1}{2}(h_2 k_1 + h_1 k_2). \end{cases}$$

After cancelling the cross terms $h_1 k_1 h_2 k_2$ of opposite signs, v and w are seen to satisfy

$$|v|^2 = |w|^2 = \frac{1}{4}(h_1^2 k_1^2 + h_2^2 k_2^2 + h_2^2 k_1^2 + h_1^2 k_2^2) = \frac{1}{4}|\mathbf{h}|^2 |\mathbf{k}|^2.$$

Thus, for such a matrix H and any A, we have

$$U_p \circ \Gamma(A + tH) = U_p(x + tv, y + tw), \qquad |v| = |w|,$$

where the last equality is a particular case of the condition $|w| \leqslant |v|$ under which the concavity property (4.54) is valid. This readily implies (see the proof of Lemma 4.5.20):

Proposition 4.6.14. *The function* $U_p \circ \Gamma : \mathbb{R}^{2\times 2} \to \mathbb{R}$ *is rank-one concave.*

The transformation Γ from (4.55) is particularly suited to considerations involving the Jacobian matrix

$$D\mathbf{u} = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}$$

of a function $\mathbf{u} = (u, v) : \mathbb{R}^2 \to \mathbb{R}^2$. We identify $\mathbf{u}$ with $f := u + iv : \mathbb{C} \to \mathbb{C}$, and recall the complex derivatives

$$\begin{aligned} \partial f &= \tfrac{1}{2}(f_x - i f_y) = \tfrac{1}{2}(u_x + v_y) + i\tfrac{1}{2}(v_x - u_y) \\ \bar\partial f &= \tfrac{1}{2}(f_x + i f_y) = \tfrac{1}{2}(u_x - v_y) + i\tfrac{1}{2}(v_x + u_y). \end{aligned}$$

A moment's comparison with (4.55) reveals that

$$\Gamma(D\mathbf{u}) = (\partial f, \bar\partial f), \qquad U_p \circ \Gamma(Du) = U_p(\partial f, \bar\partial f).$$

In this context, the right side is frequently expressed in terms of

$$|Df| := |\partial f| + |\bar\partial f|, \quad Jf := \det(D\mathbf{u}) = |\partial f|^2 - |\bar\partial f|^2 = (|\partial f| - |\bar\partial f|)|Df|$$

as

$$\begin{aligned} U_p(\partial f, \bar\partial f) &= \alpha_p(|\bar\partial f| - \beta_p|\partial f|)(|\bar\partial f| + |\partial f|)^{p-1} \\ &= \alpha_p\Big(\frac{1}{2}\big(|Df| - \frac{Jf}{|Df|}\big) - \frac{\beta_p}{2}\big(|Df| + \frac{Jf}{|Df|}\big)\Big)|Df|^{p-1} \\ &= -\alpha_p\Big(\frac{p^*}{2}Jf + \big(\frac{p^*}{2} - 1\big)|Df|^2\Big)|Df|^{p-2}. \end{aligned}$$

A long-standing conjecture in complex analysis, first stated by Iwaniec [1982], claims that

$$\|\bar\partial f\|_p \leqslant \beta_p \|\partial f\|_p \qquad \forall f \in C_c^\infty(\mathbb{C}). \tag{4.56}$$

By the pointwise bound $|\bar\partial f|^p - \beta_p^p|\partial f|^p \leqslant U_p(\partial f, \bar\partial f)$, this would follow from the stronger conjecture that

$$\int_{\mathbb{R}^2} U_p \circ \Gamma(D\mathbf{u})\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{C}} U_p(\partial f, \bar\partial f)\,\mathrm{d}x\,\mathrm{d}y \leqslant 0 \quad \forall \mathbf{u} \in C_c^\infty(\mathbb{R}^2; \mathbb{R}^2). \tag{4.57}$$

Using the integral representation of Lemma 4.5.19 and a Fubini argument it also suffices to check the same estimate for the p-independent function u introduced there. The inequality (4.57), in turn, is an instance of the following notion:

Definition 4.6.15. *A function* $F : \mathbb{R}^{n\times n} \to \mathbb{R}$ *is* quasi-concave *at* $A \in \mathbb{R}^{n\times n}$ *if*

$$\fint_{\Omega} F(A + D\mathbf{u}) \leqslant F(A) \tag{4.58}$$

for every bounded domain $\Omega \subseteq \mathbb{R}^n$*, and every* $\mathbf{u} \in C_c^\infty(\Omega; \mathbb{R}^n)$.

Since the work of Morrey [1952, 1966], it is known that quasi-concavity implies rank-one concavity, but it remained open for several decades whether the converse holds. In Šverák [1992] and Pedregal and Šverák [1998] it was shown that this converse is wrong for $n \geqslant 3$. The case $n = 2$ is still open and usually referred to as Morrey's problem. We refer to Astala, Iwaniec, and Martin [2009] and Astala, Iwaniec, Prause, and Saksman [2012] for details and further references. Thus recalling Proposition 4.6.14, an even stronger conjecture that implies (4.57) and hence (4.56) is the following:

Conjecture 4.6.16. Every continuous rank-one concave function $F : \mathbb{R}^{2\times 2} \to \mathbb{R}$ is quasi-concave at every $A \in \mathbb{R}^{2\times 2}$.

Indeed, the bound (4.57) corresponds to quasi-concavity at $A = 0$. In contrast to this, there is the following partial result at $A = I$, the identity matrix.

Theorem 4.6.17 (Astala, Iwaniec, Prause, and Saksman [2012]). *For every bounded domain $\Omega \subseteq \mathbb{R}^2$, the quasi-concavity estimate (4.58) at $A = I$ holds for $F = U_p \circ \Gamma$ and every $\mathbf{u} = (u, v) \in C_c^\infty(\Omega; \mathbb{R}^2)$ such that*

$$U_p \circ \Gamma(I + D\mathbf{u}) = U_p(1 + \partial f, \bar{\partial} f) \leqslant 0$$

everywhere, where $f = u + iv$.

Pisier's inequality

Among diverse applications of the UMD property is the following. Let $\Omega := \{-1, 1\}^d$ be equipped with its natural probability measure. For $f : \Omega \to X$, the ith discrete partial derivative is defined by

$$\partial_i f(\varepsilon) := \frac{1}{2}(f(\varepsilon) - f(S_i \varepsilon)),$$

where S_i swaps the ith component of $\varepsilon = (\varepsilon_j)_{j=1}^d \in \Omega$. For an arbitrary Banach space X, Pisier [1986b] obtained the following estimate, which may be thought of as a discrete Poincaré inequality, and Talagrand [1993] proved that the logarithmic growth of the constant is unavoidable in general: for $p \in [1, \infty)$ one has

$$\|f - \langle f \rangle\|_p \leqslant 2e(\log d) \cdot \Big(\mathbb{E}_r \Big\| \sum_{i=1}^d r_i \partial_i f \Big\|_p^p\Big)^{1/p}, \tag{4.59}$$

where $(r_i)_{i=1}^d$ is a real Rademacher sequence. In contrast to this, Naor and Schechtman [2002] showed that there is a d-independent estimate in (4.59), depending instead on the UMD constant $\beta_{p,X}$, provided that X is a UMD space and $p \in (1, \infty)$. (Observe a smaller range of p compared to Pisier's original inequality.) This is applied to questions of *non-linear type* of X. A further improvement, involving only the constant β^+_{p',X^*}, is due to Hytönen and Naor [2013].

Analytic UMD spaces

Let X be a *complex* Banach space. A sequence of X-valued functions $(f_n)_{n=0}^\infty$ is called an *analytic martingale*, if it is adapted to a filtration $\mathscr{F}_n = \sigma(\varepsilon_1, \ldots, \varepsilon_n)$ generated by a sequence of independent *complex* Rademacher (or Steinhaus) variables $(\varepsilon_n)_{n=1}^\infty$, and moreover the difference sequence $d_n = f_n - f_{n-1}$, $n \geqslant 1$, has the special form

$$d_n = \phi_n(\varepsilon_1, \ldots, \varepsilon_{n-1})\varepsilon_n.$$

Note that, replacing the complex variables ε_n by the real r_n, leads to the definition of Paley–Walsh martingales which, as we have seen, are sufficiently representative of all martingales for many purposes. In contrast to this, the class of analytic martingales is much more specific.

Requiring the inequality

$$\Big\| \sum_{n=1}^N \epsilon_n df_n \Big\|_{L^p(\Omega;X)} \leqslant C \|f_N\|_{L^p(\Omega;X)} \tag{4.60}$$

for all analytic martingales $(f_n)_{n\geqslant 0}$ and all scalars $|\epsilon_n| \leqslant 1$ leads to the notion of *analytic UMD* or *AUMD* spaces. This class was first introduced, under the name AMT (for analytic martingale transform) spaces, by Garling [1988], who showed that the estimate (4.60) holds for all $p \in (0, \infty)$ as soon as it holds for one. Other basic properties established by Garling [1988] include:

- All UMD spaces and all L^1 spaces are AUMD.
- If X is an AUMD space, then so is $L^p(\mu; X)$ for $p \in [1, \infty)$.
- AUMD spaces have finite cotype.

Piasecki [1997a,b] obtained characterisations of AUMD spaces analogous to those of UMD spaces by tangent martingale inequalities (Theorem 4.4.11 and Proposition 4.4.14), and by ζ-convexity (Theorem 4.6.11). Further results on AUMD spaces in the context of Fourier multipliers will be discussed in the Notes of Chapter 5 (see page 490).

5

Hilbert transform and Littlewood–Paley theory

In this final chapter, we investigate the extension of several operators and inequalities of classical Euclidean harmonic analysis to functions in $L^p(\mathbb{R}^d; X)$. As it turns out, a necessary and sufficient condition for this programme to work is that X is a UMD space, providing another manifestation of the importance of this class of spaces. From the point of view of Banach space theory, we obtain several characterisations of UMD spaces in terms of inequalities of Fourier analysis, including the boundedness of the Hilbert transform, the Mihlin multiplier theorem, and different forms of Littlewood–Paley inequalities. From the point of view of applications, all these estimates become powerful tools for the analysis of functions taking values in a UMD space. The Mihlin multiplier theorem, in particular, is extremely useful in many different contexts, a couple of which we illustrate in the text.

On the technical side, to carry out this programme, it is necessary to relate the Hilbert transform and other Fourier analytic objects with the martingale theory in an efficient way. The Euclidean space is certainly not short of natural filtrations, and a useful first model is already provided by the filtration generated by the dyadic cubes (or intervals on the line) discussed in a number of examples in the earlier chapters. However, on a closer look, this model on its own is somewhat inadequate to capture one of the key features of classical Fourier analysis, the invariance under translations: a non-integer translation of a dyadic interval $[k, k+1)$ is no longer an interval of the same form.

To compensate for this drawback, we add another structural level to the theory developed in the previous chapters: a fixed filtration is replaced by an ensemble of filtrations, indexed by a probability space over which we may compute averages and thereby recover the translation invariance "on average". In a sense, we apply probability theory on two different levels: the general theory of martingales on the one hand, and a random selection of the relevant martingales on the other hand.

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 63, DOI 10.1007/978-3-319-48520-1_5

5.1 The Hilbert transform as a singular integral

The Hilbert transform is formally defined as

$$\text{“}Hf(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{f(y)}{x-y}\,\mathrm{d}y\text{”},$$

but to make precise sense of the right-hand side, one needs some care. Hence we introduce the *truncated Hilbert transforms*

$$H_{\varepsilon,R}f(x) := \frac{1}{\pi}\int\limits_{\varepsilon<|x-y|<R} \frac{f(y)}{x-y}\,\mathrm{d}y$$

which make good sense for all $f \in L^1_{\mathrm{loc}}(\mathbb{R};X)$. The actual Hilbert transform is then defined as

$$Hf(x) := \lim_{\substack{\varepsilon\downarrow 0\\ R\to\infty}} H_{\varepsilon,R}f(x) \tag{5.1}$$

if the limit exists for almost every $x \in \mathbb{R}$. Here ε and R are allowed to approach their respective limits independently, and it is asked that the limit of $H_{\varepsilon,R}f(x)$ be independent of this approach. When this is the case, we say that “the Hilbert transform of f exists” and denote it by Hf. For functions $f \in C^1_{\mathrm{c}}(\mathbb{R};X)$ the Hilbert transform always exists, for any choice of Banach space X. Indeed, this follows from the simple identity

$$H_{\varepsilon,R}f(x) = \frac{1}{\pi}\int_{\varepsilon}^{R} \frac{f(x-t)-f(x+t)}{t}\,\mathrm{d}t, \quad x \in \mathbb{R},$$

noting that if $f \in C^1_{\mathrm{c}}(\mathbb{R};X)$ this integral converges as $R \to \infty$ and $\varepsilon \downarrow 0$.

We are more interested in determining the Banach spaces X such that Hf exists for every $f \in L^p(\mathbb{R};X)$ and defines a bounded operator on this space. In Example 2.1.16 we have already seen that the spaces c_0 and ℓ^1 do not fall into this class. The first main result of this chapter is the following theorem, which shows that the UMD property of X is necessary and sufficient for the boundedness of the Hilbert transform on $L^p(\mathbb{R};X)$:

Theorem 5.1.1 (Burkholder and Bourgain). *Let X be a Banach space and let $p \in (1,\infty)$ be fixed. The following assertions are equivalent:*

(1) *X is a UMD space;*
(2) *for every $f \in L^p(\mathbb{R};X)$, the limit (5.1) exists in $L^p(\mathbb{R};X)$ and satisfies*

$$\|Hf\|_p \lesssim \|f\|_p$$

with a constant independent of f.

When these conditions hold, for every $f \in L^p(\mathbb{R};X)$ the limit (5.1) also exists pointwise almost everywhere.

These conditions are furthermore equivalent to similar statements concerning a periodic analogue of the Hilbert transform on $L^p(\mathbb{T};X)$, and there are quantitative estimates between the constants involved in the different conditions. This will be made precise later on in the chapter. A two-sided estimate of the norm of the Hilbert transform in terms of the UMD constants will be provided in Corollary 5.2.11.

The reason of restricting ourselves to $p \in (1,\infty)$ is explained by the following:

Example 5.1.2. Already for scalar functions, the Hilbert transform is unbounded on both $L^1(\mathbb{R})$ and $L^\infty(\mathbb{R})$. In fact, for $f = \mathbf{1}_{(a,b)} \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$, it is directly verified that

$$H\mathbf{1}_{(a,b)}(x) = \frac{1}{\pi}\log\Big|\frac{x-a}{x-b}\Big|.$$

In the neighbourhood of the points $c \in \{a,b\}$, the logarithmic singularity belongs to $L^p([c-\varepsilon, c+\varepsilon])$ for all $p \in [1,\infty)$, but of course not to $L^\infty(\mathbb{R})$. As $x \to \pm\infty$,

$$\log\Big|\frac{x-a}{x-b}\Big| = \log\Big|\frac{1-a/x}{1-b/x}\Big| = \log(1-\frac{a}{x}) - \log(1-\frac{b}{x}) = \frac{b-a}{x} + O(\frac{1}{x^2}),$$

which is in $L^p(\complement[-r,r])$ for all $p \in (1,\infty]$ but not in $L^1(\mathbb{R})$.

5.1.a Dyadic shifts and their averages

We start by introducing several notions that serve as a basis of dyadic analysis on the line $\mathbb{R}$. The operators and estimates that emerge in this way may seem a bit exotic at first sight, but in the next subsection we will proceed to identify them with more familiar objects, in particular the Hilbert transform, showing that the exotic bounds we here obtain are no less than the $L^p(\mathbb{R};X)$-boundedness of the Hilbert transform when X is a UMD space, as asserted in the implication (1)⇒(2) of Theorem 5.1.1.

It will be helpful to recall some facts from Subsection 4.2.b. A *dyadic system* (of intervals on $\mathbb{R}$) is a family $\mathscr{D} = \bigcup_{j\in\mathbb{Z}} \mathscr{D}_j$, where each $\mathscr{D}_j$ is a partition of $\mathbb{R}$ consisting of intervals of the form $[x, x+2^{-j})$, and each interval $I \in \mathscr{D}_j$ is a union of two intervals I_- and I_+ (its left and right halves) from $\mathscr{D}_{j+1}$. The particular example $\mathscr{D}^0 = \bigcup_{j\in\mathbb{Z}} \mathscr{D}^0_j$ with $\mathscr{D}^0_j = \{2^{-j}[k,k+1) : k \in \mathbb{Z}\}$ is called the *standard dyadic system*. We further recall that to every interval I, we have associated the *Haar function*

$$h_I := |I|^{-1/2}(\mathbf{1}_{I_-} - \mathbf{1}_{I_+}),$$

where I_- and I_+ are the left and right halves of I. We observe that

$$h_I = |I|^{-1/2} h\Big(\frac{x - \inf I}{|I|}\Big), \tag{5.2}$$

where $h := h_{[0,1)} = \mathbf{1}_{[0,\frac{1}{2})} - \mathbf{1}_{[\frac{1}{2}-1)}$.

Two intervals $I, J \in \mathscr{D}$ are either disjoint or one is contained in the other, and if, say, $I \subsetneq J$, then h_J is constant on the support of h_I. As we have seen, this implies that the family $(D_I)_{I\in\mathscr{D}}$ of *Haar projections*

$$D_I f := h_I \langle h_I, f\rangle$$

is a pre-decomposition (in the sense of Definition 4.1.8) of $L^p(\mathbb{R}; X)$, for any $p \in [1,\infty]$ and any Banach space X.

Remark 5.1.3. Consider a formal series of the form

$$\sum_{I\in\mathscr{D}} f_I, \qquad \operatorname{supp} f_I \subseteq I. \tag{5.3}$$

For every $x \in \mathbb{R}$ and $j \in \mathbb{Z}$, there is a unique $I = I_j(x)$ such that $x \in I \in \mathscr{D}_j$; this is the unique $I \in \mathscr{D}_j$ such that $f_I(x)$ can be non-zero. Thus, pointwise, the series (5.3) takes the form

$$\sum_{j\in\mathbb{Z}} f_{I_j(x)}(x). \tag{5.4}$$

By the pointwise (almost everywhere) convergence of (5.3) we always understand the convergence of (5.4) in the usual sense, i.e., the existence of the limits of the finite sums $\sum_{j=m}^{n}$ as $m \to -\infty$ and $n \to +\infty$.

We recall from Definition 4.1.2 that an indexed family $(x_i)_{i\in I}$ of elements in a Banach space X is said to be *summable* to an element $x \in X$ if for all $\varepsilon > 0$ there is a finite subset $F_\varepsilon \subseteq I$ such that if $F \subseteq I$ is a finite set containing F_ε, then

$$\Big\| x - \sum_{i\in F} x_i \Big\| < \varepsilon.$$

Proposition 5.1.4. *Let X be a UMD space , $p \in (1,\infty)$, and $\mathscr{D}$ be a dyadic system on $\mathbb{R}$. For every $f \in L^p(\mathbb{R}; X)$, we have*

$$f = \sum_{i\in\mathscr{D}} D_I f = \sum_{I\in\mathscr{D}} h_I\langle h_I, f\rangle$$

in the sense of summability in $L^p(\mathbb{R}; X)$, and almost everywhere in the sense of Remark 5.1.3.

Moreover, the following estimates hold:

$$\begin{aligned}\frac{1}{\beta^-_{p,X}}\|f\|_{L^p(\mathbb{R};X)} &\leqslant \Big\| \sum_{I\in\mathscr{D}} \varepsilon_I h_I\langle h_I, f\rangle \Big\|_{L^p(\mathbb{R}\times\Omega;X)} \\ &= \Big\| \sum_{I\in\mathscr{D}} \varepsilon_{|I|} h_I\langle h_I, f\rangle \Big\|_{L^p(\mathbb{R}\times\Omega;X)} \leqslant \beta^+_{p,X}\|f\|_{L^p(\mathbb{R};X)},\end{aligned} \tag{5.5}$$

where $(\varepsilon_I)_{I\in\mathscr{D}}$ and $(\varepsilon_{2^j})_{j\in\mathbb{Z}}$ are Rademacher sequences with different index sets on a probability space Ω.

Proof. The $L^p(\mathbb{R};X)$-summability $f = \sum_{I\in\mathscr{D}} h_I\langle h_I, f\rangle$ and the upper and lower bounds for the first of the two norms in the middle of (5.5) are just restatements from Theorem 4.2.13. Since $D_{I_j(x)}f(x) = \langle f\rangle_{I_{j+1}(x)} - \langle f\rangle_{I_j(x)}$ (cf. the notation for averages introduced in (4.13)), we have a telescopic sum

$$\sum_{j=m}^{n} D_{I_j(x)}f(x) = \langle f\rangle_{I_{n+1}(x)} - \langle f\rangle_{I_m(x)},$$

where, as $n\to+\infty$ and $m\to-\infty$, the first term tends to $f(x)$ at every Lebesgue point x of f by Lebesgue's differentiation theorem, while the second term tends to zero by an elementary estimate using Hölder's inequality.

Turning to the yet unproved equality in (5.5), for either choice of $\eta_I\in\{\varepsilon_I,\varepsilon_{|I|}\}$, we have from Remark 5.1.3 that

$$\begin{aligned}\Big\|\sum_{I\in\mathscr{D}}\eta_I D_I f\Big\|_{L^p(\mathbb{R}\times\Omega;X)} &= \Big(\int_{\mathbb{R}}\mathbb{E}\Big\|\sum_{I\in\mathscr{D}}\eta_I D_I f(x)\Big\|_X^p\,\mathrm{d}x\Big)^{1/p}\\ &= \Big(\int_{\mathbb{R}}\mathbb{E}\Big\|\sum_{j\in\mathbb{Z}}\eta_{I_j(x)} D_{I_j(x)} f(x)\Big\|_X^p\,\mathrm{d}x\Big)^{1/p}.\end{aligned}$$

At every $x\in\mathbb{R}$, we observe that $|I_j(x)| = 2^{-j}$ takes a different value for every $j\in\mathbb{Z}$, and hence $(\eta_{I_j(x)})_{j\in\mathbb{Z}}$ is a Rademacher sequence, and thus has the same distribution, for both choices of η_I. This proves the equality. $\square$

We next consider so-called dyadic shift operators, which will be denoted by S_k. By the summable representation $f = \sum_{I\in\mathscr{D}} h_I\langle h_I, f\rangle$, any bounded linear operator on $L^p(\mathbb{R};X)$ can be defined by specifying its values on the vector multiples $h_I\otimes x$ of the Haar functions. Our dyadic shifts will be operators of a special form such that the Haar functions h_I are mapped into another function system k_I similarly determined by a single function

$$k_I(x) = |I|^{-1/2}k\Big(\frac{x-\inf I}{|I|}\Big), \qquad \operatorname{supp} k\subseteq[0,1) \tag{5.6}$$

that is *$\mathscr{D}$-admissible* in the sense that it has the form

$$k\in\operatorname{span}\{h_J : J\in\mathscr{D}_\ell([0,1))\}, \quad \|k\|_\infty\leqslant 1 \tag{5.7}$$

for some fixed $\ell\in\mathbb{N}$. The reader should compare (5.6) with the expression for h_I of (5.2). The shift S_k associated with k is then defined by

$$S_k f = S_k\sum_{I\in\mathscr{D}} h_I\langle h_I, f\rangle := \sum_{I\in\mathscr{D}} k_I\langle h_I, f\rangle. \tag{5.8}$$

We next prove that this formal series defines a bounded operator on the Bochner spaces $L^p(\mathbb{R};X)$.

Proposition 5.1.5. *Let X be a UMD space and let $p \in (1, \infty)$ be given. Let k be a $\mathscr{D}$-admissible function and let k_I be the functions defined above, that is,* (5.6) *and* (5.7) *are satisfied. Under these assumptions, for every $f \in L^p(\mathbb{R}; X)$ the series* (5.8) *is summable in $L^p(\mathbb{R}; X)$ and convergent almost everywhere in the sense Remark 5.1.3, and satisfies*

$$\frac{1}{\beta_{p,X}^-}\|S_k f\|_{L^p(\mathbb{R};X)} \leqslant \Big\| \sum_{I\in\mathscr{D}} \varepsilon_{|I|} k_I \langle h_I, f\rangle \Big\|_{L^p(\mathbb{R}\times\Omega;X)} \leqslant \beta_{p,X}^+ \|f\|_{L^p(\mathbb{R};X)}. \tag{5.9}$$

Proof. By the assumptions on the functions k_I, if $I \in \mathscr{D}_j$, we have

$$k_I = \sum_{j\in\Theta} \alpha_j h_{I_j},$$

where Θ is a finite index set, α_j are coefficients, and $I_j \subseteq I$ are dyadic subintervals of length $|I_j| = 2^{-\ell}|I|$.

If f has a finite Haar expansion, then $S_k f$ is clearly well defined, and by Proposition 5.1.4 we have

$$\begin{aligned}\frac{1}{\beta_{p,X}^-}\|S_k f\|_{L^p(\mathbb{R};X)} &= \frac{1}{\beta_{p,X}^-}\Big\| \sum_{j\in\Theta}\sum_{I\in\mathscr{D}} \alpha_j h_{I_j}\langle h_I, f\rangle \Big\|_{L^p(\mathbb{R};X)} \\ &\leqslant \Big\| \sum_{j\in\Theta}\sum_{I\in\mathscr{D}} \alpha_j \varepsilon_{|I_j|} h_{I_j}\langle h_I, f\rangle \Big\|_{L^p(\mathbb{R}\times\Omega;X)} =: A.\end{aligned}$$

The sequence $\varepsilon_{|I_j|} = \varepsilon_{2^{-\ell}|I|}$, which is independent of j, is another Rademacher sequence like $\varepsilon_{|I|}$, and hence by equality of distributions, we have

$$A = \Big\| \sum_{j\in\Theta}\sum_{I\in\mathscr{D}} \alpha_j \varepsilon_{|I|} h_{I_j}\langle h_I, f\rangle \Big\|_{L^p(\mathbb{R}\times\Omega;X)} = \Big\| \sum_{I\in\mathscr{D}} \varepsilon_{|I|} k_I\langle h_I, f\rangle \Big\|_{L^p(\mathbb{R}\times\Omega;X)}.$$

This completes the proof of the first bound in (5.9). For the second bound, we use the contraction principle pointwise at every $x \in \mathbb{R}$, observing that $|k_I(x)| \leqslant |h_I(x)|$:

$$\Big\| \sum_{I\in\mathscr{D}} \varepsilon_{|I|} k_I\langle h_I, f\rangle \Big\|_{L^p(\mathbb{R}\times\Omega;X)} \leqslant \Big\| \sum_{I\in\mathscr{D}} \varepsilon_{|I|} h_I\langle h_I, f\rangle \Big\|_{L^p(\mathbb{R}\times\Omega;X)},$$

and this is bounded by $\beta_{p,X}^+\|f\|_{L^p(\mathbb{R};X)}$ by Proposition 5.1.4.

Now that the quantitative part of the assertion is established for a dense class of functions, the $L^p(\mathbb{R}; X)$ summability and the estimate for all $f \in L^p(\mathbb{R}; X)$ easily follows.

As for pointwise convergence, we can apply Proposition 5.1.4 to $S_k f$ in place of f to get the almost everywhere convergence of the Haar expansion of $S_k f$. But, by simple shift of indexing,

$$\sum_{j\in\mathbb{Z}} D_{I_j(x)} S_k f(x) = \sum_{j\in\mathbb{Z}} D_{I_{j+\ell}(x)} S_k f(x) = \sum_{j\in\mathbb{Z}} k_{I_j(x)}\langle h_{I_j(x)}, f\rangle,$$

so the convergence of the first sum is equivalent to the convergence of the last one. □

Recall that our eventual aim is to relate the shifts S_k to the Hilbert transform H. To motivate the next step of our programme, we pause for a moment to discuss some basic feature of the operator that we would like to study:

Remark 5.1.6 (Invariance properties of the Hilbert transform). Recall the dilation and translation operators $\delta_r f(x) := f(rx)$ and $\tau_h f(x) := f(x-h)$ for $r \in (0,\infty)$ and $h \in \mathbb{R}$. Both these are clearly bounded operators on $L^p(\mathbb{R};X)$ for all $p \in [1,\infty]$ and all Banach spaces X.

Simple changes of variables in the defining formula

$$H_{\varepsilon,R}f(x) = \frac{1}{\pi} \int\limits_{\varepsilon<|y-x|<R} \frac{f(y)}{x-y}\,\mathrm{d}y$$

of the truncated Hilbert transform show that

$$H_{\varepsilon,R}\delta_r f = \delta_r H_{\varepsilon r,Rr} f, \qquad H_{\varepsilon,R}\tau_h f = \tau_h H_{\varepsilon,R} f.$$

Hence, if the actual Hilbert transform Hf exists, so do $H\delta_r f$ and $H\tau_h f$, and

$$H\delta_r f = \delta_r H f, \qquad H\tau_h f = \tau_h H f.$$

These properties are referred to as the invariance of H under dilations and translations

Any shift operator S_k related to a fixed dyadic system fails such invariances miserably: If $f = h_I$ for an $I \in \mathscr{D}$, then $S_k f = k_I$, but if $f = h_J$, where $J \notin \mathscr{D}$ is a slightly translated or dilated version of I, then $S_k f$ is given by a much more complicated expression. We wish to overcome this shortcoming by considering all possible dyadic systems, and eventually averaging over all of them. To this end, we first seek to parametrise the family of all dyadic systems in a convenient way to make it a probability space:

Lemma 5.1.7. *Let $\mathscr{D}^0$ be the standard dyadic system*

$$\mathscr{D}^0 = \bigcup_{j\in\mathbb{Z}} \mathscr{D}_j^0, \qquad \mathscr{D}_j^0 = \{2^{-j}[k,k+1) : \ k \in \mathbb{Z}\}.$$

Then every dyadic system $\mathscr{D}$ has the form

$$\mathscr{D}^\omega = \bigcup_{j\in\mathbb{Z}} \mathscr{D}_j^\omega, \qquad \mathscr{D}_j^\omega := \mathscr{D}_j^0 + \sum_{i>j} 2^{-i}\omega_i$$

for some $\omega \in \{0,1\}^{\mathbb{Z}}$.

Proof. It is easy to see that $\mathscr{D}_j$ has to be of the form $\mathscr{D}_j^0 + x_j$ for some $x_j \in \mathbb{R}$. If one adds an integer multiple of 2^{-j} to x_j, the collection $\mathscr{D}_j^0 + x_j$ does not change, so one can demand that $x_j \in [0, 2^{-j})$. Then x_j is actually the unique end-point of intervals in $\mathscr{D}_j$ which belongs to the interval $[0, 2^{-j})$. Since this is also an end-point of the intervals in $\mathscr{D}_{j+1}$, there must hold $x_j - x_{j+1} \in \{0, 2^{-j-1}\}$. Let us write $\omega_{j+1} := 2^{j+1}(x_j - x_{j+1}) \in \{0,1\}$ so that $x_j = 2^{-j-1}\omega_{j+1} + x_{j+1}$, and by iteration

$$x_j = \sum_{i>j} 2^{-i}\omega_i, \qquad \omega = (\omega_i)_{i\in\mathbb{Z}} \in \{0,1\}^{\mathbb{Z}}.$$

The claim follows. □

The parametrising set $\{0,1\}^{\mathbb{Z}}$ will be equipped with the product σ-algebra, i.e., the smallest σ-algebra so that the coordinate functions $\omega \mapsto \omega_i$ are measurable. In the sequel we will also need *dilated* dyadic systems

$$t\mathscr{D}^\omega := \{tI = [ta, tb) : I = [a,b) \in \mathscr{D}^\omega\}.$$

Note that $2^j\mathscr{D}^\omega = \mathscr{D}^{\omega'}$ for another $\omega' \in \{0,1\}^{\mathbb{Z}}$, so only the dilation factors $t \in [1,2)$ will be relevant.

Consider next the family of dyadic shifts $S_k = S_k^{\omega,r}$ as above for every dyadic system $t\mathscr{D}^\omega$, where we use the same function k as in (5.7) for every (ω, r). We shall consider the average of these operators with respect to a product probability measure on $\{0,1\}^{\mathbb{Z}} \times [1,2)$. On $\{0,1\}^{\mathbb{Z}}$, we use the natural probability, denoted simply by $\mathrm{d}\omega$, which makes the coordinates ω_i independent with the values 0 and 1 equally likely. On $[1,2)$ we may, for the moment, proceed with an arbitrary probability measure ν, although only two specific choices will be relevant later on.

The average dyadic shift related to ν is the following integral:

$$\langle S_k\rangle^\nu f(x) := \int_{[1,2)}\int_{\{0,1\}^{\mathbb{Z}}} S_k^{\omega,r} f(x)\,\mathrm{d}\omega\,\mathrm{d}\nu(r).$$

Lemma 5.1.8. *Let X be a UMD space, let $p \in (1,\infty)$ be given, and let k be $\mathscr{D}$-admissible. Then for all $f \in L^p(\mathbb{R};X)$, the integral defining $\langle S_k\rangle^\nu f$ exists in the following senses:*

(1) *The function $(\omega,r) \mapsto S_k^{\omega,r} f(x) \in X$ is Bochner integrable for almost every $x \in \mathbb{R}$.*
(2) *The function $(\omega,r) \mapsto S_k^{\omega,r} f \in L^p(\mathbb{R};X)$ is Bochner integrable.*

Moreover,

$$\langle S_k\rangle^\nu f(x) = \sum_{j=-\infty}^{\infty}\int_{[1,2)}\int_{\{0,1\}^{\mathbb{Z}}} \sum_{I\in r\mathscr{D}_j^\omega} k_I(x)\langle h_I, f\rangle\,\mathrm{d}\omega\,\mathrm{d}\nu(r), \tag{5.10}$$

both pointwise for almost every $x \in \mathbb{R}$ and in the sense of convergence in $L^p(\mathbb{R};X)$. The series in j is understood as the limit of the truncated series $\sum_{j=-m}^{n}$ as $m,n \to \infty$ independently.

Proof. Recall that

$$S_k^{\omega,r} f(x) = \sum_{j=-\infty}^{\infty} \sum_{I \in r\mathscr{D}_j^{\omega}} k_I(x)\langle h_I, f\rangle, \tag{5.11}$$

and the inner sum may be written with the summation variable $m \in \mathbb{Z}$, so that

$$I = 2^{-j} r\big([0,1) + m + \omega^{(j)}\big), \qquad \omega^{(j)} := \sum_{i>j} 2^{j-i}\omega_i \in [0,1]. \tag{5.12}$$

The value of

$$\langle h_I, f\rangle = |I|^{-1/2}\Big(\int_{I_-} f(y)\,\mathrm{d}y - \int_{I_+} f(y)\,\mathrm{d}y\Big), \quad I = 2^{-j} r\big([0,1) + m + \omega^{(j)}\big)$$

depends continuously on $(\omega^{(j)}, r)$, hence measurably on (ω, r), and the function

$$k_I(x) = 2^{j/2} r^{-1/2} k(2^j r^{-1} x - m - \omega^{(j)})$$

is measurable with respect to (ω, r, x). Then also (5.11) is measurable in both required senses, as a limit, both pointwise and in $L^p(\mathbb{R}; X)$, of measurable functions. □

Having settled the meaningfulness of the average dyadic shift, we turn to the investigation of what this object actually looks like. As it turns out, the averaging over the translation parameter $\omega \in \{0,1\}^{\mathbb{Z}}$ produces a convolution operator. The following lemma is just a computation with no deep convergence issues, and can therefore be stated in a greater generality than the previous two.

Lemma 5.1.9. *Let X be a Banach space, let k be a $\mathscr{D}$-admissible function, and let k_I be as before. Then, for all $f \in L^1_{\mathrm{loc}}(\mathbb{R}; X)$,*

$$\int_{\{0,1\}^{\mathbb{Z}}} \sum_{I \in r\mathscr{D}_j^{\omega}} k_I(x)\langle h_I, f\rangle\,\mathrm{d}\omega = \phi_{2^{-j}r} * f(x), \tag{5.13}$$

where $\phi_t(x) := t^{-1}\phi(t^{-1}x)$ is the L^1-dilation of the function

$$\phi(x) := \int_{-\infty}^{\infty} k(x+u)h(u)\,\mathrm{d}u, \qquad h := h_{[0,1)} := \mathbf{1}_{[0,1/2)} - \mathbf{1}_{[\frac{1}{2},1)}. \tag{5.14}$$

Proof. Observe that $\omega^{(j)}$, as defined in (5.12), is uniformly distributed on $[0,1]$ under the probability $\mathrm{d}\omega$. Hence, with $t := 2^{-j}r$,

$$\int_{\{0,1\}^{\mathbb{Z}}} \sum_{I \in \mathscr{D}_j^{\omega}} k_I(x)\langle h_I, f\rangle\,\mathrm{d}\omega$$

$$
\begin{aligned}
&= \int_{\{0,1\}^{\mathbb{Z}}} \sum_{m\in\mathbb{Z}} k_{t([0,1)+m+\omega^{(j)})}(x)\langle h_{t([0,1)+m+\omega^{(j)})}, f\rangle \,\mathrm{d}\omega \\
&= \int_0^1 \sum_{m\in\mathbb{Z}} k_{t([0,1)+m+u)}(x)\langle h_{t([0,1)+m+u)}, f\rangle \,\mathrm{d}u \\
&= \int_{-\infty}^{\infty} k_{t([0,1)+v)}(x)\langle h_{t([0,1)+v)}, f\rangle \,\mathrm{d}v \\
&= \int_{-\infty}^{\infty} t^{-1/2}k(t^{-1}x-v)\int_{-\infty}^{\infty} t^{-1/2}h(t^{-1}y-v)f(y)\,\mathrm{d}y\,\mathrm{d}v \\
&= \int_{-\infty}^{\infty} t^{-1}\int_{-\infty}^{\infty} k(t^{-1}x-v)h(t^{-1}y-v)\,\mathrm{d}v f(y)\,\mathrm{d}y \\
&=: \int_{-\infty}^{\infty} \varPhi(t,x,y)f(y)\,\mathrm{d}y,
\end{aligned}
$$

and, substituting $u = t^{-1}y - v$,

$$
\begin{aligned}
\varPhi(t,x,y) &= t^{-1}\int_{-\infty}^{\infty} k(t^{-1}(x-y)+u)h(u)\,\mathrm{d}u \\
&= t^{-1}\phi(t^{-1}(x-y)) = \phi_t(x-y)
\end{aligned}
$$

so that

$$
\int_{-\infty}^{\infty} \varPhi(t,x,y)f(y)\,\mathrm{d}y = \phi_t * f(x),
$$

as claimed. □

Let us put the pieces together. Under the assumptions (5.6) and (5.7) on the function system k_I in the definition (5.8) of the dyadic shift, we have, by chaining (5.10) and (5.13), that

$$
\langle S_k\rangle^{\nu} f = \sum_{j=-\infty}^{\infty} \int_{[1,2)} \phi_{2^{-j}r} * f \,\mathrm{d}\nu(r). \tag{5.15}
$$

Let us further extend the measure ν from $[1,2)$ to $(0,\infty)$ by setting $\nu(2^{-j}A) := \nu(A)$ for $j\in\mathbb{Z}$ and $A\subseteq[1,2)$. Then (5.15) can be rewritten as

$$
\langle S_k\rangle^{\nu} f = \int_0^{\infty} \phi_t * f \,\mathrm{d}\nu(t) := \lim_{m,n\to\infty}\int_{2^{-m}}^{2^n} \phi_t * f\,\mathrm{d}\nu(t), \tag{5.16}
$$

where the indefinite integral should be interpreted in terms of the pointwise limit along the powers of two as indicated, as only this convergence is directly justified by the convergence of the series in (5.15). With two particular choices for the measure ν, we then have

$$\begin{aligned} \langle S_k\rangle^{\delta_1} f &= \sum_{j=-\infty}^{\infty} \phi_{2^j} * f, \\ \langle S_k\rangle^{c\,\mathrm{d}r/r} f &= c\int_0^\infty \phi_t * f \,\frac{\mathrm{d}t}{t}, \qquad c = \frac{1}{\log 2}, \end{aligned} \tag{5.17}$$

where δ_1 is the Dirac measure at $t = 1$ and the choice of c ensures that $c\int_1^2 \mathrm{d}r/r = 1$.

We have the following proposition, where the first estimate, with a specific choice of k, will be used for the proof of the Hilbert transform boundedness (assuming UMD), while the second is our preliminary version of the Littlewood–Paley inequality.

Proposition 5.1.10. *Let X be a UMD space and let $p \in (1,\infty)$. Let k be $\mathscr{D}$-admissible function in the sense of (5.7), and set*

$$\phi(x) := \int_{-\infty}^{\infty} k(x+y)h(y)\,\mathrm{d}y \tag{5.18}$$

with h as in (5.14). Then the series and the integral in (5.17), the latter interpreted as in (5.16), converge both pointwise almost everywhere and in $L^p(\mathbb{R};X)$, and the following inequalities hold:

$$\Big\| \int_0^\infty \phi_t * f \,\frac{\mathrm{d}t}{t}\Big\|_{L^p(\mathbb{R};X)} \leqslant \log 2 \cdot \beta^-_{p,X}\beta^+_{p,X}\|f\|_{L^p(\mathbb{R};X)},$$

$$\Big\| \sum_{j=-\infty}^{\infty} \varepsilon_j \phi_{2^j} * f\Big\|_{L^p(\mathbb{R}\times\Omega;X)} \leqslant \beta^+_{p,X}\|f\|_{L^p(\mathbb{R};X)}.$$

Proof. The first estimate follows from Proposition 5.1.5 by observing that $\langle S_k\rangle^{c\,\mathrm{d}r/r} f$ is exactly the indefinite integral that we wanted to estimate by the second equality of (5.17).

$$\begin{aligned} \|\langle S_k\rangle^{c\,\mathrm{d}r/r} f\|_p &\leqslant \frac{1}{\log 2}\int_{[1,2)}\int_{\{0,1\}^{\mathbb{Z}}} \|S_k^{\omega,r} f\|_p \,\mathrm{d}\omega\,\frac{\mathrm{d}r}{r} \\ &\leqslant \frac{1}{\log 2}\int_{[1,2)}\int_{\{0,1\}^{\mathbb{Z}}} \beta^-_{p,X}\beta^+_{p,X}\|f\|_p\,\mathrm{d}\omega\,\frac{\mathrm{d}r}{r} = \beta^-_{p,X}\beta^+_{p,X}\|f\|_p. \end{aligned}$$

The second estimate is proved similarly by using only the second, randomised bound in Proposition 5.1.5, and the first equality of (5.17). □

Example 5.1.11. Two specific choices of the function k will be of relevance to us below, namely $h_{[0,1)} = \mathbf{1}_{[0,1/2)} - \mathbf{1}_{[\frac12,1)}$ and

$$2^{-1/2}(h_{[0,1/2)} - h_{[\frac12,1)}) = \mathbf{1}_{[0,1/4)\cup[3/4,1)} - \mathbf{1}_{[1/4,3/4)}. \tag{5.19}$$

One easily computes the corresponding functions ϕ defined through (5.18) In both cases, ϕ is a piecewise affine function supported on $[-1,1]$; for the first choice of k, it is even, takes the values $1, -\frac{1}{2}, 0$ at the points $0, \frac{1}{2}, 1$ and interpolates linearly between them, whereas for the second choice it is odd and interpolates between the values $0, -\frac{3}{4}, 0, \frac{1}{4}, 0$ at $0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$. (See Figure 5.1.)

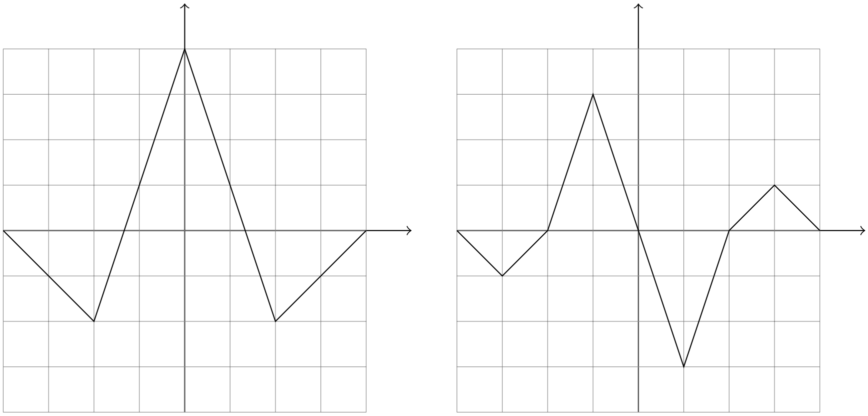

Fig. 5.1: Graphs of the two choices of ϕ in Example 5.1.11. The grid size is $\frac{1}{4}$ units.

5.1.b The Hilbert transform from the dyadic shifts

It is time to relate the exotic bounds that we have derived to something of more familiar form, in particular, the Hilbert transform.

We now fix k as the function in (5.19), and $\phi(x) = \int_{-\infty}^{\infty} k(x+u)h(u)\,\mathrm{d}u$ correspondingly. The aim is to identify the average dyadic shift

$$\langle S_k\rangle^{c\,\mathrm{d}r/r} f(x) = \frac{1}{\log 2} \lim_{m,n\to\infty} \int_{2^{-m}}^{2^n} \phi_t * f(x)\,\frac{\mathrm{d}t}{t}, \quad \phi_t(x) = \frac{1}{t}\phi\Big(\frac{x}{t}\Big), \quad (5.20)$$

where $f \in L^p(\mathbb{R}; X)$, with a constant multiple of the Hilbert transform Hf of f.

Changing the order of integration in (5.20), we have

$$\int_{\varepsilon}^{R} \phi_t * f(x)\,\frac{\mathrm{d}t}{t} = \int_{-\infty}^{\infty} \Big[\int_{\varepsilon}^{R} \phi_t(y)\,\frac{\mathrm{d}t}{t}\Big] f(x-y)\,\mathrm{d}y$$

and

$$\int_{\varepsilon}^{R} \phi_t(y)\,\frac{\mathrm{d}t}{t} = \int_{\varepsilon}^{R} \phi\Big(\frac{y}{t}\Big)\frac{1}{t^2}\,\mathrm{d}t = \frac{1}{y}\Big[\Phi\Big(\frac{y}{\varepsilon}\Big) - \Phi\Big(\frac{y}{R}\Big)\Big],$$

where

$$\Phi(x) := \int_0^x \phi(s)\,\mathrm{d}s \tag{5.21}$$

is the primitive of ϕ. Here $f \in L^p(\mathbb{R};X)$, we have $p \in (1,\infty)$, and X could be any Banach space for the above computation, but the convergence of these integrals as $\varepsilon \downarrow 0$ and $R \to \infty$ along the dyadic powers 2^n, $n \in \mathbb{Z}$, is only guaranteed when X is a UMD space.

From the fact that ϕ is odd, it follows that Φ is even. Since ϕ is supported on $[-1,1]$, its primitive Φ is a constant on the complement, and in fact $\Phi(x) = -\frac{1}{8}$ for $|x| \geqslant 1$. Write

$$\psi(x) := x^{-1}\Phi(x)\mathbf{1}_{[-1,1]}(x), \tag{5.22}$$

which is again an odd and bounded function. Then

$$\frac{1}{y}\Phi\Big(\frac{y}{\varepsilon}\Big) = \frac{1}{\varepsilon}\frac{\varepsilon}{y}\Big(\Phi\Big(\frac{y}{\varepsilon}\Big)\mathbf{1}_{[-1,1]}\Big(\frac{y}{\varepsilon}\Big) - \frac{1}{8}\mathbf{1}_{\complement[-1,1]}\Big(\frac{y}{\varepsilon}\Big)\Big) = \psi_\varepsilon(y) - \frac{1}{8y}\mathbf{1}_{|y|>\varepsilon},$$

hence

$$\frac{1}{y}[\Phi(\frac{y}{\varepsilon}) - \Phi(\frac{y}{R})] = \psi_\varepsilon(y) - \psi_R(y) - \frac{1}{8y}\mathbf{1}_{\varepsilon<|y|\leqslant R}.$$

Therefore

$$\int_\varepsilon^R \phi_t * f\,\frac{\mathrm{d}t}{t} = \psi_\varepsilon * f - \psi_R * f - \frac{\pi}{8}H_{\varepsilon,R}f, \quad f \in L^p(\mathbb{R};X),$$

and finally

$$\begin{aligned} -\frac{\pi}{8}H_{\varepsilon,R}f &= \int_{2^{\lfloor \log_2 \varepsilon\rfloor}}^{2^{\lfloor \log_2 R\rfloor}} \phi_t * f\,\frac{\mathrm{d}t}{t} \\ &\quad + \Big(\int_{2^{\lfloor \log_2 R\rfloor}}^{R} - \int_{2^{\lfloor \log_2 \varepsilon\rfloor}}^{\varepsilon}\Big)\phi_t * f\,\frac{\mathrm{d}t}{t} - \psi_\varepsilon * f + \psi_R * f. \end{aligned} \tag{5.23}$$

As $\varepsilon \downarrow 0$ and $R \to \infty$, the first term on the right-hand side approaches $\log 2\cdot \langle S_k\rangle^{(\log 2)^{-1}\,\mathrm{d}r/r}$, both almost everywhere and in $L^p(\mathbb{R};X)$. So to complete the proof of the existence of the Hilbert transform Hf, i.e., the limit of $H_{\varepsilon,R}f$ as $\varepsilon \downarrow 0$ and $R \to \infty$, it remains to establish the convergence of the error terms on the right-hand side of (5.23). And indeed we will prove:

Lemma 5.1.12. *Let X be a Banach space, $p \in (1,\infty)$, and $f \in L^p(\mathbb{R};X)$. Then all the terms on the second line of (5.23) converge to 0, both pointwise almost everywhere and in $L^p(\mathbb{R};X)$, as $\varepsilon \downarrow 0$ and $R \to \infty$.*

Assuming this for the moment, we have everything to complete the proof of the implication "⇒" of Theorem 5.1.1, and more precisely we get the following quantitative version of this statement:

Theorem 5.1.13. *Let X be a UMD space, $p \in (1,\infty)$, and $f \in L^p(\mathbb{R};X)$. Then the Hilbert transform Hf exists and satisfies*

$$Hf = -\frac{8}{\pi} \lim_{\substack{\varepsilon\to 0\\ R\to\infty}} \int_\varepsilon^R \phi_t * f \,\frac{\mathrm{d}t}{t},$$

and

$$\|Hf\|_p \leqslant 2\beta^+_{p,X}\beta^-_{p,X}\|f\|_p, \tag{5.24}$$

where the limit in the identity is attained both pointwise almost everywhere and in $L^p(\mathbb{R};X)$, and ϕ is the odd, piecewise affine function supported on $[-1,1]$, which interpolates linearly between the values $0,-\frac{3}{4},0,\frac14,0$ at $0,\frac14,\frac12,\frac34,1$.

Proof. The identity is immediate from (5.23) and Lemma 5.1.12. The inequality is then a consequence of Proposition 5.1.10. □

Remark 5.1.14. The numerical factor 2 in (5.24) is an artifact of our discretisation scheme, and can be improved to 1 by other methods that we discuss in the Notes. (In fact, our proof would give the constant $8/\pi\cdot\log 2\approx 1.7651$, but in view of the remark just made, there is no point in insisting too much on this value.) A more serious question is whether the product $\beta^+_{p,X}\beta^-_{p,X}$, which is of course dominated by $\beta^2_{p,X}$, can be replaced by a linear bound $C_p\beta_{p,X}$ in terms of the UMD constant. See Problem O.6 and the discussion in the Notes at the end of this chapter.

The error terms

We will now give the postponed proof of Lemma 5.1.12. The function ϕ appearing in the assertion is as in Theorem 5.1.13, i.e., it is the odd function supported on $[-1,1]$ which interpolates linearly between the values $0,-\frac34,0,\frac14,0$ at $0,\frac14,\frac12,\frac34,1$, and ψ was defined in (5.21) and (5.22) as

$$\psi(x) := \mathbf{1}_{[-1,1]}(x)\cdot\frac{1}{x}\int_0^x \phi(y)\,\mathrm{d}y.$$

Both ϕ and ψ are bounded functions supported on $[-1,1]$. This readily implies that, for any $a>0$,

$$\|\psi_t * f(x)\| \leqslant CMf(x), \qquad \int_a^{2a} \|\phi_t * f(x)\|\,\frac{\mathrm{d}t}{t} \leqslant CMf(x),$$

where M is the Hardy–Littlewood maximal operator (see Section 2.3); the first inequality follows from Proposition 2.3.9 and the second from this by noting that $\int_a^{2a}\mathrm{d}t/t = \log 2$ is independent of a.

Lemma 5.1.15. *Under the assumptions of Lemma 5.1.12, we have the following limits both almost everywhere and in $L^p(\mathbb{R};X)$:*

$$\lim_{R\to\infty}\psi_R * f \to 0, \qquad \lim_{\substack{a,b\to\infty\\ a\leqslant b\leqslant 2a}} \int_a^b \phi_t * f\,\frac{\mathrm{d}t}{t} \to 0.$$

Proof. By Hölder's inequality,

$$\|\psi_R * f\|_\infty \leqslant \|\psi_R\|_{p'}\|f\|_p = R^{-1/p}\|\psi\|_{p'}\|f\|_p \underset{R\to\infty}{\longrightarrow} 0,$$

and the integrated version with ϕ is similar. Due to the maximal function control and dominated convergence, we also get the convergence in the $L^p(\mathbb{R};X)$ norm. □

Concerning the limit at zero, the standard mollification result of Proposition 1.2.32 (for the L^p-convergence) and Theorem 2.3.8 (for the almost everywhere convergence) gives

$$\lim_{\varepsilon\downarrow 0}\psi_\varepsilon * f = 0,$$

where we used that ψ is odd. Moreover, standard arguments give

$$\lim_{\substack{a,b\to 0\\ a<b\leqslant 2a}} \int_a^b \theta_t * f \,\frac{\mathrm{d}t}{t} = 0,$$

pointwise almost everywhere and in $L^p(\mathbb{R};X)$. This implies:

Lemma 5.1.16. *Under the assumptions of Lemma 5.1.12, the following limits exist both almost everywhere and in $L^p(\mathbb{R};X)$:*

$$\lim_{\varepsilon\downarrow 0}\psi_\varepsilon * f \to 0, \qquad \lim_{\substack{a,b\to 0\\ a\leqslant b\leqslant 2a}} \int_a^b \phi_t * f \,\frac{\mathrm{d}t}{t} \to 0.$$

This completes the proof of Lemma 5.1.12, and then of Theorem 5.1.13.

5.2 Fourier analysis of the Hilbert transform

Having investigated the Hilbert transform as a singular convolution operator, we now adopt a different point of view which reveals several further features of this fundamental transform, and also leads to far-reaching generalisations. The point of view that we have in mind is that of Fourier analysis, the relevant elements of which we have presented in Section 2.4. Due to the presence of complex exponentials in the very defining formulae of Fourier integrals, whenever dealing with the Fourier transform, we need to assume that X is a complex Banach space.

Nevertheless, many of the operators that we consider, notably the Hilbert transform itself, make perfect sense for real Banach spaces as well, and we shall occasionally take a few additional steps to verify that a number of key results remain valid for real spaces, even if complex methods were used as an intermediate step. The tool for achieving this is provided by the complexification method described in Subsection 2.1.b. In particular, since the Hilbert transform maps real-valued functions to real-valued ones, it is an operator in the scope of Corollary 2.1.13, which guarantees that, for real Banach spaces X, we have

$$\|H\|_{\mathscr{L}(L^p(\mathbb{R};X^{\gamma,p}_{\mathbb{C}}))} = \|H\|_{\mathscr{L}(L^p(\mathbb{R};X))},$$

where $X^{\gamma,p}_{\mathbb{C}}$ is the complexification of X endowed with the particular norm defined in Subsection 2.1.b. We shall use this observation with other similar operators as well.

5.2.a The Hilbert transform via the Fourier transform

In the Fourier analysis of the Hilbert transform, it will be convenient to work with the space $\check{L}^1(\mathbb{R}^d;X)$ of functions that are the inverse Fourier transform of a function in $L^1(\mathbb{R}^d;X)$. This space has been introduced in Subsection 2.4.a. We start by considering

$$H_{\varepsilon,R}f = k_{\varepsilon,R} * f$$

for functions $f = \check{h} \in \check{L}^1(\mathbb{R};X)$, where

$$k_{\varepsilon,R}(x) = (\pi x)^{-1}\mathbf{1}_{\varepsilon<|x|<R}.$$

Since $k_{\varepsilon,R} \in L^1(\mathbb{R})$, the operators $H_{\varepsilon,R}$ map $\check{L}^1(\mathbb{R};X)$ to itself by Lemma 2.4.8.

Lemma 5.2.1. *For $f \in \check{L}^1(\mathbb{R};X)$, the Fourier transform of the truncated Hilbert transform satisfies*

$$\widehat{H_{\varepsilon,R}f}(\xi) = -i\,\mathrm{sgn}(\xi)m_{\varepsilon,R}(\xi)\widehat{f}(\xi),$$

where $\|m_{\varepsilon,R}\|_\infty \leqslant 2$ and $m_{\varepsilon,R}(\xi) \to 1$ as $\varepsilon \downarrow 0$ and $R \to \infty$.

Proof. By Lemma 2.4.8, $\widehat{H_{\varepsilon,R}f}(\xi) = \widehat{k}_{\varepsilon,R}(\xi)\widehat{f}(\xi)$, and it remains to compute

$$\pi\widehat{k}_{\varepsilon,R}(\xi) = \int_{\varepsilon<|x|<R} e^{-2\pi i x\cdot\xi}\,\frac{\mathrm{d}x}{x} = -\,\mathrm{sgn}(\xi)\int_{i[-R,-\varepsilon]\cup i[\varepsilon,R]} e^{2\pi|\xi|z}\,\frac{\mathrm{d}z}{z}.$$

By Cauchy's theorem, the integration path may be shifted from $i[-R,-\varepsilon]\cup i[\varepsilon,R]$ to the union of the two semicircles of radii ε and R in the left half plane. Thus the above integral is equal to

$$\int_{i[-R,-\varepsilon]\cup i[\varepsilon,R]} e^{2\pi|\xi|z}\,\frac{\mathrm{d}z}{z} = \int_{\pi/2}^{3\pi/2}\big(e^{2\pi|\xi|\varepsilon e^{i\theta}} - e^{2\pi|\xi|Re^{i\theta}}\big)i\,\mathrm{d}\theta.$$

The second term is dominated by

$$\Big|\int_{\pi/2}^{3\pi/2} e^{2\pi|\xi|Re^{i\theta}}\,\mathrm{d}\theta\Big| \leqslant \int_{\pi/2}^{3\pi/2} e^{2\pi|\xi|R\cos\theta}\,\mathrm{d}\theta = 2\int_0^{\pi/2} e^{-2\pi|\xi|R\sin\theta}\,\mathrm{d}\theta$$

$$\leqslant 2\int_0^{\pi/2} e^{-4|\xi|R\theta}\,\mathrm{d}\theta = \frac{1-e^{-2\pi|\xi|R}}{2|\xi|R},$$

where the estimate $\sin\theta \geqslant 2/\pi\cdot\theta$ for $\theta\in[0,\frac{1}{2}\pi]$ was used. The result is uniformly bounded by π (since $1-e^{-x}\leqslant x$) and tends to zero as $R\to\infty$.

The same upper bound applies to the first term, which clearly tends to

$$\int_{\pi/2}^{3\pi/2} i\,\mathrm{d}\theta = i\pi$$

as $\varepsilon\downarrow 0$. This proves the lemma. □

This leads to the desired formula for $\widehat{Hf}$. As a byproduct, we obtain the existence of Hf for many functions in an arbitrary Banach space X.

Proposition 5.2.2. *Let X be a complex Banach space and $f\in \check{L}^1(\mathbb{R};X)$. Then the Hilbert transform Hf exists, belongs to $\check{L}^1(\mathbb{R};X)$, and satisfies*

$$\widehat{Hf}(\xi) = -i\,\mathrm{sgn}(\xi)\widehat{f}(\xi).$$

Proof. We already checked in Lemma 5.2.1 that $H_{\varepsilon,R}f\in\check{L}^1(\mathbb{R};X)$ with $\widehat{H_{\varepsilon,R}f}(\xi) = -i\,\mathrm{sgn}(\xi)m_{\varepsilon,R}(\xi)\widehat{f}(\xi)$; hence

$$H_{\varepsilon,R}f(x) = \int_{\mathbb{R}} -i\,\mathrm{sgn}(\xi)m_{\varepsilon,R}(\xi)\widehat{f}(\xi)e^{2\pi i x\cdot\xi}\,\mathrm{d}\xi.$$

Passing to the limit $\varepsilon\downarrow 0$, $R\to\infty$ in the above equality and using dominated convergence on the right, we conclude that

$$H_{\varepsilon,R}f(x) \to \int_{\mathbb{R}} -i\,\mathrm{sgn}(\xi)\widehat{f}(\xi)e^{2\pi i x\cdot\xi}\,\mathrm{d}\xi,$$

which gives the existence of $Hf(x)$ with the asserted formula. □

Corollary 5.2.3. *Let X be a complex Banach space.*

(1) *$H^2 f = -f$ for all $f\in\check{L}^1(\mathbb{R};X)$.*
(2) *If H maps $L^p(\mathbb{R};X)\cap\check{L}^1(\mathbb{R};X)$ into $L^p(\mathbb{R};X)$ and extends to a bounded operator on $L^p(\mathbb{R};X)$, then $H^2 f = -f$ for all $f\in L^p(\mathbb{R};X)$.*

Proof. For $f\in\check{L}^1(\mathbb{R};X)$ we have $Hf\in\check{L}^1(\mathbb{R};X)$ with $\widehat{Hf}(\xi) = -i\,\mathrm{sgn}(\xi)\widehat{f}(\xi)$, and hence $H^2 f\in\check{L}^1(\mathbb{R};X)$ with

$$\widehat{H^2 f}(\xi) = -i\,\mathrm{sgn}(\xi)\widehat{Hf}(\xi) = (-i\,\mathrm{sgn}(\xi))^2\widehat{f}(\xi) = -\widehat{f}(\xi).$$

Thus $H^2 f = (\widehat{H^2 f})^{\vee} = (-\widehat{f})^{\vee} = -f$ for $f\in\check{L}^1(\mathbb{R};X)$.

For $f\in L^p(\mathbb{R};X)$, we just approximate it by a sequence in $L^p(\mathbb{R};X)\cap\check{L}^1(\mathbb{R};X)$, and use the continuity of H, and thus of H^2, on $L^p(\mathbb{R};X)$. □

5.2.b Periodic Hilbert transform and Fourier series

Let X be a complex Banach space. The *Fourier coefficients* $\widehat{f}(k) \in X$ of a periodic function $f \in L^1(\mathbb{T}^d; X)$ are defined by

$$\widehat{f}(k) := \int_{\mathbb{T}^d} f(t) e^{-2\pi i k \cdot t} \,\mathrm{d}t, \quad k \in \mathbb{Z}^d.$$

Although we use the same notation as for the Fourier transform on $\mathbb{R}^d$, this should hopefully cause no confusion, as the meaning should be clear from the context. Since on the probability space $\mathbb{T}^d$ we have $L^p(\mathbb{T}^d; X) \subseteq L^1(\mathbb{T}^d; X)$ for $p \geqslant 1$, the Fourier coefficients are in particular defined for all $f \in L^p(\mathbb{T}^d; X)$. When f has Fourier coefficients $\widehat{f}(k)$, it is traditionally customary to write

$$f \sim \sum_{k \in \mathbb{Z}^d} \widehat{f}(k) e_k,$$

and it is a principal question of classical Fourier analysis to determine when, and in what sense, we may replace the heuristic "$\sim$" by a proper identity "$=$".

An X-valued *trigonometric polynomial* is a function $f : \mathbb{T}^d \to X$ of the form

$$f = \sum_{k \in \mathbb{Z}^d} a_k e_k$$

with finitely many non-zero coefficients $a_k \in X$; here

$$e_k(t) := \exp(i 2\pi k \cdot t).$$

It is immediate from the orthogonality of the exponentials e_k in $L^2(\mathbb{T}^d)$ that the Fourier coefficients of f are given by $\widehat{f}(k) = a_k$.

Since

$$e_k(t) = \cos(2\pi k \cdot t) + i \sin(2\pi k \cdot t),$$

where cos is even and sin is odd, a trigonometric polynomial can also be expressed in the form

$$f = a_0 + \sum_{\substack{k \in \mathbb{Z}^d \\ k \neq 0}} b_k \cos(2\pi k \cdot t) + \sum_{\substack{k \in \mathbb{Z}^d \\ k \neq 0}} c_k \sin(2\pi k \cdot t), \tag{5.25}$$

with $a_0 \in X$ and finitely many non-zero coefficients $b_k, c_k \in X$. This latter form makes no reference to complex multiplication, and may be taken as a definition of trigonometric polynomials in all Banach spaces, real or complex. With this understanding, the following result is valid in the same generality:

Proposition 5.2.4. *Let X be Banach space and let $p \in [1, \infty)$. Then trigonometric polynomials are dense in $L^p(\mathbb{T}^d; X)$.*

Proof. This follows from the fact that $L^p(\mathbb{T}^d)\otimes X$ is dense in $L^p(\mathbb{T}^d;X)$, and the well-known scalar-valued result that trigonometric polynomials are dense in $L^p(\mathbb{T}^d)$. □

The Fourier analytic point of view to the Hilbert transforms suggests a natural analogue $\widetilde{H}$ acting on $L^p(\mathbb{T};X)$. We set

$$\widetilde{H}\Big(\sum_{k\in\mathbb{Z}} a_k e_k\Big) := -i\sum_{k\in\mathbb{Z}} \operatorname{sgn}(k) a_k e_k,$$

when $\sum_{k\in\mathbb{Z}} a_k e_k$ is a trigonometric polynomial (a finitely non-zero sum). Note that, unlike on $\mathbb{R}$, it is here necessary to be specific about the meaning of $\operatorname{sgn}(0)$, and we follow the convention that $\operatorname{sgn}(0) := 0$. Thus $\widetilde{H}$ maps into functions of mean zero.

Writing sine and cosine as linear combinations of exponentials, we find that

$$\widetilde{H}\big(\cos(2\pi k\cdot)\big) = \sin(2\pi k\cdot), \qquad \widetilde{H}\big(\sin(2\pi k\cdot)\big) = -\cos(2\pi k\cdot).$$

Expressed in terms of the notation in (5.25), we get

$$\begin{aligned}\widetilde{H}\Big(a_0 + \sum_{\substack{k\in\mathbb{Z}^d\\ k\neq 0}} b_k\cos(2\pi k\cdot t) + \sum_{\substack{k\in\mathbb{Z}^d\\ k\neq 0}} c_k\sin(2\pi k\cdot t)\Big)\\ = -\sum_{\substack{k\in\mathbb{Z}^d\\ k\neq 0}} c_k\cos(2\pi k\cdot t) + \sum_{\substack{k\in\mathbb{Z}^d\\ k\neq 0}} b_k\sin(2\pi k\cdot t),\end{aligned}$$

which can be taken as the definition of $\widetilde{H}$ for real Banach spaces; it shows in particular that the periodic $\widetilde{H}$, like its analogue H on $\mathbb{R}$, maps real-valued functions to real-valued ones.

For real and complex Banach spaces alike, we say that $\widetilde{H}$ is bounded on $L^p(\mathbb{T};X)$ if

$$\|\widetilde{H}f\|_{L^p(\mathbb{T};X)} \leqslant C\|f\|_{L^p(\mathbb{T};X)}$$

uniformly for all trigonometric polynomials $f\in L^p(\mathbb{T};X)$. By density, this then gives a bounded extension of $\widetilde{H}$ to all $f\in L^p(\mathbb{T};X)$.

The following result is an example of a more general method of *transference*, but it takes a particularly clean form in the case of the Hilbert transform.

Proposition 5.2.5. *Let X be a Banach space and let $p\in(1,\infty)$ be given. If H is bounded on $L^p(\mathbb{R};X)$, then $\widetilde{H}$ is bounded on $L^p_0(\mathbb{T};X)$ and $L^p(\mathbb{T};X)$, and*

$$\|\widetilde{H}\|_{\mathscr{L}(L^p_0(\mathbb{T};X))} \leqslant \|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};X))} \leqslant \|H\|_{\mathscr{L}(L^p(\mathbb{R};X))},$$

where $L^p_0(\mathbb{T};X) := \{\phi\in L^p(\mathbb{T};X) : \int_{\mathbb{T}^d}\phi = 0\}$.

The first bound is obvious, since it says that the restriction of $\widetilde{H}$ to a smaller subspace has smaller norm, but it is recorded here, since it is interesting to observe that some further deductions can be made by using this smaller quantity, instead of the norm of $\widetilde{H}$ on all $L^p(\mathbb{T};X)$. In Corollary 5.7.7 we will show that all three norms above are actually equal.

For the proof of Proposition 5.2.5, we need a simple connection between L^p-norms on $\mathbb{T}$ and on $\mathbb{R}$:

Lemma 5.2.6. *For $f \in L^p(\mathbb{T};X)$ and $\phi \in \mathscr{S}(\mathbb{R})$, we have*

$$\lim_{\varepsilon\downarrow 0} \varepsilon^{1/p}\|f(\cdot)\phi(\varepsilon\cdot)\|_{L^p(\mathbb{R};X)} = \|f\|_{L^p(\mathbb{T};X)}\|\phi\|_{L^p(\mathbb{R};X)}.$$

Proof. The pth power of the left-hand side can be written as

$$\lim_{\varepsilon\downarrow 0}\int_0^1 \|f(x)\|^p\Big(\sum_{k\in\mathbb{Z}}\varepsilon|\phi(\varepsilon(x+k))|^p\Big)\,\mathrm{d}x.$$

For each fixed $x\in\mathbb{R}$, the sum over $k\in\mathbb{Z}$ is a Riemann sum of $|\phi(y)|^p$; these converge to the integral $\int_{\mathbb{R}}|\phi(y)|^p\,\mathrm{d}y$, and one can also check that they remain uniformly bounded in $x\in\mathbb{R}$ and $\varepsilon>0$. Thus the claim follows by dominated convergence. ☐

Proof of Proposition 5.2.5. Complex case. Consider a trigonometric polynomial $f=\sum_{k\in\mathbb{Z}}a_k e_k \in L^p(\mathbb{T};X)$. Viewed as a function on $\mathbb{R}$, we multiply it by $e_{\pm\frac12}\phi$, where $\phi\in\mathscr{S}(\mathbb{R})$ is chosen so that $\widehat{\phi}$ is compactly supported in $(-\frac12,+\frac12)$, and $\|\phi\|_{L^p(\mathbb{R})}=1$. Thus we arrive at the function

$$\phi e_{\pm\frac12}f=\sum_{k\in\mathbb{Z}}a_k e_{k\pm\frac12}\phi,$$

where $\widehat{e_{k\pm\frac12}\phi}(\xi)=\widehat{\phi}(\xi-(k\pm\frac12))$ is supported in a $\frac12$-neighbourhood around $k\pm\frac12$. In particular, $\operatorname{sgn}(\xi)=\operatorname{sgn}(k\pm\frac12)$ for all $\xi\in\operatorname{supp}(\widehat{e_{k\pm\frac12}\phi})$, and hence

$$H(\phi e_{\pm\frac12}f)=\sum_{k\in\mathbb{Z}}a_k H(e_{k\pm\frac12}\phi)=\Big(\sum_{k\in\mathbb{Z}}-i\operatorname{sgn}(k\pm\tfrac12)a_k e_k\Big)e_{\pm\frac12}\phi. \tag{5.26}$$

From Lemma 5.2.6 and (5.26), the latter with $\phi(\varepsilon\cdot)$ in place of ϕ (which is permitted since the derivation of the identity only used that $\widehat{\phi}$ is supported in $(-\frac12,+\frac12)$, and this is also true for $\widehat{\phi(\varepsilon\cdot)}=\varepsilon^{-1}\widehat{\phi}(\varepsilon^{-1}\cdot)$ as long as $0<\varepsilon\leqslant 1$) we deduce that

$$\begin{aligned}&\Big\|\sum_{k\in\mathbb{Z}}-i\operatorname{sgn}(k\pm\tfrac12)a_k e_k\Big\|_{L^p(\mathbb{T};X)}\\ &=\Big\|\sum_{k\in\mathbb{Z}}-i\operatorname{sgn}(k\pm\tfrac12)a_k e_k e_{\pm\frac12}\Big\|_{L^p(\mathbb{T};X)}\|\phi\|_{L^p(\mathbb{R})}\end{aligned}$$

$$
\begin{aligned}
&= \lim_{\varepsilon \downarrow 0} \varepsilon^{1/p} \Big\| \Big(\sum_{k \in \mathbb{Z}} -i \operatorname{sgn}(k \pm \tfrac{1}{2}) a_k e_k \Big) e_{\pm\frac{1}{2}} \phi(\varepsilon \cdot) \Big\|_{L^p(\mathbb{R};X)} \\
&= \lim_{\varepsilon \downarrow 0} \varepsilon^{1/p} \| H(\phi(\varepsilon \cdot) e_{\pm\frac{1}{2}} f) \|_{L^p(\mathbb{R};X)} \\
&\leqslant \lim_{\varepsilon \downarrow 0} \varepsilon^{1/p} \| H \|_{\mathscr{L}(L^p(\mathbb{R};X))} \| \phi(\varepsilon \cdot) e_{\pm\frac{1}{2}} f \|_{L^p(\mathbb{R};X)} \\
&= \| H \|_{\mathscr{L}(L^p(\mathbb{R};X))} \| f \|_{L^p(\mathbb{T};X)}.
\end{aligned}
$$

We note that $\operatorname{sgn}(k \pm \frac{1}{2}) = \operatorname{sgn}(k)$ for all $k \in \mathbb{Z} \setminus \{0\}$, and it is only $\operatorname{sgn}(0 \pm \frac{1}{2}) = \pm 1$ that differs depending on the choice of $\pm\frac{1}{2}$ in the above computation. Taking the average of the two functions $\sum_{k \in \mathbb{Z}} -i \operatorname{sgn}(k \pm \frac{1}{2}) a_k e_k$ and using the convexity of the norm, we finally arrive at

$$
\Big\| \sum_{k \in \mathbb{Z}} -i \operatorname{sgn}(k) a_k e_k \Big\|_{L^p(\mathbb{T};X)} \leqslant \| H \|_{\mathscr{L}(L^p(\mathbb{R};X)} \| f \|_{L^p(\mathbb{T};X)},
$$

which was the claim.

Real case. If X is a real Banach space, we apply the proven complex case to the complexification $X_{\mathbb{C}}^{\gamma,p}$, together with Corollary 2.1.13, to find that

$$
\| \widetilde{H} \|_{\mathscr{L}(L^p(\mathbb{T};X))} = \| \widetilde{H} \|_{\mathscr{L}(L^p(\mathbb{T};X_{\mathbb{C}}^{\gamma,p}))} \leqslant \| H \|_{\mathscr{L}(L^p(\mathbb{R};X_{\mathbb{C}}^{\gamma,p}))} = \| H \|_{\mathscr{L}(L^p(\mathbb{R};X))}.
$$

□

Convergence of Fourier series

Equipped with the periodic Hilbert transform, we now return to the question of convergence of Fourier series. On $L^p(\mathbb{T};X)$ let

$$
P_n f = \sum_{|k| \leqslant n} \widehat{f}(k) e_k, \quad n \geqslant 0,
$$

where $\widehat{f}(k)$ are the Fourier coefficients of f. Since $\|\widehat{f}(k)\| \leqslant \|f\|_1 \leqslant \|f\|_p$, we find that each P_n is well defined with norm at most $2n+1$ on $L^p(\mathbb{T};X)$ for any Banach space X. The operator P_0 is just the projection onto constant functions in $L^p(\mathbb{T};X)$.

Due to the one-to-one correspondence of the pairs of trigonometric functions $\{e_k, e_{-k}\}$ and $\{g_k, g_{-k}\}$ for every $k \in \mathbb{Z}_+$ (and $e_0 = g_0$ for $k = 0$), the reader can routinely verify that P_n can also be expressed in terms of the real functions g_k and, by the resulting expression, maps real-valued functions to real-valued ones. (For the last assertion it is important that we consider the symmetric partial sums over $|k| \leqslant n$.) Thus these operators, too, are meaningful for all Banach spaces.

Proposition 5.2.7. *For a Banach space X and a fixed $p \in (1,\infty)$, the following are equivalent:*

(1) *the periodic Hilbert transform* $\widetilde{H}$ *is bounded on* $L^p(\mathbb{T};X)$*;*
(2) *the partial sum projections* $(P_n)_{n\geqslant 1}$ *are uniformly bounded on* $L^p(\mathbb{T};X)$*;*
(3) *For all* $f \in L^p(\mathbb{T};X)$ *one has* $P_n f \to f$ *in* $L^p(\mathbb{T};X)$*.*

In this case the following estimate holds:

$$\frac{1}{2}\|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};X))} - 1 \leqslant \sup_{n\geqslant 1} \|P_n\|_{\mathscr{L}(L^p(\mathbb{T};X))} \leqslant \|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};X))} + 1. \tag{5.27}$$

Since we will show later on that the boundedness of the (periodic) Hilbert transform is equivalent to the UMD property, it follows that all of the above the assertions are equivalent to X being a UMD space.

Proof. Complex case. For $a \in \mathbb{Z}$ let M_a be the operator on $L^p(\mathbb{T};X)$ given by $M_a f := e_a f$, and for $a \leqslant b \in \mathbb{Z} \cup \{\infty\}$ let $\widetilde{\Delta}_{[a,b)}$ be the partial sum projection on $L^p(\mathbb{T};X)$ given by

$$\widetilde{\Delta}_{[a,b)} f := \sum_{a\leqslant k<b} \widehat{f}(k) e_k.$$

Let us derive some identities, first for trigonometric polynomials only, so that there are no issues of convergence or well-definedness of the operators in the first place. Recall that

$$i\widetilde{H}f = \sum_{k\in\mathbb{Z}} \operatorname{sgn}(k)\widehat{f}(k)e_k = \sum_{k\geqslant 1} \widehat{f}(k)e_k - \sum_{k\leqslant -1} \widehat{f}(k)e_k, \qquad P_0 f = \widehat{f}(k)e_0,$$

so that

$$\begin{aligned} \frac{1}{2}(I + P_0 + i\widetilde{H})f &= \sum_{k\geqslant 0} \widehat{f}(k)e_k = \widetilde{\Delta}_{[0,\infty)} f, \\ M_a \widetilde{\Delta}_{[0,\infty)} M_{-a} f &= \sum_{k\geqslant a} \widehat{f}(k)e_k = \widetilde{\Delta}_{[a,\infty)} f, \end{aligned}$$

and finally

$$\begin{aligned} \widetilde{\Delta}_{[a,b)} &= \widetilde{\Delta}_{[a,\infty)} - \widetilde{\Delta}_{[b,\infty)} \\ &= \frac{1}{2}(M_a P_0 M_{-a} - M_b P_0 M_{-b}) + \frac{i}{2}(M_a \widetilde{H} M_{-a} + M_b \widetilde{H} M_{-b}). \end{aligned} \tag{5.28}$$

Solving for $\widetilde{H}$ instead, we get

$$\widetilde{H} = -i(2\widetilde{\Delta}_{[0,\infty)} - I - P_0). \tag{5.29}$$

(1)⇒(2): From (5.28) it follows that

$$\|P_n\|_{\mathscr{L}(L^p(\mathbb{T};X))} = \|\widetilde{\Delta}_{[-n,n+1)}\|_{\mathscr{L}(L^p(\mathbb{T};X))} \leqslant 1 + \|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};X))}.$$

(2)⇒(1): For a trigonometric polynomial $f = \sum_{|k|\leqslant m} a_k e_k$, it is clear that $\widetilde{\Delta}_{[0,\infty)} f = \widetilde{\Delta}_{[0,n)} f$ for any $n > m$, and therefore, from (5.29), we have

$$\widetilde{H} f = -i(2\widetilde{\Delta}_{[0,2m+1)} - I - P_0) f = -i(2M_m \widetilde{\Delta}_{[-m,m+1)} M_{-m} - I - P_0) f$$

and

$$\begin{aligned}\|\widetilde{H} f\|_{L^p(\mathbb{T};X)} &\leqslant (2\|P_m\|_{\mathscr{L}(L^p(\mathbb{T};X))} + 2)\|f\|_{L^p(\mathbb{T};X)} \\ &\leqslant 2\Big(\sup_{n\geqslant 1} \|P_n\|_{\mathscr{L}(L^p(\mathbb{T};X))} + 1\Big)\|f\|_{L^p(\mathbb{T};X)}.\end{aligned}$$

(2)⇒(3): For trigonometric polynomials $f : \mathbb{T} \to X$, the convergence is obvious. For general $f \in L^p(\mathbb{T}; X)$, the result follows by density and the uniform boundedness of $(P_n)_{n\geqslant 1}$.

(3)⇒(2): This follows from the uniform boundedness principle.

Real case. The norm comparison (5.27) is immediate from the already proven complex case applied to the complexification $X_{\mathbb{C}}^{\gamma,p}$ of X and Corollary 2.1.13, which shows that

$$\|T\|_{\mathscr{L}(L^p(\mathbb{T};X))} = \|T\|_{\mathscr{L}(L^p(\mathbb{T};X_{\mathbb{C}}^{\gamma,p}))}, \qquad T \in \widetilde{H}, P_n.$$

This proves that (1)⇔(2), and the proof of (2)⇔(3) follows the complex case *verbatim*. □

5.2.c Necessity of the UMD condition

We now turn to the converse implication of Theorem 5.1.1, the deduction of the UMD condition from the Hilbert transform boundedness. The method of proof is interesting in its own right. It can be generalised to a number of other situations where the defining condition of UMD is derived from the boundedness of some operators, and indeed it is essentially the only known method applicable for this purpose.

The key intermediate estimate is the following lifting of the boundedness of the Hilbert transform from individual functions to certain sequences of functions, indeed, martingale difference sequences of a suitable form on the product space $\mathbb{T}^n$. For $k \leqslant n$, let $\mathbb{T}_k$ denote the kth copy of $\mathbb{T}$ in the product $\mathbb{T}^n = \mathbb{T} \times \cdots \times \mathbb{T} = \mathbb{T}_1 \times \cdots \times \mathbb{T}_n$, and $\mathbb{T}^k := \mathbb{T}_1 \times \cdots \times \mathbb{T}_k$ be the product of the first k copies. These conventions are relevant when we identify functions of $k \leqslant n$ variables with functions of n variables, constant with respect to the remaining ones.

Proposition 5.2.8. *Let X be a complex Banach space, let $p \in (1,\infty)$, and suppose that $\widetilde{H}$ is bounded on $L^p_0(\mathbb{T}; X)$. Then for all signs $\epsilon_1, \dots, \epsilon_n \in \{-1, 1\}$ and all $f_k \in L^p_0(\mathbb{T}_k; L^p(\mathbb{T}^{k-1}; X))$, $k = 1, \dots, n$, we have*

$$\Big\| \sum_{k=1}^{n} \epsilon_k \widetilde{H}^{(k)} f_k \Big\|_{L^p(\mathbb{T}^n;X)} \leqslant \|\widetilde{H}\|_{\mathscr{L}(L^p_0(\mathbb{T};X))} \Big\| \sum_{k=1}^{n} f_k \Big\|_{L^p(\mathbb{T}^n;X)},$$

where $L^p(\mathbb{T}^0;X) := X$ *and* $\widetilde{H}^{(k)} = I \otimes \cdots \otimes I \otimes \widetilde{H}$ *is the tensor extension of* $\widetilde{H}$ *acting in the* k*th coordinate; we make the natural identifications* $L^p_0(\mathbb{T}_k; L^p(\mathbb{T}^{k-1};X)) \subseteq L^p(\mathbb{T}^k;X) \subseteq L^p(\mathbb{T}^n;X)$.

This, in turn, is deduced with the help of a transformation that introduces a new auxiliary variable.

Lemma 5.2.9. *For a sequence* $N = (N_k)_{k=1}^n$ *of non-zero integers, let* $\mathscr{B}$ *denote the following transformation, lifting functions on* $\mathbb{T}^n$ *to functions on* $\mathbb{T}^{n+1}$*, where we index the new coordinate as the zeroth:*

$$\mathscr{B}g(s, t_1 \dots, t_n) := g(t + Ns) = g(t_1 + N_1 s, \dots, t_n + N_n s).$$

(1) *For all* $p \in [1,\infty)$, $\mathscr{B}$ *is an isometry from* $L^p(\mathbb{T}^n;X)$ *into* $L^p(\mathbb{T}^{n+1};X)$ *which maps trigonometric polynomials into trigonometric polynomials.*
(2) *For any given sequence of trigonometric polynomials*

$$f_k \in L^p_0(\mathbb{T}_k; L^p(\mathbb{T}^{k-1};X)) \subseteq L^p(\mathbb{T}^n;X),$$

if $|N_1|, \dots, |N_n|$ *is sufficiently rapidly increasing, then we have* $\mathscr{B}f_k \in L^p_0(\mathbb{T}_0; L^p(\mathbb{T}^k;X))$ *and*

$$\widetilde{H}^{(0)} \mathscr{B} f_k = \operatorname{sgn}(N_k) \mathscr{B} \widetilde{H}^{(k)} f_k.$$

Proof. The isometry is an immediate computation

$$\begin{aligned}\|\mathscr{B}g\|_{L^p(\mathbb{T}^{n+1};X)} &= \Big(\int_{\mathbb{T}} \int_{\mathbb{T}^n} \|g(t_1 + N_1 s, \dots, t_n + N_n s)\|^p \,\mathrm{d}t \,\mathrm{d}s \Big)^{1/p} \\ &= \Big(\int_{\mathbb{T}} \int_{\mathbb{T}^n} \|g(t_1, \dots, t_n)\|^p \,\mathrm{d}t \,\mathrm{d}s \Big)^{1/p} = \|g\|_{L^p(\mathbb{T}^n;X)},\end{aligned}$$

as the dependence on s disappears by a change of variables and the translation-invariance of the Haar measure on $\mathbb{T}$.

We turn to the identity in (2). We first consider the complex case of a complex Banach space X. Let

$$f_k(t) = \sum_{\substack{m \in \mathbb{Z}^k \\ m_k \neq 0}} \widehat{f}_k(m) e_m(t) = \sum_{\substack{m \in \mathbb{Z}^k \\ m_k \neq 0}} \widehat{f}_k(m) e_{(m_1,\dots,m_{k-1})}(t_1, \dots, t_{k-1}) e_{m_k}(t_k),$$

where the restriction $m_k \neq 0$ comes from $f_k \in L^p_0(\mathbb{T}_k; L^p(\mathbb{T}^{k-1};X))$, we observe that

$$\widetilde{H}^{(k)} f_k(t) = -i \sum_{\substack{m \in \mathbb{Z}^k \\ m_k \neq 0}} \widehat{f}_k(m) \operatorname{sgn}(m_k) e_m(t),$$

$$\mathscr{B}\widetilde{H}^{(k)} f_k(s,t) = -i \sum_{\substack{m\in\mathbb{Z}^k \\ m_k\neq 0}} \widehat{f}_k(m) \operatorname{sgn}(m_k) e_m(t+Ns). \tag{5.30}$$

On the other hand,

$$\mathscr{B} f_k(s,t) = f_k(t+Ns) = \sum_{\substack{m\in\mathbb{Z}^k \\ m_k\neq 0}} \widehat{f}_k(m) e_m(t) e_{m\cdot N}(s),$$

$$\widetilde{H}^{(0)}\mathscr{B} f_k(s,t) = -i \sum_{\substack{m\in\mathbb{Z}^k \\ m_k\neq 0}} \widehat{f}_k(m) \operatorname{sgn}(m\cdot N) e_m(t+Ns). \tag{5.31}$$

The expressions (5.30) and (5.31) differ in the factors $\operatorname{sgn}(m_k)$ and $\operatorname{sgn}(m\cdot N)$ only.

Now $m\cdot N = m_1N_1 + \ldots + m_{k-1}N_{k-1} + m_kN_k$, and we would like that the sign is determined by the last term only. Since $|m_k| \geqslant 1$, this is ensured if

$$|N_k| > A(|N_1| + \cdots + |N_{k-1}|), \quad A := \max\{|m|_\infty : \ \widehat{f}_k(m)\neq 0,\ k=1,\ldots,n\},$$

since then

$$|m_1N_1 + \ldots + m_{k-1}N_{k-1}| \leqslant A(|N_1| + \ldots + |N_{k-1}|) < |N_k| \leqslant |N_k m_k|,$$

and thus $\operatorname{sgn}(m\cdot N) = \operatorname{sgn}(m_kN_k) = \operatorname{sgn}(m_k)\operatorname{sgn}(N_k)$. Substituting this into (5.31), we obtain the identity asserted in claim (2).

If X is a real Banach space, we simply view it as a subset of its complexification $X_{\mathbb{C}}$. The identities in (2) that are true for $X_{\mathbb{C}}$-valued trigonometric polynomials must in particular hold for X-valued ones, noting that $\mathscr{B}$, $\widetilde{H}^{(0)}$ and $\widetilde{H}^{(k)}$ are all real operators that send real-valued functions to real-valued ones. (In this algebraic verification, we do not even need to worry about the choice of a particular norm on $X_{\mathbb{C}}$.) $\square$

Proof of Proposition 5.2.8. It follows from Fubini's theorem that $\widetilde{H}^{(k)}$ is continuous on the space $L^p_0(\mathbb{T}; L^p(\mathbb{T}^{k-1}; X))$ and satisfies

$$\|\widetilde{H}^{(k)}\|_{\mathscr{L}(L^p_0(\mathbb{T};L^p(\mathbb{T}^{k-1};X)))} = \|\widetilde{H}\|_{\mathscr{L}(L^p_0(\mathbb{T};X))}.$$

By the density of trigonometric polynomials in $L^p_0(\mathbb{T}; Z)$ and in $L^p(\mathbb{T}^{k-1}; X))$, they are also dense in $L^p_0(\mathbb{T}; L^p(\mathbb{T}^{k-1}; X))$: we first approximate a function in this space by a trigonometric polynomial with $L^p(\mathbb{T}^{k-1}; X)$-coefficients, and then approximate each of these coefficients by a trigonometric polynomial with X-coefficients. So it suffices to prove the claim under the assumption that the f_k, and then also $\widetilde{H}^{(k)} f_k$, are trigonometric polynomials in the respective spaces.

We then pick non-zero integers $N_1,\ldots,N_n$, such that $|N_1|,\ldots,|N_n|$ increase sufficiently rapidly to apply Lemma 5.2.9, and $\operatorname{sgn}(N_k) = \epsilon_k$. Then

$$\begin{aligned}\Big\|\sum_{k=1}^{n}\epsilon_k\widetilde{H}^{(k)}f_k\Big\|_{L^p(\mathbb{T}^n;X)} &= \Big\|\mathscr{B}\sum_{k=1}^{n}\epsilon_k\widetilde{H}^{(k)}f_k\Big\|_{L_0^p(\mathbb{T};L^p(T^n;X))}\\ &= \Big\|\widetilde{H}^{(0)}\mathscr{B}\sum_{k=1}^{n}f_k\Big\|_{L_0^p(\mathbb{T};L^p(\mathbb{T}^n;X))}\\ &\leqslant \|\widetilde{H}^{(0)}\|\Big\|\mathscr{B}\sum_{k=1}^{n}f_k\Big\|_{L_0^p(\mathbb{T};L^p(\mathbb{T}^n;X))}\\ &= \|\widetilde{H}\|\Big\|\sum_{k=1}^{n}f_k\Big\|_{L^p(\mathbb{T}^n;X)}.\end{aligned}$$

□

The deduction of the UMD property from the Hilbert transform boundedness is now a simple corollary to Proposition 5.2.8.

Theorem 5.2.10 (Bourgain). *Let X be a Banach space, $p \in (1,\infty)$, and let the Hilbert transform act boundedly on $L_0^p(\mathbb{T};X)$. Then X is a UMD space, and more precisely*

$$\beta_{p,X}^{\mathbb{R}} \leqslant \|\widetilde{H}\|^2_{L_0^p(\mathbb{T};X)}.$$

Note that the quantitative estimate is in terms of the real UMD constant $\beta_{p,X}^{\mathbb{R}}$, even when X is a complex Banach space. Namely, the signs $\epsilon_k \in \{-1,1\}$ implicit in the assertion are ultimately produced by the eigenvalues ± 1 of the Hilbert transform, whereas there is no equally direct way of relating the Hilbert transform to more general complex ϵ_k of modulus one.

Proof. By Theorem 4.2.5 it suffices to estimate the dyadic UMD constant. In order to most conveniently connect this with Fourier analysis, we choose a model of the Rademacher system $(r_k)_{k=1}^n$, where the probability space is $\mathbb{T}^n$, and $r_k = r_k(t_k)$ is a function of the kth coordinate only. Then it is (more than) sufficient to prove that

$$\Big\|\sum_{k=1}^{n}\epsilon_k f_k\Big\|_{L^p(\mathbb{T}^n;X)} \leqslant \|\widetilde{H}\|^2_{L_0^p(\mathbb{T};X)}\Big\|\sum_{k=1}^{n}f_k\Big\|_{L^p(\mathbb{T}^n;X)} \tag{5.32}$$

for all $f_k \in L_0^p(\mathbb{T}_k;L^p(\mathbb{T}^{k-1};X))$, since in particular each $\phi_k(r_1,\dots,r_{k-1})r_k$ is of this form, and the latter are precisely the martingale differences of Paley–Walsh martingales (see Proposition 3.1.10).

Now (5.32) follows from two (as suggested by the power of $\|\widetilde{H}\|_{L_0^p(\mathbb{T};X)}$) applications of Proposition 5.2.8, first with the same signs ϵ_k as in (5.32), and then with all ϵ_k set to 1. Namely, observing that $\widetilde{H}^{(k)}\circ\widetilde{H}^{(k)} = -I$ on $L_0^p(\mathbb{T}_k;L^p(\mathbb{T}^{k-1};X))$, we have

$$\Big\|\sum_{k=1}^{n}\epsilon_k f_k\Big\|_{L^p(\mathbb{T}^n;X)} = \Big\|\sum_{k=1}^{n}\epsilon_k\widetilde{H}^{(k)}(\widetilde{H}^{(k)}f_k)\Big\|_{L^p(\mathbb{T}^n;X)}$$

$$
\begin{aligned}
&\leqslant \|\widetilde{H}\|_{\mathscr{L}(L^p_0(\mathbb{T};X))}\Big\| \sum_{k=1}^n \widetilde{H}^{(k)} f_k \Big\|_{L^p(\mathbb{T}^n;X)} \\
&\leqslant \|\widetilde{H}\|^2_{\mathscr{L}(L^p_0(\mathbb{T};X))}\Big\| \sum_{k=1}^n f_k \Big\|_{L^p(\mathbb{T}^n;X)}.
\end{aligned}
$$

□

Corollary 5.2.11. *Let X be a Banach space and let $p \in (1,\infty)$. Then*

$$
\begin{aligned}
(\beta^{\mathbb{R}}_{p,X})^{1/2} \leqslant \|\widetilde{H}\|_{\mathscr{L}(L^p_0(\mathbb{T};X))} &\leqslant \|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};X))} \\
&\leqslant \|H\|_{\mathscr{L}(L^p(\mathbb{R};X))} \leqslant 2(\beta^{\mathbb{R}}_{p,X})^2,
\end{aligned}
\tag{5.33}
$$

and in particular X is a UMD space if and only if any of the three Hilbert transform norms is finite.

Recalling from Proposition 4.2.10 that $\beta^{\mathbb{R}}_{p,X} \leqslant \beta_{p,X} \leqslant \pi/2 \cdot \beta^{\mathbb{R}}_{p,X}$, we also obtain comparisons between the constants in (5.33) and $\beta_{p,X}$.

Proof. The estimate (5.33) is a combination of Theorem 5.1.13, Proposition 5.2.5 and Theorem 5.2.10, except that the rightmost bound coming from Theorem 5.1.13 was there stated with $\beta_{p,X}$ instead of $\beta^{\mathbb{R}}_{p,X}$, where the latter is defined using martingale transforms with signs in $\{-1,1\}$ only, rather than scalars from $\{z \in \mathbb{K} : |z| = 1\}$. However, applying this mentioned Theorem 5.1.13 to the real Banach space $X_{\mathbb{R}}$ in place of X (the same space, where we simply 'forget' about the complex multiplication), we have $\|H\|_{\mathscr{L}(L^p(\mathbb{R};X_{\mathbb{R}}))} \leqslant 2\beta^2_{p,X_{\mathbb{R}}}$, and it remains to observe that $\|H\|_{\mathscr{L}(L^p(\mathbb{R};X_{\mathbb{R}}))} = \|H\|_{\mathscr{L}(L^p(\mathbb{R};X))}$ and $\beta_{p,X_{\mathbb{R}}} = \beta^{\mathbb{R}}_{p,X}$, since X and $X_{\mathbb{R}}$ have the same norm. □

Later in this chapter, as a particular instance of a more general theorem concerning periodic and non-periodic operator norms, we find that all three quantities in the middle of (5.33) are actually equal. Also, the factor 2 on the right of (5.33) was an artefact of our discretisation scheme, and can be avoided by other methods that we indicate in the Notes. Aside from these remarks, (5.33) displays the best quantitative information currently known on the relation of the different quantities. In particular, it is an open problem whether either of the first or the last estimates in (5.33) could be improved to a linear bound between the respective quantities.

5.3 Fourier multipliers

The representation of the Hilbert transform as a multiplication by $-i\,\mathrm{sgn}(\xi)$ "on the Fourier transform side" leads to a natural generalisation. We work intensively with the Fourier transform, and therefore stay exclusively in the realm of complex Banach spaces for the core developments of this theory.

A reader interested in real space will, however, find in Proposition 5.3.11 a condition under which a class of multipliers have a meaningful interpretation in such spaces as well.

Recall that $\check{L}^1(\mathbb{R}^d;X)$ is the space of inverse Fourier transforms of functions in $L^1(\mathbb{R}^d;X)$.

Definition 5.3.1. *Given $m \in L^\infty(\mathbb{R}^d;\mathscr{L}(X,Y))$, its associated* Fourier multiplier operator *is*

$$T_m : \check{L}^1(\mathbb{R}^d;X) \to \check{L}^1(\mathbb{R}^d;Y),\ f \mapsto (m\widehat{f})\check{}.$$

The function m is called the multiplier *or the* symbol *of this operator.*

To see that this definition makes sense, it is enough to observe that the pointwise multiplication $g \mapsto mg$ by $m \in L^\infty(\mathbb{R}^d;\mathscr{L}(X,Y))$ maps $L^1(\mathbb{R}^d;X)$ into $L^1(\mathbb{R}^d;Y)$. An important special case appears when $X = Y$ and $m \in L^\infty(\mathbb{R}^d)$.

In this section we begin a systematic study of Fourier multiplier operators. For later reference, the basic theory will be set up over $\mathbb{R}^d$, as this presents rarely any complications compared to the line $\mathbb{R}$. However, when it comes to some deeper results, such as Mihlin's multiplier theorem, we specialise to the one-dimensional setting. The multiplier theory in several variables is taken up again in Section 5.5.

5.3.a General theory

Here we survey some relatively direct consequences of the definition of Fourier multipliers.

The function $T_m f$, as defined in Definition 5.3.1 with $f \in \check{L}^1(\mathbb{R}^d;X)$ and $m \in L^\infty(\mathbb{R}^d;\mathscr{L}(X,Y))$, is continuous (in fact, an element of $C_0(\mathbb{R}^d;Y)$) and is explicitly given by the Bochner integral

$$T_m f(x) = \int_{\mathbb{R}^d} m(\xi)\widehat{f}(\xi)e^{2\pi i\xi\cdot x}\,\mathrm{d}\xi.$$

Also, T_m is bounded from $\check{L}^1(\mathbb{R}^d;X)$ to $\check{L}^1(\mathbb{R}^d;Y)$: this follows from the factorisation

$$T_m : \check{L}^1(\mathbb{R}^d;X) \xrightarrow{\ \widehat{}\ } L^1(\mathbb{R}^d;X) \xrightarrow{m\cdot} L^1(\mathbb{R}^d;Y) \xrightarrow{\ \check{}\ } \check{L}^1(\mathbb{R}^d;Y),$$

where the two Fourier transforms are isometric and the multiplication with an $L^\infty(\mathbb{R}^d;\mathscr{L}(X,Y))$-function maps $L^1(\mathbb{R}^d;X)$ into $L^1(\mathbb{R}^d;Y)$.

Lemma 5.3.2 (Algebraic properties). *Given $m_1, m_2 \in L^\infty_{\mathrm{so}}(\mathbb{R}^d;\mathscr{L}(X,Y))$ and $m_3 \in L^\infty_{\mathrm{so}}(\mathbb{R}^d;\mathscr{L}(Y,Z))$, the following operator identities are valid:*

$$\begin{aligned} T_{m_1+m_2} &= T_{m_1} + T_{m_2} \quad \text{in} \quad \mathscr{L}(\check{L}^1(\mathbb{R}^d;X),\check{L}^1(\mathbb{R}^d;Y)),\\ T_{m_3 m_2} &= T_{m_3}T_{m_2} \quad \text{in} \quad \mathscr{L}(\check{L}^1(\mathbb{R}^d;X),\check{L}^1(\mathbb{R}^d;Z)). \end{aligned}$$

Proof. Linearity is clear, and multiplicativity follows from the fact that $(\check{g})\widehat{\ } = g$ for $g = m_2\widehat{f} \in L^1(\mathbb{R}^d; Y)$. $\square$

The principal issue of interest about multiplier operators is whether they are bounded from $L^p(\mathbb{R}^d; X)$ into $L^p(\mathbb{R}^d; Y)$; more precisely, whether a constant $C \geqslant 0$ exists such that

$$\|T_m f\|_p \leqslant C\|f\|_p \qquad \forall f \in L^p(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}^d; X).$$

If this is the case for some $1 \leqslant p < \infty$, then T_m extends uniquely to $L^p(\mathbb{R}^d; X)$, by the density of $L^p(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}^d; X)$ in $L^p(\mathbb{R}^d; X)$ (Lemma 2.4.7). We then call m a *Fourier multiplier* from $L^p(\mathbb{R}^d; X)$ to $L^p(\mathbb{R}^d; Y)$.

Sometimes it is convenient to the check the *a priori* estimate

$$\|T_m f\|_p \leqslant C\|f\|_p$$

on some subspace $E \subseteq L^p(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}^d; X)$ first. As long as E is dense in $L^p(\mathbb{R}^d; X)$, this yields the existence and uniqueness of a bounded extension $\bar{T}_m$ to $L^p(\mathbb{R}^d; X)$. On the other hand, we always have a direct definition of $T_m f$ for all $f \in L^p(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}^d; X)$, and one may wonder whether this coincides with the abstract extension $\bar{T}_m f$, when both make sense. Since T_m is continuous on $\check{L}^1(\mathbb{R}^d; X)$ and $\bar{T}_m$ on $L^p(\mathbb{R}^d; X)$, the answer is affirmative provided that E is dense in $L^p(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}^d; X)$. Proposition 2.4.23 settles this matter for the most common choices

$$E = \mathscr{S}(\mathbb{R}^d; X), \quad E = \mathscr{D}(\mathbb{R}^d; X), \quad E = \check{\mathscr{D}}(\mathbb{R}^d \setminus \{0\}; X).$$

Definition 5.3.3 (Multipliers from $L^p(\mathbb{R}^d; X)$ into $L^p(\mathbb{R}^d; Y)$). *The space of all $m \in L^\infty(\mathbb{R}^d; \mathscr{L}(X, Y))$ for which T_m has a bounded extension from $L^p(\mathbb{R}^d; X)$ to $L^p(\mathbb{R}^d; Y)$ will be denoted by $\mathfrak{M}L^p(\mathbb{R}^d; X, Y)$.*

By $\mathfrak{M}L^p(\mathbb{R}^d; X) \subseteq \mathfrak{M}L^p(\mathbb{R}^d; X, X)$ we denote the subspace of scalar-valued multipliers $m \in L^\infty(\mathbb{R}^d)$ *of $L^p(\mathbb{R}^d; X)$, and by $\mathfrak{M}L^p(\mathbb{R}^d) := \mathfrak{M}L^p(\mathbb{R}^d; \mathbb{C})$ the* multipliers of the scalar-valued space $L^p(\mathbb{R}^d)$.

Endowed with the norm

$$\|m\|_{\mathfrak{M}L^p(\mathbb{R}^d; X, Y)} := \|T_m\|_{\mathscr{L}(L^p(\mathbb{R}^d; X), L^p(\mathbb{R}^d; Y))},$$

$\mathfrak{M}L^p(\mathbb{R}^d; X, Y)$ is a normed vector space. We shall see shortly (in Proposition 5.3.16) that it is actually a Banach space.

Example 5.3.4. Simple interrelations of the multiplier classes include:

(1) By Plancherel's theorem

$$\mathfrak{M}L^2(\mathbb{R}^d) = L^\infty(\mathbb{R}^d), \qquad \|m\|_{\mathfrak{M}L^2(\mathbb{R}^d)} = \|m\|_{L^\infty(\mathbb{R}^d)}.$$

(2) For all $p \in [1,\infty)$,

$$\mathfrak{M}L^p(\mathbb{R}^d;X) \subseteq \mathfrak{M}L^p(\mathbb{R}^d), \quad \|m\|_{\mathfrak{M}L^p(\mathbb{R}^d)} \leqslant \|m\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)}. \tag{5.34}$$

This follows by considering one-dimensional subspaces of X.

(3) If H is a Hilbert space and $p \in [1,\infty)$, then by Theorem 2.1.9

$$\mathfrak{M}L^p(\mathbb{R}^d;H) = \mathfrak{M}L^p(\mathbb{R}^d).$$

(4) If X is a Banach space and $p \in [1,\infty)$, then by Fubini's theorem

$$\mathfrak{M}L^p(\mathbb{R}^d;L^p(S;X)) = \mathfrak{M}L^p(\mathbb{R}^d;X).$$

Proposition 5.3.5. *If μ is a complex Borel measure on $\mathbb{R}^d$, its Fourier transform*

$$\widehat{\mu}(\xi) := \int_{\mathbb{R}^d} e^{-2\pi i x\cdot\xi}\,\mathrm{d}\mu(x), \quad \xi \in \mathbb{R}^d,$$

belongs to $\mathfrak{M}L^p(\mathbb{R}^d;X)$ for every Banach space X and $p \in [1,\infty)$, and

$$\|\widehat{\mu}\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)} \leqslant \|\mu\|,$$

where $\|\mu\|$ is the total variation of μ.

Proof. Defining, for $f \in L^p(\mathbb{R}^d;X)$,

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

we have *Young's inequality*

$$\begin{aligned}
\|\mu * f\|_p^p &\leqslant \|\mu\|^p \int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d} |f(x-y)|\,\frac{\mathrm{d}|\mu|(y)}{\|\mu\|}\Big)^p\,\mathrm{d}x \\
&\leqslant \|\mu\|^p \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(x-y)|^p\,\frac{\mathrm{d}|\mu|(y)}{\|\mu\|}\,\mathrm{d}x \\
&= \|\mu\|^p \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(x-y)|^p\,\mathrm{d}x\,\frac{\mathrm{d}|\mu|(y)}{\|\mu\|} = \|\mu\|^p\|f\|_p^p.
\end{aligned}$$

Proceeding as in Lemma 2.4.8, for $f \in L^p(\mathbb{R}^d;X) \cap \check{L}^1(\mathbb{R}^d;X)$ we obtain $(\mu * f)\widehat{\ } = \widehat{\mu}\widehat{f}$. Hence,

$$T_{\widehat{\mu}} f = (\widehat{\mu}\widehat{f})\check{\ } = \mu * f.$$

Recalling from Lemma 2.4.7 that $L^p(\mathbb{R}^d;X)\cap\check{L}^1(\mathbb{R}^d;X)$ is dense in $L^p(\mathbb{R}^d;X)$ we infer that $\widehat{\mu} \in \mathfrak{M}L^p(\mathbb{R}^d;X)$ and

$$\|\widehat{\mu}\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)} \leqslant \|\mu\|.$$

□

Remark 5.3.6 (The multiplier space $\mathfrak{M}L^1(\mathbb{R}^d;X)$). In the scalar case, it is known that all Fourier multipliers of $L^1(\mathbb{R}^d)$ arise from Fourier transforms of complex Borel measures, as in the previous result, and $\|T_{\widehat{\mu}}\|_{\mathfrak{M}L^1(\mathbb{R}^d)} = \|\mu\|$. In combination with (5.34) we then find

$$\|\mu\| = \|T_{\widehat{\mu}}\|_{\mathfrak{M}L^1(\mathbb{R}^d)} \leqslant \|T_{\widehat{\mu}}\|_{\mathfrak{M}L^1(\mathbb{R}^d;X)} \leqslant \|\mu\|,$$

which shows that

$$\mathfrak{M}L^1(\mathbb{R}^d;X) = \mathfrak{M}L^1(\mathbb{R}^d) = \widehat{M}(\mathbb{R}^d)$$

isometrically, for any Banach space X, where $\widehat{M}(\mathbb{R}^d)$ is the Banach space of Fourier transforms of complex Borel measures, equipped with the total variation norm of the corresponding measure.

By the previous remark we conclude that the L^1-theory of Fourier multipliers is relatively uninteresting (except in the operator-valued setting), and our primary concern in the following is the case $p \in (1,\infty)$.

From the fact that the reflection $f \mapsto \widetilde{f}$,

$$\widetilde{f}(x) := f(-x),$$

commutes with the Fourier transform we infer that $m \in \mathfrak{M}L^p(\mathbb{R}^d;X)$ if and only if $\widetilde{m} \in \mathfrak{M}L^p(\mathbb{R}^d;X)$ and

$$T_{\widetilde{m}} f = (T_m \widetilde{f})^{\sim}, \qquad \|\widetilde{m}\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)} = \|m\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)}.$$

Proposition 5.3.7. *For any Banach spaces X and Y and all $p \in (1,\infty)$, if $m \in \mathfrak{M}L^p(\mathbb{R}^d;X,Y)$, then the pointwise adjoint $m^* \in L^\infty(\mathbb{R}^d;\mathscr{L}(Y^*,X^*))$ satisfies $m^* \in \mathfrak{M}L^{p'}(\mathbb{R}^d;Y^*,X^*)$ and*

$$(T_m)^*|_{L^{p'}(\mathbb{R}^d;X^*)} = T_{\widetilde{m}^*} \in \mathscr{L}(L^{p'}(\mathbb{R}^d;X^*), L^{p'}(\mathbb{R}^d;Y^*)).$$

As a consequence,

$$\mathfrak{M}L^p(\mathbb{R}^d;X) = \mathfrak{M}L^{p'}(\mathbb{R}^d;X^*).$$

isometrically.

Proof. For $f \in \mathscr{S}(\mathbb{R}^d;X)$ and $g \in \mathscr{S}(\mathbb{R}^d;Y^*)$ we have

$$\langle g, T_m f\rangle = \langle \breve{g}, m\widehat{f}\rangle = \langle m\breve{g}, \widehat{f}\rangle = \langle \widetilde{m}\widehat{g}, \breve{f}\rangle = \langle T_{\widetilde{m}} g, f\rangle.$$

Since $\mathscr{S}(\mathbb{R}^d;X)$ and $\mathscr{S}(\mathbb{R}^d;Y^*)$ are dense in $L^p(\mathbb{R}^d;X)$ and $L^{p'}(\mathbb{R}^d;Y^*)$, respectively, the first assertion follows from Proposition 1.3.1.

For $X = Y$ and scalar-valued multipliers $m \in \mathfrak{M}L^p(\mathbb{R}^d;X)$, we also get the contractive inclusion $\mathfrak{M}L^p(\mathbb{R}^d;X) \subseteq \mathfrak{M}L^{p'}(\mathbb{R}^d;X^*)$. The reverse inclusion follows by applying the inclusion to p' and X^*, and noting that $\mathfrak{M}L^p(\mathbb{R}^d;X^{**}) \subseteq \mathfrak{M}L^p(\mathbb{R}^d;X)$ by restriction. $\square$

In the following we collect some basic invariance results for multiplier spaces.

Proposition 5.3.8 (Invariances of the space $\mathfrak{M}L^p(\mathbb{R}^d;X,Y)$). *Let $m \in \mathfrak{M}L^p(\mathbb{R}^d;X,Y)$ and $a \in \mathbb{R}^d$ be given, and let A be an invertible $(d \times d)$ matrix with real entries satisfying $\det A = 1$. Then $m(\cdot - a)$, $e_a m$ and $m \circ A$ belong to $\mathfrak{M}L^p(\mathbb{R}^d;X,Y)$ with multiplier norms equal to that of m. The corresponding multiplier operators are given by*

$$\begin{aligned} T_{m(\cdot-a)} &= M_a T_m M_{-a}, && \text{with } M_a f(x) := e_a(x) f(x) = e^{2\pi i a x} f(x), \\ T_{e_a m} &= \tau_{-a} T_m, && \text{with } \ \tau_a f(x) := f(x-a), \\ T_{m\circ A} &= D_{A^{\mathrm{t}}} T_m D_{A^{-\mathrm{t}}}, && \text{with } D_A f(x) := f(A^{-1}x). \end{aligned} \tag{5.35}$$

Here A^{t} denotes the transpose of A and $A^{-\mathrm{t}}$ its inverse.

Proof. The claims on the multiplier norms are immediate from the operator identities and the corresponding invariances of L^p-norms under the corresponding transformations, which are easily checked by direct computations. We provide the details of the first identity in (5.35), as this formula will play an important role in the subsequent developments:

$$\begin{aligned} T_m M_{-a} f = (m\widehat{M_{-a}f})^{\vee} &= (m(\cdot)\widehat{f}(\cdot + a))^{\vee} \\ &= ((m(\cdot - a)\widehat{f})(\cdot + a))^{\vee} = M_{-a} T_{m(\cdot - a)} f. \end{aligned}$$

Now multiply both sides with M_a. $\square$

A particular consequence of (5.35) is a representation for the multipliers arising from indicators of intervals; this is a continuous analogue of considerations that we already encountered in a somewhat *ad hoc* manner in the context of Fourier series.

Lemma 5.3.9. *We have, on $\check{L}^1(\mathbb{R}^d;X)$,*

$$\Delta_{[a,b)} := T_{\mathbf{1}_{[a,b)}} = \frac{i}{2}\Big(M_a H M_{-a} - M_b H M_{-b}\Big),$$

where H is the Hilbert transform.

Proof. Recall that $\Delta_{[0,\infty)} = \frac{1}{2}(I + iH)$; indeed, the symbols of both multiplier operators coincide except at $\xi = 0$, but this one point has no effect on the action of the operators, which ultimately depends on the equivalence class of $m \in L^\infty(\mathbb{R})$ only.

Observe that $\mathbf{1}_{[a,\infty)} = \mathbf{1}_{[0,\infty)}(\cdot - a)$, which by the previous lemma implies that $\Delta_{[a,\infty)} = M_a \Delta_{[0,\infty)} M_{-a}$. Thus

$$\begin{aligned} \Delta_{[a,b)} = \Delta_{[a,\infty)} - \Delta_{[b,\infty)} &= M_a \Delta_{[0,\infty)} M_{-a} - M_b \Delta_{[0,\infty)} M_{-b} \\ &= \frac{i}{2}\Big(M_a H M_{-a} - M_b H M_{-b}\Big), \end{aligned}$$

since $M_a I M_{-a} = I$ cancels with $M_b I M_{-b} = I$. $\square$

Lemma 5.3.9 leads in particular to the following simple result on the $L^p(\mathbb{R};X)$-boundedness of the operators $\Delta_{[a,b)}$. For more substantial consequences, however, we will require somewhat more than the mere boundedness, and the explicit representation provided by Lemma 5.3.9 will play a critical role.

Proposition 5.3.10. *For a Banach space X and $p \in (1,\infty)$ the following are equivalent:*

(1) *the Hilbert transform H is bounded on $L^p(\mathbb{R};X)$;*
(2) *$\Delta_{[a,b)}$ is bounded on $L^p(\mathbb{R};X)$ for all $a < b$;*
(3) *$\Delta_{[0,1)}$ is bounded on $L^p(\mathbb{R};X)$.*

Moreover, in this case $\|\Delta_{[0,1)}\| = \|\Delta_{[a,b)}\|$ and

$$\frac{1}{2}\|H\|_{\mathscr{L}(L^p(\mathbb{R};X))} \leqslant \|\Delta_{[a,b)}\|_{\mathscr{L}(L^p(\mathbb{R};X))} \leqslant \|H\|_{\mathscr{L}(L^p(\mathbb{R};X))}.$$

By Theorem 5.1.1, all these properties are characterisations of X being a UMD space.

Proof. (1)⇒(2): This follows from Lemma 5.3.9.

(2)⇔(3): Since $\mathbf{1}_{[a,b)}$ can be realised as a translation and dilation of $\mathbf{1}_{[0,1)}$, Proposition 5.3.8 shows that $\|\Delta_{[a,b)}\| = \|\Delta_{[0,1)}\|$ for all $a < b$.

(2)⇒(1): If $f \in \check{\mathscr{D}}(\mathbb{R};X)$ is supported on $(-n,n)$, then

$$\|\Delta_{[0,\infty)}f\|_p = \|\Delta_{[0,n)}f\|_p \leqslant \|\Delta_{[0,n)}\|\,\|f\|_p = \|\Delta_{[0,1)}\|\,\|f\|_p.$$

From the reflection invariance of the space of multipliers, it follows that also $\|\Delta_{(-\infty,0]}\| = \|\Delta_{[0,\infty)}\| \leqslant \|\Delta_{[0,1)}\|$. Hence

$$\|H\| = \|-i(\Delta_{[0,\infty)} - \Delta_{(-\infty,0]})\| \leqslant \|\Delta_{[0,\infty)}\| + \|\Delta_{(-\infty,0]}\| \leqslant 2\|\Delta_{[0,1)}\|.$$

□

Multiplier theory in real Banach spaces

We conclude the subsection by describing a class of (scalar-valued) Fourier multipliers for which a meaningful theory can developed in real Banach spaces X as well. We do this via the notion of *complexification* $X_{\mathbb{C}} = \{x + iy : x, y \in X\}$ of such a space as given in Appendix B.4. Our purpose is to prove:

Proposition 5.3.11 (Multipliers on real Banach spaces). *Let X be a real Banach space, and $X_{\mathbb{C}}$ its complexification. If $m \in L^\infty(\mathbb{R}^d)$ has an even real part and an odd imaginary part, then $T_m \in \mathscr{L}(\check{L}^1(\mathbb{R}^d;X_{\mathbb{C}}))$ restricts to an operator $T_m \in \mathscr{L}(\check{L}^1(\mathbb{R}^d;X))$, where*

$$\check{L}^1(\mathbb{R}^d;X) := \big\{f \in \check{L}^1(\mathbb{R}^d;X_{\mathbb{C}}) : f \text{ takes values in } X\big\}.$$

Such a multiplier induces a bounded $T_m \in \mathscr{L}(L^p(\mathbb{R}^d;X))$ *if and only if it induces a bounded* $T_m \in \mathscr{L}(L^p(\mathbb{R}^d;X_{\mathbb{C}}))$, *and if* $X_{\mathbb{C}}$ *is equipped with the norm* $\| \ \|_{X_{\mathbb{C}}^{\gamma,p}}$ *defined in Subsection 2.1.b, then*

$$\|T_m\|_{\mathscr{L}(L^p(\mathbb{R}^d;X)} = \|T_m\|_{\mathscr{L}(L^p(\mathbb{R}^d;X_{\mathbb{C}}^{\gamma,p})}.$$

The multiplier $-i\,\mathrm{sgn}(\xi)$ of the Hilbert transform is, of course, a prime example of the class appearing in the proposition.

Proof. On $X_{\mathbb{C}}$, we can define the operator of *conjugation* $z = x+iy \mapsto \overline{z} = x-iy$, which is conjugate-linear, namely $\overline{\alpha z} = \overline{\alpha}\overline{z}$, where $\overline{\alpha}$ is the usual complex conjugate of $\alpha \in \mathbb{C}$. This conjugation has a natural pointwise extension on $L^p(\mathbb{R}^d;X_{\mathbb{C}})$, which obeys the same algebraic rules, and f takes values in X if and only if $\overline{f} = f$. It is then immediate from the definition of the Fourier transform that

$$\overline{\widehat{f}(\xi)} = (\overline{f})^{\vee}(\xi), \qquad \overline{\check{f}(\xi)} = (\overline{f})^{\wedge}(\xi).$$

The reflection $\widetilde{f}(\xi) := f(-\xi)$ obeys a similar rule

$$\widetilde{\widehat{f}(\xi)} = (\widetilde{f})^{\vee}(\xi), \qquad \widetilde{\check{f}(\xi)} = (\widetilde{f})^{\wedge}(\xi).$$

Putting these together for $f \in \check{L}^1(\mathbb{R}^d;X_{\mathbb{C}})$ and $m \in L^\infty(\mathbb{R}^d)$, we compute

$$\overline{T_m f} = \overline{(m\widehat{f})^{\vee}} = (\overline{m}(\overline{f})^{\vee})^{\wedge} = (\widetilde{\overline{m}}(\overline{f})^{\wedge})^{\vee} = T_{\widetilde{\overline{m}}}\overline{f}.$$

If m is as in the assumptions of the proposition, then $\widetilde{\overline{m}} = m$, and hence $\overline{T_m f} = T_m\overline{f}$ for all $f \in \check{L}^1(\mathbb{R}^d;X_{\mathbb{C}})$, and $\overline{f} = f$ if f is X-valued. But then $\overline{T_m f} = T_m f$, and $T_m f$ is also X-valued. Thus $\check{L}^1(\mathbb{R}^d;X) \subseteq \check{L}^1(\mathbb{R}^d;X_{\mathbb{C}})$ is an invariant subspace for T_m, and $T_m \in \mathscr{L}(\check{L}^1(\mathbb{R}^d;X))$ by restriction.

The L^p-boundedness result follows from Proposition 2.1.12, since T_m on $L^p(\mathbb{R}^d;X_{\mathbb{C}}^{\gamma,p})$ may be interpreted as the complex extension $(T_m)_{\mathbb{C}}$ of T_m on $L^p(\mathbb{R}^d;X)$, in the sense of the quoted proposition. □

Remark 5.3.12. For the special multipliers appearing in Proposition 5.3.11, it is not difficult to give an intrinsic description (avoiding the complexification $X_{\mathbb{C}}$) of the operator T_m on $L^p(\mathbb{R}^d;X)$ in terms of the *sine* and *cosine transforms*

$$Sf(\xi) := \int_{\mathbb{R}^d} f(x)\sin(2\pi x\cdot\xi)\,\mathrm{d}x, \qquad Cf(\xi) := \int_{\mathbb{R}^d} f(x)\cos(2\pi x\cdot\xi)\,\mathrm{d}x.$$

However, since the resulting formulae would hardly provide additional insight, and since everything interesting that we are able to say about the multipliers depends on the complexified version anyway, we do not insist too much on this point.

5.3.b *R*-boundedness: a necessary condition for multipliers

A far-reaching necessary condition for membership in $\mathfrak{M}L^p(\mathbb{R}^d;X,Y)$ can be stated in terms of the following operator-theoretic notion that we study systematically in Volume II. For our immediate needs, we require little more than the definition:

Definition 5.3.13 (R-boundedness). *A family of operators $\mathscr{T}\subseteq\mathscr{L}(X,Y)$ is called* R-bounded *if for some $p\in[1,\infty)$ and $C<\infty$ we have*

$$\Big\|\sum_{n=1}^N \varepsilon_n T_n x_n\Big\|_{L^p(\Omega;Y)} \leqslant C\Big\|\sum_{n=1}^N \varepsilon_n x_n\Big\|_{L^p(\Omega;X)}$$

for all choices of $x_1,\dots,x_N\in X$, $T_1,\dots,T_N\in\mathscr{T}$, and $N\in\mathbb{Z}_+$.

The last admissible constant C is denoted by $\mathscr{R}_p(\mathscr{T})$ and is called the $\mathscr{R}_p$-bound of $\mathscr{T}$, notation $\mathscr{R}_p(\mathscr{T})$.

Taking $N=1$ in the definition, we see that every R-bounded family of operators is uniformly bounded and

$$\sup_{T\in\mathscr{T}}\|T\|_{\mathscr{L}(X,Y)}\leqslant\mathscr{R}_p(\mathscr{T}).$$

If X and Y are Hilbert spaces, then also the converse holds. The Kahane contraction principle asserts that bounded sets in $\mathbb{K}$ are R-bounded (identifying scalars $c\in\mathbb{K}$ with the operators $x\mapsto cx$ on X).

Remark 5.3.14. Let us make two further easy observations.

(1) If $\mathscr{T}$ and $\mathscr{S}$ are R-bounded, then so are $\mathscr{T}\mathscr{S}:=\{TS:T\in\mathscr{T},S\in\mathscr{S}\}$ and $\mathscr{T}+\mathscr{S}:=\{T+S:T\in\mathscr{T},S\in\mathscr{S}\}$, and

$$\begin{aligned}\mathscr{R}_p(\mathscr{T}\mathscr{S})&\leqslant\mathscr{R}_p(\mathscr{T})\mathscr{R}_p(\mathscr{S}),\\ \mathscr{R}_p(\mathscr{T}+\mathscr{S})&\leqslant\mathscr{R}_p(\mathscr{T})+\mathscr{R}_p(\mathscr{S}).\end{aligned}$$

(2) By the Kahane–Khintchine inequality (Theorem 3.2.23), once the R-boundedness inequality holds for one $p\in[1,\infty)$, then it holds for all $p\in[1,\infty)$, and

$$\mathscr{R}_p(\mathscr{T})\leqslant\kappa_{p,q}\kappa_{q,p}\mathscr{R}_q(\mathscr{T}).$$

In terms of this notion, we have:

Theorem 5.3.15 (Clément–Prüss). *Let X,Y be a Banach spaces and let $p\in(1,\infty)$. For all $m\in\mathfrak{M}L^p(\mathbb{R}^d;X,Y)$ the set*

$$\{m(\xi):\xi \text{ is a Lebesgue point of } m\}$$

is R-bounded, and

$$\begin{aligned}\|m\|_{L^\infty(\mathbb{R}^d;\mathscr{L}(X,Y))}&\leqslant\mathscr{R}_p(m(\xi):\xi \text{ is a Lebesgue point of } m)\\ &\leqslant\|m\|_{\mathfrak{M}L^p(\mathbb{R}^d;X,Y)}.\end{aligned}$$

Proof. The first estimate is immediate from Remark 5.3.14 and the fact that almost every point is a Lebesgue point (Theorem 2.3.4). We concentrate on the second bound.

Let $\xi_1, \dots, \xi_N$ be Lebesgue points of m. We consider two Schwartz functions $\phi, \psi \in \mathscr{S}(\mathbb{R}^d)$ with the property that $\int \widehat{\phi}\widecheck{\psi} \, \mathrm{d}\xi = 1$. Since m is assumed only locally integrable first, let $\widehat{\phi}$, $\widecheck{\psi}$ be compactly supported. Then by Theorem 2.3.8

$$
\begin{aligned}
m(\xi_k)x_k &= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon^d} \int_{\mathbb{R}^d} m(\xi)\widehat{\phi}\Big(\frac{\xi-\xi_k}{\varepsilon}\Big)x_k\widecheck{\psi}\Big(\frac{\xi-\xi_k}{\varepsilon}\Big) \, \mathrm{d}\xi \\
&= \lim_{\varepsilon\to 0} \varepsilon^d \int_{\mathbb{R}^d} m(\xi)\mathscr{F}(e_{\xi_k}\cdot\phi(\varepsilon\cdot))(\xi)x_k\mathscr{F}^{-1}(e_{-\xi_k}\cdot\psi(\varepsilon\cdot))(\xi) \, \mathrm{d}\xi \\
&= \lim_{\varepsilon\to 0} \varepsilon^d \int_{\mathbb{R}^d} T_m[e_{\xi_k}\cdot\phi(\varepsilon\cdot)x_k](y)e_{-\xi_k}(y)\psi(\varepsilon y) \, \mathrm{d}y
\end{aligned}
$$

and hence, using the contraction principle to pull out the bounded factors $e_{\mp\xi_k}(y)$ and the assumed boundedness of the operator T_m, we find that

$$
\begin{aligned}
&\Big\| \sum_{k=1}^N \varepsilon_k m(\xi_k)x_k \Big\|_{L^p(\Omega;Y)} \\
&\leqslant \liminf_{\varepsilon\to 0} \varepsilon^d \int_{\mathbb{R}^d} \Big\| \sum_{k=1}^N \varepsilon_k e_{-\xi_k}(y)T_m[e_{\xi_k}\cdot\phi(\varepsilon\cdot)x_k](y) \Big\|_{L^p(\Omega;Y)} |\psi(\varepsilon y)| \, \mathrm{d}y \\
&\leqslant \liminf_{\varepsilon\to 0} \varepsilon^d \Big\| \sum_{k=1}^N \varepsilon_k e_{-\xi_k} T_m[e_{\xi_k}\cdot\phi(\varepsilon\cdot)x_k] \Big\|_{L^p(\mathbb{R}^n\times\Omega;Y)} \|\psi(\varepsilon\cdot)\|_{L^{p'}(\mathbb{R}^n)} \\
&\leqslant \liminf_{\varepsilon\to 0} \varepsilon^d \Big\| T_m \sum_{k=1}^N \varepsilon_k e_{\xi_k}\cdot\phi(\varepsilon\cdot)x_k \Big\|_{L^p(\mathbb{R}^n\times\Omega;Y)} \|\psi(\varepsilon\cdot)\|_{p'} \\
&\leqslant \liminf_{\varepsilon\to 0} \varepsilon^d \|m\|_{\mathfrak{M}L^p(\mathbb{R}^d;X,Y)} \Big\| \sum_{k=1}^N \varepsilon_k \phi(\varepsilon\cdot)x_k \Big\|_{L^p(\mathbb{R}^d\times\Omega;X)} \|\psi(\varepsilon\cdot)\|_{p'} \\
&= \|m\|_{\mathfrak{M}L^p(\mathbb{R}^d;X,Y)} \Big\| \sum_{k=1}^N \varepsilon_k x_k \Big\|_{L^p(\Omega;X)} \liminf_{\varepsilon\to 0} \varepsilon^d \|\phi(\varepsilon\cdot)\|_p \|\psi(\varepsilon\cdot)\|_{p'}.
\end{aligned}
$$

By change of variable,

$$
\varepsilon^d \|\phi(\varepsilon\cdot)\|_p \|\psi(\varepsilon\cdot)\|_{p'} = \|\phi\|_p \|\psi\|_{p'}
$$

is independent of $\epsilon > 0$. Let us choose specifically $\phi(x) = e^{-\pi|x|^2/p}$, $\psi(x) = e^{-\pi|x|^2/p'}$, which satisfy

$$
1 = \|\phi\|_p = \|\psi\|_{p'} = \int \phi\psi \, \mathrm{d}y = \int \widehat{\phi}\widecheck{\psi} \, \mathrm{d}\xi.
$$

This gives the right constants as in the assertion and completes the proof. □

The following Fatou type result for $\mathfrak{M}L^p(\mathbb{R}^d;X)$ should be regarded as part of the general theory of the previous subsection, although we prove it here as an application of the previous theorem. Note, however, that we only use the estimate between the L^∞ and $\mathfrak{M}L^p$ norms of m, not the stronger statement concerning R-boundedness.

Proposition 5.3.16. *Let $(m_k)_{n\geqslant 1}$ be a bounded sequence in $\mathfrak{M}L^p(\mathbb{R}^d;X,Y)$. If for all $x \in X$ we have almost everywhere $m_k x \to mx$, then $m \in \mathfrak{M}L^p(\mathbb{R}^d;X,Y)$ and*

$$\|m\|_{\mathfrak{M}L^p(\mathbb{R}^d;X,Y)} \leqslant \liminf_{k\to\infty} \|m_k\|_{\mathfrak{M}L^p(\mathbb{R}^d;X,Y)}.$$

Moreover, $\mathfrak{M}L^p(\mathbb{R}^d;X,Y)$ is a Banach algebra with respect to pointwise multiplication.

Proof. By Theorem 5.3.15, $(m_k)_{n\geqslant 1}$ is bounded in $L^\infty(\mathbb{R};\mathscr{L}(X,Y))$. Let now $f \in \check{L}^1(\mathbb{R}^d)\otimes X$ be given. By the dominated convergence theorem,

$$T_{m_k}f(x) = \int_{\mathbb{R}^d} e^{2\pi i\xi\cdot x} m_k(\xi)\widehat{f}(\xi)\,\mathrm{d}\xi \to \int_{\mathbb{R}^d} e^{2\pi i\xi\cdot x} m(\xi)\widehat{f}(\xi)\,\mathrm{d}\xi = T_m f(x)$$

for all $x\in\mathbb{R}^d$ as $k\to\infty$. Therefore, by Fatou's lemma,

$$\begin{aligned}\|T_m f\|_{L^p(\mathbb{R}^d;Y)} &= \liminf_{k\to\infty}\|T_{m_k}f\|_{L^p(\mathbb{R}^d;Y)}\\ &\leqslant \liminf_{k\to\infty}\|m_k\|_{\mathfrak{M}L^p(\mathbb{R}^d;X,Y)}\|f\|_{L^p(\mathbb{R}^d;X)}.\end{aligned}$$

This implies the required estimate for m.

To prove that $\mathfrak{M}L^p(\mathbb{R}^d;X,Y)$ is a Banach space let $(m_k)_{k\geqslant 1}$ be a Cauchy sequence in $\mathfrak{M}L^p(\mathbb{R}^d;X,Y)$. Then it is a Cauchy sequence in $L^\infty(\mathbb{R}^d;\mathscr{L}(X,Y))$ and hence convergent to some $m \in L^\infty(\mathbb{R}^d,\mathscr{L}(X,Y))$. By the above, $m \in \mathfrak{M}L^p(\mathbb{R}^d;X,Y)$, and for all $k\geqslant 1$ we have

$$\|m-m_k\|_{\mathfrak{M}L^p(\mathbb{R}^d;X,Y)} \leqslant \liminf_{j\to\infty}\|m_j-m_k\|_{\mathfrak{M}L^p(\mathbb{R}^d;X,Y)}.$$

That $\mathfrak{M}L^p(\mathbb{R}^d;X,Y)$ is a Banach algebra is immediate from Lemma 5.3.2. $\square$

5.3.c Mihlin's multiplier theorem on $\mathbb{R}$

The problem of finding useful characterisations of the class $\mathfrak{M}L^p(\mathbb{R}^d;X)$ of all Fourier multipliers of $L^p(\mathbb{R}^d;X)$ for $p\in(1,\infty)\setminus\{2\}$ is open even in the scalar case, and the main interest has therefore been in finding workable sufficient conditions that cover a wide range of applications. While the class $\mathfrak{M}L^p(\mathbb{R}^d;X)$ itself is invariant under translations, it is not uncommon for the multipliers arising in applications that they exhibit a singular behaviour in the neighbourhood of a distinguished point of their domain, the origin, and this has also guided the formulation of the multiplier theorems. The main result of this section is the following sufficient condition for multipliers on the line, which has turned out very useful in applications. The appearance of R-boundedness in this condition is natural in view of the necessary condition given in Theorem 5.3.15.

Definition 5.3.17 (Mihlin's class). *A function $m \in L^\infty(\mathbb{R}; \mathscr{L}(X,Y))$ belongs to* Mihlin's class $\mathfrak{M}(\mathbb{R}; X, Y)$ *if m is continuous and piecewise continuously differentiable on each*

$$I \in \mathscr{I} := \{(2^j, 2^{j+1}), (-2^{j+1}, -2^j) : j \in \mathbb{Z}\},$$

and the following Mihlin norm is finite:

$$\begin{aligned}\|m\|_{\mathfrak{M}_p(\mathbb{R};X,Y)} := \sup_{|\xi|\in(1,2)} &\mathscr{R}_p(\{m(2^k\xi)\}_{k\in\mathbb{Z}}) \\ &+ \sup_{|\xi|\in(1,2)} \mathscr{R}_p(\{2^k\xi m'(2^k\xi)\}_{k\in\mathbb{Z}}).\end{aligned}$$

We write simply $\mathfrak{M}(\mathbb{R}) := \mathfrak{M}(\mathbb{R}; \mathbb{C}, \mathbb{C})$ for the scalar-valued case.

Clearly $\|m\|_{\mathfrak{M}_p(\mathbb{R};X,Y)}$ is dominated by the quantity

$$\mathscr{R}_p(\{m(\xi) : \xi \in \mathbb{R} \setminus \{0, \pm 2^k\}_{k\in\mathbb{Z}}\}) + \mathscr{R}_p(\{\xi m'(\xi) : \xi \in \mathbb{R} \setminus \{0, \pm 2^k\}_{k\in\mathbb{Z}}\}),$$

and for most purposes a multiplier theorem involving this larger quantity would be quite sufficient. It is perhaps mainly a theoretical curiosity that R-boundedness, rather than mere uniform boundedness, is only needed over the dyadic scales. However, this stronger form of the theorem follows with negligible additional effort. For scalar-valued multipliers, both versions coincide, and reduce to the uniform bound

$$\|m\|_{\mathfrak{M}(\mathbb{R})} := \sup_{\xi\in\mathbb{R}\setminus\{0,\pm 2^k\}_{k\in\mathbb{Z}}} |m(\xi)| + \sup_{\xi\in\mathbb{R}\setminus\{0,\pm 2^k\}_{k\in\mathbb{Z}}} |\xi m'(\xi)|.$$

It is the finiteness of the Mihlin norm $\|m\|_{\mathfrak{M}_p(\mathbb{R};X,Y)}$ rather than the precise sense of differentiability that is most relevant to the subject at hand. In essence, what we really need is that the function m may be locally recovered as the integral of its derivative, and the educated reader can surely find weaker assumptions that guarantee such a property. Our chosen definition is sufficiently general to cover a plenitude of applications, some of them treated later in these volumes, yet sufficiently strong to allow straightforward calculus when working with this condition. Note that the assumptions on m easily imply the existence of one-sided limits at the end-points of the intervals $I \in \mathscr{I}$.

Theorem 5.3.18 (Mihlin's multiplier theorem). *Let X and Y be complex UMD spaces and let $p \in (1, \infty)$. Then $\mathfrak{M}(\mathbb{R}; X, Y) \subseteq \mathfrak{M}L^p(\mathbb{R}; X, Y)$, and*

$$\begin{aligned}\|m\|_{\mathfrak{M}L^p(\mathbb{R};X,Y)} &\leqslant 400\min(\hbar_{p,X},\hbar_{p,Y})\beta^+_{p,X}\beta^+_{p',Y^*}\|m\|_{\mathfrak{M}_p(\mathbb{R};X,Y)}\\ &\leqslant 400\min(\hbar_{p,X},\hbar_{p,Y})\beta_{p,X}\beta_{p,Y}\|m\|_{\mathfrak{M}_p(\mathbb{R};X,Y)},\end{aligned}$$

where $\hbar_{p,X}$ is the norm of the Hilbert transform on $L^p(\mathbb{R};X)$. For $X=Y$ and scalar-valued m with $\overline{m(-\xi)}=m(\xi)$, the result is valid for real UMD spaces X as well.

It is clear that UMD is a necessary assumption for Theorem 5.3.18, since it is already necessary in the simplest case $m=-i\,\mathrm{sgn}\in\mathfrak{M}(\mathbb{R})$ of the Hilbert transform. We do not know whether the order of the quantitative bound of Theorem 5.3.18 can be improved; however, it is essentially optimal (except for the numerical factor, which could be somewhat reduced by more careful calculations) for the method of proof presented here, which involves two applications of an unconditionality estimate, each producing one of the UMD constants $\beta^+_{p,X}\leqslant\beta_{p,X}$ and $\beta^+_{p',Y^*}\leqslant\beta_{p',Y^*}=\beta_{p,Y}$, together with an application of the Hilbert transform boundedness, which can be performed in either space, as one likes.

As an illustration of multipliers covered by this theorem, we give:

Corollary 5.3.19. *Let X be a complex UMD space and let $p\in(1,\infty)$. Then the imaginary powers $m_s(\xi):=|\xi|^{is}$, $s\in\mathbb{R}$, are Fourier multipliers for $L^p(\mathbb{R};X)$, and*

$$\|T_{m_s}\|_{\mathscr{L}(L^p(\mathbb{R};X))}\leqslant 400\,\hbar_{p,X}\beta^2_{p,X}(1+|s|).$$

Proof. It suffices to check that $\|m_s\|_{\mathfrak{M}(\mathbb{R})}\leqslant 1+|s|$. This (in fact, with equality) follows from $m_s'(\xi)=|\xi|^{is}is/\xi$, so that

$$|m_s(\xi)|=1,\qquad |\xi m_s'(\xi)|=|s|,\qquad \forall\xi\in\mathbb{R}\setminus\{0\}.$$

□

The discontinuities at $\pm 2^j$ (and at 0) allowed for m in Mihlin's multiplier class are not only useful for a number of applications, but also natural in view of the method of proof: in particular, they allow an easy piecewise construction of auxiliary multipliers, avoiding the need of introducing additional smooth truncations at certain places of the argument. For later uses, especially in the shortly forthcoming proof of Mihlin's theorem, we record the following easy invariance properties of the multiplier class $\mathfrak{M}(\mathbb{R};X,Y)$.

Proposition 5.3.20 (Invariances of the Mihlin class). *Let $m,m_1,m_2\in\mathfrak{M}(\mathbb{R};X,Y)$, $m_3\in\mathfrak{M}(\mathbb{R};Y,Z)$ and $m_I\in\mathfrak{M}(\mathbb{R})$. Then the following multipliers also belong to the Mihlin class (for appropriate spaces), with the indicated estimates for their norms:*

(1) *The pointwise sum and product m_1+m_2 and m_3m_2, with*

$$\begin{aligned}\|m_1+m_2\|_{\mathfrak{M}(\mathbb{R};X,Y)} &\leqslant \|m_1\|_{\mathfrak{M}(\mathbb{R};X,Y)}+\|m_2\|_{\mathfrak{M}(\mathbb{R};X,Y)},\\ \|m_3m_2\|_{\mathfrak{M}(\mathbb{R};X,Z)} &\leqslant \|m_3\|_{\mathfrak{M}(\mathbb{R};Y,Z)}\|m_2\|_{\mathfrak{M}(\mathbb{R};X,Y)}.\end{aligned}$$

(2) *The dyadic dilations and reflection* $m(\pm 2^j\cdot)$, $j\in\mathbb{Z}$, *with*

$$\|m(\pm 2^j\cdot)\|_{\mathfrak{M}(\mathbb{R};X,Y)} = \|m\|_{\mathfrak{M}(\mathbb{R};X,Y)}.$$

(3) *The sum of dyadic pieces* $\sum_{I\in\mathscr{I}}\mathbf{1}_I m_I$ *of multipliers* m_I *of uniformly bounded Mihlin norm:*

$$\Big\|\sum_{I\in\mathscr{I}}\mathbf{1}_I m_I\Big\|_{\mathfrak{M}(\mathbb{R})} \leqslant \sup_{I\in\mathscr{I}}\|m_I\|_{\mathfrak{M}(\mathbb{R})}.$$

Proof. The piecewise continuity and differentiability properties of the new multipliers are immediate in each case, so it remains to verify the quantitative bounds. For the sum m_1+m_2 this is clear, and for the product m_3m_2 it follows from the sum and product rule of R-bounds and the computation

$$\begin{aligned}
&\mathscr{R}_p(m_1(u\xi)m_2(u\xi):u)+\mathscr{R}_p(u\xi(m_1m_2)'(u\xi):u)\\
&\quad\leqslant \mathscr{R}_p(m_1(u\xi)m_2(u\xi))+\mathscr{R}_p(u\xi m_1'(u\xi)m_2(u\xi):u)\\
&\qquad+\mathscr{R}_p(u\xi m_1(u\xi)m_2'(u\xi):u)\\
&\quad\leqslant \mathscr{R}_p(m_1(u\xi):u)\mathscr{R}_p(m_2(u\xi):u)+\mathscr{R}_p(u\xi m_1'(u\xi):u)\mathscr{R}_p(m_2(u\xi):u)\\
&\qquad+\mathscr{R}_p(m_1(u\xi):u)\mathscr{R}_p(u\xi m_2'(u\xi):u)\\
&\quad\leqslant \Big(\mathscr{R}_p(m_1(u\xi):u)+\mathscr{R}_p(u\xi m_1'(u\xi):u)\Big)\\
&\qquad\times\Big(\mathscr{R}_p(m_2(u\xi):u)+\mathscr{R}_p(u\xi m_2'(u\xi):u)\Big),
\end{aligned}$$

where u ranges over $u=2^k$, $k\in\mathbb{Z}$, in each case. For dilations the factor ξ in front of the derivative is critical for the invariance, and we get

$$\xi\frac{d}{d\xi}(\xi\mapsto m(\pm 2^j\xi)) = \eta m'(\eta), \qquad \eta=\pm 2^j\xi,$$

which again yields the asserted bound by taking relevant R-bounds. For case (3), we just observe that, at every point $\xi\in I$ with any one of the intervals $I\in\mathscr{I}$, the value $|m(\xi)|+|\xi m'(\xi)|$ for the new multiplier m coincides with a similar expression with one of the m_I in place of m. □

Proof of Mihlin's multiplier theorem

We only consider the complex case, as the real version is then an immediate corollary of Proposition 5.3.11. The strategy of proof of Theorem 5.3.18 is to decompose m into sums and products of simpler pieces that we can estimate either in terms of the Hilbert transform, or the Littlewood–Paley inequality of Proposition 5.1.10. In doing so, the algebra properties recorded in Lemma 5.3.2 are very handy, as they immediately allow to deduce operator identities from the corresponding identities for the functions m. We shall often apply this reasoning without explicit mentioning.

A rough outline of this strategy is as follows:

- From the Hilbert transform, we first build the multipliers of arbitrary intervals, $\Delta_I := T_{\mathbf{1}_I}$.
- A piecewise C^1 function on an interval is expressed as a superposition of indicators according to

$$\begin{aligned}\mathbf{1}_{[a,b)}(\xi)m(\xi) &= \mathbf{1}_{[a,b)}(\xi)\Big(m(a)+\int_a^\xi m'(\eta)\,\mathrm{d}\eta\Big)\\ &= \mathbf{1}_{[a,b)}(\xi)m(a)+\int_a^b \mathbf{1}_{[\eta,b)}(\xi)m'(\eta)\,\mathrm{d}\eta,\end{aligned}\tag{5.36}$$

  which leads to a formula for $T_{\mathbf{1}_I m}$ in terms of Δ_J's for $J \subseteq I$.
- Finally, we invoke the Littlewood–Paley inequality to sum up the pieces $\mathbf{1}_I m$ with $I \in \mathscr{I}$ with appropriate control.

We now turn to the details. If we expand the function $m = \sum_{I\in\mathscr{I}} \mathbf{1}_I m$ (off the countably many points 0 and $\pm 2^j$), the boundedness of T_m asks that

$$\Big\| \sum_{I\in\mathscr{I}} T_{\mathbf{1}_I m} f_I \Big\|_{L^p(\mathbb{R};Y)} \lesssim \Big\| \sum_{I\in\mathscr{I}} f_I \Big\|_{L^p(\mathbb{R};X)}, \qquad f_I := \Delta_I f = T_{\mathbf{1}_I} f.$$

As a first major step towards this goal, we first prove the following randomised version, which only requires one of the spaces X and Y to be UMD:

Lemma 5.3.21. *Let X or Y be a UMD space, $p \in (1,\infty)$, and $m \in \mathfrak{M}(\mathbb{R}; X, Y)$. For all choices of functions $f_I \in L^p(\mathbb{R}; X)$, $I \in \mathscr{I}$, and $\sigma \in \{-,+\}$ we have*

$$\begin{aligned}&\Big\| \sum_{I\in\mathscr{I}_\sigma} \varepsilon_I T_{\mathbf{1}_I m} f_I \Big\|_{L^p(\Omega\times\mathbb{R};Y)}\\ &\leqslant 2\|m\|_{\mathfrak{M}(\mathbb{R};X,Y)} \min(\hbar_{p,X}, \hbar_{p,Y}) \Big\| \sum_{I\in\mathscr{I}_\sigma} \varepsilon_I f_I \Big\|_{L^p(\Omega\times\mathbb{R};X)},\end{aligned}$$

where $\mathscr{I}_\sigma := \{I \in \mathscr{I} : I \subseteq \mathbb{R}_\sigma\}$.

In contrast to the motivating discussion above, we are not assuming any specific form of the functions $f_I \in L^p(\mathbb{R}; X)$ here.

Proof. In order to conveniently handle the positive and negative half-lines at one strike, we use the notation $(-t)(a,b) = (-tb,-ta)$ and $t(a,b) = (ta,tb)$ for $t > 0$ and let $\sigma = \pm 1$. From (5.36) and a change of variable, we have

$$\mathbf{1}_{\sigma(2^j,2^{j+1})}(\xi)m(\xi) = \mathbf{1}_{\sigma(2^j,2^{j+1})}(\xi)m(\sigma 2^j) + \int_1^2 \mathbf{1}_{\sigma(2^j t,2^{j+1})}(\xi)m'(\sigma 2^j t)\sigma 2^j\,\mathrm{d}t,$$

and thus

$$T_{\mathbf{1}_{\sigma(2^j,2^{j+1})}m} f = m(\sigma 2^j)\Delta_{\sigma(2^j,2^{j+1})} f + \int_1^2 \sigma 2^j m'(\sigma 2^j t)\Delta_{\sigma(2^j t,2^{j+1})} f\,\mathrm{d}t$$

$$= \frac{i}{2} \sum_{k=1}^{2} (-1)^{k-1} \Big(m(\sigma 2^j) e_{\sigma k 2^j} H(e_{-\sigma k 2^j} f) \\ + \int_1^2 \sigma 2^j m'(\sigma 2^j t) e_{\sigma(k \vee t) 2^j} H(-e_{\sigma(k\vee t)2^j} f) \, \mathrm{d}t \Big),$$

noting that

$$(k \vee t) 2^j = \begin{cases} t2^j & \text{if } k = 1, \\ 2^{j+1} & \text{if } k = 2. \end{cases}$$

Summing over all $I = \sigma(2^j, 2^{j+1}) = \sigma(|I|, 2|I|) \in \mathscr{I}_\sigma$, i.e., over all $j \in \mathbb{Z}$, we find that

$$\begin{aligned} &\Big\| \sum_{I \in \mathscr{I}_\sigma} \varepsilon_I T_{\mathbf{1}_I m} f_I \Big\|_p \\ &\leqslant \frac{1}{2} \sum_{k=1}^{2} \Big(\Big\| \sum_{I \in \mathscr{I}_\sigma} \varepsilon_I m(\sigma_I |I|) e_{\sigma k|I|} H(e_{-\sigma_I k|I|} f_I) \Big\|_p \\ &\quad + \int_1^2 \Big\| \sum_{I \in \mathscr{I}_\sigma} \varepsilon_I \sigma_I t |I| m'(\sigma_I t|I|) e_{\sigma(k\vee t)|I|} H(e_{-\sigma_I (k\vee t)|I|} f_I) \Big\|_p \frac{\mathrm{d}t}{t} \Big). \end{aligned} \tag{5.37}$$

For the various expressions of the form

$$\Big\| \sum_{I \in \mathscr{I}} \varepsilon_I m_I e_{a_I} H(e_{-a_I} f_I) \Big\|_p,$$

with $m_I = m(\sigma|I|)$ or $m_I = \sigma t|I| m'(\sigma t|I|)\}$, we have

$$\mathscr{R}_p(m_I : I \in \mathscr{I}) \leqslant \|m\|_{\mathfrak{M}_p(\mathbb{R};X,Y)},$$

and we apply the chain of estimates

$$\begin{aligned} &\Big\| \sum_{I \in \mathscr{I}_\sigma} \varepsilon_I m_I e_{a_I} H(e_{-a_I} f_I) \Big\|_{L^p(\mathbb{R}\times\Omega;Y)} \\ &\leqslant \|m\|_{\mathfrak{M}(\mathbb{R};X,Y)} \Big\| \sum_{I \in \mathscr{I}_\sigma} \varepsilon_I e_{a_I} H(e_{-a_I} f_I) \Big\|_{L^p(\mathbb{R}\times\Omega;X)} \\ &\leqslant \|m\|_{\mathfrak{M}(\mathbb{R};X,Y)} \Big\| \sum_{I \in \mathscr{I}_\sigma} \varepsilon_I H(e_{-a_I} f_I) \Big\|_{L^p(\mathbb{R}\times\Omega;X)} \\ &\leqslant \|m\|_{\mathfrak{M}(\mathbb{R};X,Y)} \hbar_{p,X} \Big\| \sum_{I \in \mathscr{I}_\sigma} \varepsilon_I e_{-a_I} f_I \Big\|_{L^p(\mathbb{R}\times\Omega;X)} \\ &\leqslant \|m\|_{\mathfrak{M}(\mathbb{R};X,Y)} \hbar_{p,X} \Big\| \sum_{I \in \mathscr{I}_\sigma} \varepsilon_I f_I \Big\|_{L^p(\mathbb{R}\times\Omega;X)}, \end{aligned}$$

consisting the definition of R-boundedness, the contraction principle with the factors $m_I e_{a_I}(x)$, the linearity and the boundedness of the Hilbert transform on $L^p(\mathbb{R};X)$, and another application of the contraction principle with the factors $e_{-a_I}(x)$.

Alternatively, we could have used the commutativity of the operators to write

$$m_I e_{a_I} H(e_{-a_I} f_I) = e_{a_I} H(e_{-a_I} m_I f_I),$$

using the R-boundedness as the last step rather than the first, and the boundedness of the Hilbert transform on $L^p(\mathbb{R};Y)$ instead of $L^p(\mathbb{R};X)$. This gives the factor $\hbar_{p,Y}$ instead of $\hbar_{p,X}$, and we can take the better of the two alternative estimates.

Substituting this back to (5.37), we find that

$$\begin{aligned}
&\Big\| \sum_{I\in\mathscr{I}_\sigma} \varepsilon_I T_{\mathbf{1}_I m} f_I \Big\|_{L^p(\mathbb{R}\times\Omega;Y)} \\
&\leqslant \frac{1}{2}\sum_{k=1}^{2}\Big(1+\int_1^2 \frac{\mathrm{d}t}{t}\Big)\|m\|_{\mathfrak{M}(\mathbb{R};X,Y)} \min(\hbar_{p,X},\hbar_{p,Y})\Big\| \sum_{I\in\mathscr{I}_\sigma} \varepsilon_I f_I \Big\|_{L^p(\mathbb{R}\times\Omega;X)},
\end{aligned}$$

where the numerical factor is estimated as $\frac{1}{2}\sum_{k=1}^{2}\big(1+\int_1^2 \frac{\mathrm{d}t}{t}\big) \leqslant 2$. □

In order to be able to further estimate the right side of the bound of Lemma 5.3.21 by $\|f\|_{L^p(\mathbb{R};X)}$, we need to make a suitable choice of functions f_I. For this purpose, we record here the following reparametrised version of Proposition 5.1.10 that is convenient for our further analysis:

Lemma 5.3.22. *Let*

$$\begin{aligned}
\phi(x) &:= h * \widetilde{h}(x), \qquad & h(x) &:= \mathbf{1}_{[0,\frac{1}{2})} - \mathbf{1}_{[\frac{1}{2},1)},\ \widetilde{h} := h(-\cdot), \\
\phi_I(x) &:= 2|I|\cdot\phi(2|I|x), & I &\in \mathscr{I}.
\end{aligned} \tag{5.38}$$

Then

$$\Big\| \sum_{I\in\mathscr{I}_\pm} \varepsilon_I \phi_I * f \Big\|_{L^p(\Omega\times\mathbb{R};X)} \leqslant \beta^+_{p,X}\|f\|_{L^p(\mathbb{R};X)}.$$

Proof. This is just a reparametrisation of the second bound in Proposition 5.1.10, observing that $2|I|$ goes through all dyadic powers 2^j, $j\in\mathbb{Z}$, as I goes through either $\mathscr{I}_+$ or $\mathscr{I}_-$. □

To apply Lemma 5.3.22 in the proof of Mihlin's theorem, we want to insert the convolution by ϕ_I, thus multiplication by $\widehat{\phi}_I$ on the Fourier transform side, into an expression where it was not present to begin with. So we also need to divide by $\widehat{\phi}_I$, and this introduces a need to have good estimates for this inverse:

Lemma 5.3.23. *The functions ϕ and ϕ_I from (5.38) satisfy*

$$\Big\| \sum_{I\in\mathscr{I}} \frac{\mathbf{1}_I}{\widehat{\phi}_I} \Big\|_{\mathfrak{M}(\mathbb{R})} = \sup_{I\in\mathscr{I}} \Big\| \frac{\mathbf{1}_I}{\widehat{\phi}_I} \Big\|_{\mathfrak{M}(\mathbb{R})} = \Big\| \frac{\mathbf{1}_{[\frac{1}{2},1)}}{\widehat{\phi}} \Big\|_{\mathfrak{M}(\mathbb{R})}$$

$$\leqslant \Big(\sup_{\xi\in[\frac12,1]} \frac{1}{|\widehat{\phi}(\xi)|} + \sup_{\xi\in[\frac12,1]} \Big|\frac{d}{d\xi}\frac{1}{\widehat{\phi}(\xi)}\Big|\Big) \leqslant \frac{\pi^2}{4} + 2\pi < 9.$$

Note that in the penultimate inequality bound, we have estimated the factor ξ in front of the derivative by $|\xi| \leqslant 1$; we record this explicitly, as we will make use of the last upper bound above, rather than just the estimate for the slightly smaller Mihlin norm, in the extension to several variables in Section 5.5 below.

Proof. Observe from (5.38) that $\phi_{[\frac12,1)} = \phi = \widetilde{\phi}$ (cf. Figure 5.1), so that each $\mathbf{1}_I/\widehat{\phi}_I$ is a dyadic dilation of $\mathbf{1}_{[\frac12,1)}/\widehat{\phi}$ if $I \in \mathscr{I}_+$, and a dyadic dilation with reflection if $I \in \mathscr{I}_-$. Thus the two identities are direct consequences of Proposition 5.3.20, and it remains to estimate $|1/\widehat{\phi}(\xi)| + |\xi\cdot(1/\widehat{\phi})'(\xi)|$ for $\xi\in[\frac12,1)$. A standard computation with the Fourier transform shows that

$$\widehat{\phi}(\xi) = \widehat{h}(\xi)\widehat{h}(-\xi) = \frac{\sin^4(\frac12\pi\xi)}{(\frac12\pi\xi)^2}.$$

For $\xi\in[\frac12,1]$, we have $t := \frac12\pi\xi \in [\frac14\pi, \frac12\pi]$ and here $\sin^2(t)/t \geqslant \frac{2}{\pi}$. Indeed, since $t\mapsto \sin^2(t) - \frac{2t}{\pi}$ is strictly concave on $[\frac14\pi,\frac12\pi]$, this easily follows. We find

$$\widehat{\phi}(\xi) = \frac{\sin^4 t}{t^2} \geqslant \frac{4}{\pi^2}$$

and by some elementary calculus one can check that

$$\begin{aligned}\Big|\frac{\widehat{\phi}'(\xi)}{\widehat{\phi}(\xi)^2}\Big| &= \frac{|4\sin^3 t\cdot\cos t\cdot t^2 - \sin^4 t\cdot 2t|}{t^4}\cdot\frac{\pi}{2}\cdot\frac{t^4}{\sin^8 t} \\ &= \Big|\frac{2t^2\cos t}{\sin^5(t)} - \frac{t}{\sin^4(t)}\Big|\cdot\pi \leqslant 2\pi.\end{aligned}$$

Thus

$$\sup_{\xi\in[\frac12,1]}\Big|\frac{1}{\widehat{\phi}(\xi)}\Big| + \sup_{\xi\in[\frac12,1]}\Big|\frac{d}{d\xi}\Big(\frac{1}{\widehat{\phi}(\xi)}\Big)\Big| \leqslant \frac{1}{4/\pi^2} + \Big|\frac{\widehat{\phi}'(\xi)}{\widehat{\phi}(\xi)^2}\Big| \leqslant \frac{\pi^2}{4} + 2\pi = 8.7505\ldots.$$

This completes the proof. □

Now we are ready for:

Proof of Mihlin's Multiplier Theorem 5.3.18. Fix a function $f\in\mathscr{S}(\mathbb{R};X)\subseteq \check{L}^1(\mathbb{R};X)$. Then $T_m f\in\check{L}^1(\mathbb{R};X)$. We shall estimate its $L^p(\mathbb{R};X)$ norm by dualising with $g\in\mathscr{S}(\mathbb{R};X^*)$, a norming subspace of the dual space. Using the basic properties of the Fourier transform and the estimates established above, we compute:

$$|\langle T_m f, g\rangle| = |\langle m\widehat{f}, \check{g}\rangle|$$

$$
\begin{aligned}
&= \Big|\sum_{I\in\mathscr{I}} \langle \mathbf{1}_I m \widehat{f}, \widecheck{g}\rangle\Big| \\
&= \Big|\sum_{\sigma\in\{-,+\}}\sum_{I\in\mathscr{I}_\sigma} \langle \frac{\mathbf{1}_I m}{\widehat{\phi}_I\widecheck{\phi}_I}\widehat{\phi}_I\widehat{f}, \widecheck{\phi}_I\widecheck{g}\rangle\Big| \\
&= \Big|\sum_{\sigma\in\{-,+\}} \mathbb{E}\Big\langle \sum_{I\in\mathscr{I}_\sigma} \varepsilon_I \frac{\mathbf{1}_I m}{\widehat{\phi}_I\widecheck{\phi}_I}\widehat{\phi}_I\widehat{f}, \sum_{I\in\mathscr{I}_\sigma} \bar{\varepsilon}_I\widecheck{\phi}_I\widecheck{g}\Big\rangle\Big| \\
&= \Big|\sum_{\sigma\in\{-,+\}} \mathbb{E}\Big\langle \sum_{I\in\mathscr{I}_\sigma} \varepsilon_I \mathbf{1}_I n\,\widehat{\phi}_I\widehat{f}, \sum_{I\in\mathscr{I}_\sigma} \bar{\varepsilon}_I\widecheck{\phi}_I\widecheck{g}\Big\rangle\Big| \quad \text{with } \ n := \sum_{I\in\mathscr{I}} \frac{\mathbf{1}_I m}{\widehat{\phi}_I\widecheck{\phi}_I} \\
&= \Big|\sum_{\sigma\in\{-,+\}} \mathbb{E}\Big\langle \sum_{I\in\mathscr{I}_\sigma} \varepsilon_I T_{\mathbf{1}_I n}\phi_I * f, \sum_{I\in\mathscr{I}_\sigma} \bar{\varepsilon}_I\phi_I * g\Big\rangle\Big| \\
&\leqslant \sum_{\sigma\in\{-,+\}} \Big\|\sum_{I\in\mathscr{I}_\sigma} \varepsilon_I T_{\mathbf{1}_I n}\phi_I * f\Big\|_{L^p(\mathbb{R}\times\Omega;Y)} \\
&\qquad\qquad \times \Big\|\sum_{I\in\mathscr{I}_\sigma} \bar{\varepsilon}_I\phi_I * g\Big\|_{L^{p'}(\mathbb{R}\times\Omega;Y^*)} \\
&\leqslant \sum_{\sigma\in\{-,+\}} 2\|n\|_{\mathfrak{M}(\mathbb{R};X,Y)} \min(\hbar_{p,X},\hbar_{p,Y})\Big\|\sum_{I\in\mathscr{I}_\sigma} \varepsilon_I \phi_I * f\Big\|_{L^p(\mathbb{R}\times\Omega;X)} \\
&\qquad\qquad \times \Big\|\sum_{I\in\mathscr{I}_\sigma} \bar{\varepsilon}_I\phi_I * g\Big\|_{L^{p'}(\mathbb{R}\times\Omega;Y^*)} \qquad \text{by Lemma 5.3.21} \\
&\leqslant \sum_{\sigma\in\{-,+\}} 2\|n\|_{\mathfrak{M}(\mathbb{R};X,Y)} \min(\hbar_{p,X},\hbar_{p,Y})\beta^+_{p,X}\|f\|_{L^p(\mathbb{R};X)} \\
&\qquad\qquad \times \beta^+_{p',Y^*}\|g\|_{L^{p'}(\mathbb{R};Y^*)} \qquad \text{by Lemma 5.3.22 (twice)} \\
&= 4\|n\|_{\mathfrak{M}(\mathbb{R};X,Y)} \min(\hbar_{p,X},\hbar_{p,Y})\beta^+_{p,X}\beta^+_{p',Y^*}\|f\|_{L^p(\mathbb{R};X)}\|g\|_{L^{p'}(\mathbb{R};Y^*)} \\
&\leqslant 4\|n\|_{\mathfrak{M}(\mathbb{R};X,Y)} \min(\hbar_{p,X},\hbar_{p,Y})\beta_{p,X}\beta_{p,Y}\|f\|_{L^p(\mathbb{R};X)}\|g\|_{L^{p'}(\mathbb{R};Y^*)},
\end{aligned}
$$

where we used that the conjugate sequence $(\bar{\varepsilon}_I)_{I\in\mathscr{I}}$ is a Rademacher sequence again.

It remains to observe, recalling $\widetilde{\phi}_I = \phi_I$ and hence $\widehat{\phi}_I = \widecheck{\phi}_I$, that

$$n = m\Big(\sum_{I\in\mathscr{I}} \frac{\mathbf{1}_I}{\widehat{\phi}_I}\Big)^2,$$

and therefore, by Proposition 5.3.20 and Lemma 5.3.23,

$$\|n\|_{\mathfrak{M}(\mathbb{R};X,Y)} \leqslant \|m\|_{\mathfrak{M}(\mathbb{R};X,Y)}\Big\|\sum_{I\in\mathscr{I}} \frac{\mathbf{1}_I}{\widehat{\phi}_I}\Big\|^2_{\mathfrak{M}(\mathbb{R})} \leqslant 9^2\|m\|_{\mathfrak{M}(\mathbb{R};X,Y)}.$$

The proof is completed by estimating $4\cdot 9^2 < 400$. $\square$

5.3.d Littlewood–Paley inequalities on $\mathbb{R}$

Specialising Mihlin's theorem to multipliers of the form $m_\epsilon = \sum_{I\in\mathscr{I}} \epsilon_I \mathbf{1}_I$, i.e., taking scalars $\epsilon_I \in \mathbb{K}$ of modulus one on each dyadic $I \in \mathscr{I}$, we find that

$$\|f\|_{L^p(\mathbb{R};X)} \lesssim \Big\| \sum_{I\in\mathscr{I}} \epsilon_I \Delta_I f \Big\|_{L^p(\mathbb{R};X)} \lesssim \|f\|_{L^p(\mathbb{R};X)},$$

where the "reverse" estimate on the left is obtained by applying $T_{m_{\bar\epsilon}}$ to $T_{m_\epsilon} f$, observing that $T_{m_{\bar\epsilon}} T_{m_\epsilon} = I$. Averaging the above estimate over $\epsilon_I \in S_{\mathbb{K}}$ (i.e., replacing them by the random ε_I), we obtain a two-sided estimate

$$\|f\|_{L^p(\mathbb{R};X)} \lesssim \Big\| \sum_{I\in\mathscr{I}} \varepsilon_I \Delta_I f \Big\|_{L^p(\mathbb{R}\times\Omega;X)} \lesssim \|f\|_{L^p(\mathbb{R};X)},$$

commonly known as the Littlewood–Paley inequality.

Aside from being an averaged version of inequalities for scalar-valued multiplier, the estimate on the right can also be seen as the boundedness of a single operator-valued multiplier $m \in L^\infty(\mathbb{R}; \mathscr{L}(X,Y))$, where $Y = L^p(\Omega;X)$ and

$$m(\xi) : x \mapsto \varepsilon_I \otimes x \in L^p(\Omega;X) \qquad \forall \xi \in I \in \mathscr{I}.$$

This point-of-view will become relevant later, when we transfer this inequality to a periodic setting in the proof of Theorem 5.7.10.

As for now, however, rather than viewing the Littlewood–Paley inequality as a corollary to Mihlin's theorem (as we just explained), we shall revisit parts of the actual proof, as this will lead to slightly better constants, recorded in the following:

Theorem 5.3.24 (Littlewood–Paley inequality on $\mathbb{R}$). *Let X be a UMD space and let $p \in (1,\infty)$, and let $(\varepsilon_I)_{I\in\mathscr{I}}$ be a Rademacher sequence indexed by $\mathscr{I}$ on a probability space $(\Omega,\mathbb{P})$. For all $f \in L^p(\mathbb{R};X)$ we have*

$$\frac{\|f\|_{L^p(\mathbb{R};X)}}{40\hbar_{p,X}\beta_{p,X}} \leqslant \Big\| \sum_{I\in\mathscr{I}} \varepsilon_I \Delta_I f \Big\|_{L^p(\mathbb{R}\times\Omega;X)} \leqslant 40\hbar_{p,X}\beta_{p,X}\|f\|_{L^p(\mathbb{R};X)}. \tag{5.39}$$

Proof. For $f \in (L^p \cap \check{L}^1)(\mathbb{R};X)$, we have

$$\Big(\sum_{I\in\mathscr{I}} \varepsilon_I \Delta_I f\Big)^{\widehat{}} = \sum_{I\in\mathscr{I}} \varepsilon_I \mathbf{1}_I \widehat{f} = \sum_{I\in\mathscr{I}} \varepsilon_I \frac{\mathbf{1}_I}{\widehat{\phi}_I} \widehat{\phi}_I \widehat{f} = \Big(\sum_{I\in\mathscr{I}} \varepsilon_I T_{\mathbf{1}_I m}(\phi_I * f)\Big)^{\widehat{}},$$

where $m = \displaystyle\sum_{I\in\mathscr{I}} \frac{\mathbf{1}_I}{\widehat{\phi}_I}$ is the multiplier from Lemma 5.3.23. Hence

$$\begin{aligned}
&\Big\| \sum_{I\in\mathscr{I}} \varepsilon_I \Delta_I f\Big\|_{L^p(\mathbb{R};X)} \\
&= \Big\| \sum_{I\in\mathscr{I}} \varepsilon_I T_{\mathbf{1}_I m}(\phi_I * f)\Big\|_{L^p(\mathbb{R};X)} \\
&\leqslant 2\|m\|_{\mathfrak{M}(\mathbb{R})} \hbar_{p,X} \Big\| \sum_{I\in\mathscr{I}} \varepsilon_I \phi_I * f\Big\|_{L^p(\mathbb{R};X)} && \text{by Lemma 5.3.21} \\
&\leqslant 2\cdot 9\cdot \hbar_{p,X} \sum_{\sigma\in\{+,-\}} \Big\| \sum_{I\in\mathscr{I}_\sigma} \varepsilon_I \phi_I * f\Big\|_{L^p(\mathbb{R};X)} && \text{by Lemma 5.3.23} \\
&\leqslant 2\cdot 9\hbar_{p,X}\cdot 2\beta^+_{p,X}\|f\|_{L^p(\mathbb{R};X)} && \text{by Lemma 5.3.22} \\
&\leqslant 36\hbar_{p,X}\beta_{p,X}\|f\|_{L^p(\mathbb{R};X)}.
\end{aligned} \tag{5.40}$$

The bound can be extended to all $f \in L^p(\mathbb{R}; X)$ by density.

For the reverse bound, we argue by duality that

$$\begin{aligned}
|\langle f, g\rangle| &= \Big|\mathbb{E}\Big\langle \sum_{I\in\mathscr{I}} \varepsilon_I \Delta_I f, \sum_{I\in\mathscr{I}} \bar{\varepsilon}_I \Delta_I g\Big\rangle\Big| \\
&\leqslant \Big\| \sum_{I\in\mathscr{I}} \varepsilon_I \Delta_I f\Big\|_{L^p(\Omega\times\mathbb{R};X)} \Big\| \sum_{I\in\mathscr{I}} \bar{\varepsilon}_I \Delta_I g\Big\|_{L^{p'}(\mathbb{R};X^*)} \\
&\leqslant \Big\| \sum_{I\in\mathscr{I}} \varepsilon_I \Delta_I f\Big\|_{L^p(\Omega\times\mathbb{R};X)} \cdot 36\hbar_{p',X^*}\beta^+_{p',X^*}\|g\|_{L^{p'}(\mathbb{R};X^*)},
\end{aligned}$$

where we used the bound already established for $L^{p'}(\mathbb{R}; X^*)$ instead, observing that the conjugate sequence $(\bar{\varepsilon}_I)_{I\in\mathscr{I}}$ is just another Rademacher sequence again. Recalling that $\hbar_{p',X^*} = \hbar_{p,X}$ and $\beta^+_{p',X^*} \leqslant \beta_{p',X^*} = \beta_{p,X}$ and taking the supremum over all $\|g\|_{L^{p'}(\mathbb{R};X^*)} \leqslant 1$, we have completed the proof. □

Sometimes it is useful to have the following variant, where the we group all "small frequencies" together.

Corollary 5.3.25. *Let X be a UMD space and let $p \in (1,\infty)$. For all $f \in L^p(\mathbb{R}; X)$ we have*

$$\frac{1}{40\hbar_{p,X}\beta_{p,X}}\|f\|_{L^p(\mathbb{R};X)} \leqslant \Big\| \sum_{I\in\widehat{\mathscr{I}}} \varepsilon_I \Delta_I f\Big\|_{L^p(\mathbb{R}\times\Omega;X)} \leqslant 40\hbar_{p,X}\beta_{p,X}\|f\|_{L^p(\mathbb{R};X)},$$

where

$$\widehat{\mathscr{I}} := \{(-1,1)\} \cup \{(-2^{k+1}, -2^k), (2^k, 2^{k+1}) : k \in \mathbb{N}\}.$$

Proof. Let $I_0 := (-1,1)$. By the triangle inequality and the contraction principle (observing that $\widehat{\mathscr{I}} \setminus \{I_0\} \subseteq \mathscr{I}$), we have

$$\begin{aligned}\Big\|\sum_{I\in\widehat{\mathscr{I}}}\varepsilon_I\Delta_I f\Big\|_{L^p(\mathbb{R}\times\Omega;X)} &\leqslant \Big\|\sum_{I\in\widehat{\mathscr{I}}\setminus\{I_0\}}\varepsilon_I\Delta_I f\Big\|_{L^p(\mathbb{R}\times\Omega;X)}+\|\Delta_{I_0}f\|_{L^p(\mathbb{R};X)}\\ &\leqslant \Big\|\sum_{I\in\mathscr{I}}\varepsilon_I\Delta_I f\Big\|_{L^p(\mathbb{R}\times\Omega;X)}+\|\Delta_{I_0}f\|_{L^p(\mathbb{R};X)}\\ &\leqslant 36\hbar_{p,X}\beta_{p,X}\|f\|_{L^p(\mathbb{R};X)}+\hbar_{p,X}\|f\|_{L^p(\mathbb{R};X)},\end{aligned}$$

where the last step used (5.40) from the proof of Theorem 5.3.24 for the first term, and Proposition 5.3.10 for the second. The proof of the right inequality is finished by observing that $1\leqslant\beta_{p,X}$ and $36+1<40$. The left one follows from this by a duality argument completely analogous to that in the proof of Theorem 5.3.24. □

Remark 5.3.26. Either version of the Littlewood–Paley inequality, or just their right side alone, is yet another characterisation of UMD spaces. Namely, they imply in particular (by the contraction principle, say) the $L^p(\mathbb{R};X)$-boundedness of the individual operators Δ_I, which is equivalent to the $L^p(\mathbb{R};X)$-boundedness of the Hilbert transform by Proposition 5.3.10.

In other situations, it is convenient to use a version of (5.39) involving smooth cut-off functions in place of the indicators $\mathbf{1}_I$. As a matter of fact, one could say that our preliminary "Littlewood–Paley inequality" recorded in Lemma 5.3.22 is an instance of such an estimate; in this case, the cut-offs $\widehat{\phi}_I$ are only weakly localised on $I\cup(-I)$ with "long" tails, decaying but supported on all $\mathbb{R}$. By a "smooth Littlewood–Paley inequality", one often understands a version somewhat intermediate between Lemma 5.3.22 and Theorem 5.3.24: involving smooth cut-off functions with "short" tails, localised in a neighbourhood of the intervals I. In order to avoid excessive repetition, we postpone the treatment of such estimates to Subsection 5.5.c, where we prove a version in several variables, covering the case of the line as a particular instance.

5.4 Applications to analysis in the Schatten classes

We interrupt the discussion of the general theory of Fourier multipliers on UMD spaces by illustrating one of the many beautiful applications that can already be handled by the machinery that we have developed up to this point. We have in mind some results dealing with the Schatten classes $\mathscr{C}^p$, whose basic theory is recalled in Appendix D. A reader who prefers to go on with the multiplier theory in several variables without this interruption may very well do so, as all subsequent sections are independent of the one at hand.

5.4.a The UMD property of the Schatten classes

Our point of view on the Schatten classes is biased towards aspects that depend on (or at least can be understood via) the fact that these spaces have the UMD property. Thus, the natural starting point is to establish this property. For this goal we shall exploit Theorem 5.2.10, which says that it is sufficient to prove the boundedness of the periodic Hilbert transform $\widetilde{H}$ on $L^p_0(\mathbb{T};\mathscr{C}^p)$. (Note in particular that we exploit the freedom of considering the same exponent p in the L^p space over $\mathbb{T}$ as that appearing in the space $\mathscr{C}^p$ that we are interested in.) The proof is essentially based on the original approach of M. Riesz to prove that $\widetilde{H} \in \mathscr{L}(L^p(\mathbb{T}))$, and gives us the opportunity to revisit this classical argument in a modern framework.

From the multiplier representation of $\widetilde{H}$ it is immediate that $f + i\widetilde{H}f$ has non-zero Fourier-coefficients for non-negative frequencies only. This is to some extent characteristic of the Hilbert transform:

Lemma 5.4.1. *For $u \in L^p(\mathbb{T};\mathbb{R})$, $p \in (1,\infty)$, its Hilbert transform is the unique $v \in L^p_0(\mathbb{T};\mathbb{R})$ such that $u + iv$ has only non-negative frequencies.*

Proof. We already checked that the Hilbert transform has the asserted property, as it follows from $u + iv = (I + i\widetilde{H})v$, so it remains to be shown that it is the unique such function. For a real-valued function, the Fourier coefficients satisfy $\widehat{u}(-k) = \overline{\widehat{u}(k)}$. By assumption, for all $k \geqslant 1$ there holds

$$0 = \widehat{u}(-k) + i\widehat{v}(-k) = \overline{\widehat{u}(k)} + i\overline{\widehat{v}(k)} = \overline{\widehat{u}(k) - i\widehat{v}(k)},$$

and hence

$$\widehat{v}(k) = -i\,\mathrm{sgn}(k)\widehat{u}(k) \tag{5.41}$$

for all $k \neq 0$. Since further $\widehat{v}(0) = \int_{\mathbb{T}} v \,\mathrm{d}x = 0 = \mathrm{sgn}(0)$ according to our convention, (5.41) holds for all $k \in \mathbb{Z}$, proving that $v = \widetilde{H}u$. ☐

This lemma leads to a “magical” identity concerning the Hilbert transforms of products of functions. Let u and v be real-valued trigonometric polynomials with vanishing integral. Then the function

$$\begin{aligned}-i(v + i\widetilde{H}v)(u + i\widetilde{H}u) &= -i[(v \cdot u - \widetilde{H}v \cdot \widetilde{H}u) + i(v \cdot \widetilde{H}u + \widetilde{H}v \cdot u)] \\ &= (v \cdot \widetilde{H}u + \widetilde{H}v \cdot u) + i(\widetilde{H}v \cdot \widetilde{H}u - v \cdot u)\end{aligned}$$

has only non-negative frequencies, being a product of functions with this property. Moreover, the imaginary part of this function has mean zero, since

$$\int_{\mathbb{T}} \widetilde{H}v \cdot \widetilde{H}u \,\mathrm{d}x = -\int_{\mathbb{T}} \widetilde{H}^2 v \cdot u \,\mathrm{d}x = \int_{\mathbb{T}} v \cdot u \,\mathrm{d}x,$$

using the skew-adjointness of $\widetilde{H}$ and fact that $\widetilde{H}^2 = -I$ when restricted to $L^p_0(\mathbb{T})$. Thus, the uniqueness part of Lemma 5.4.1 implies *Cotlar's identity*

$$\widetilde{H}(v \cdot \widetilde{H}u + \widetilde{H}v \cdot u) = \widetilde{H}v \cdot \widetilde{H}u - v \cdot u$$

for all real-valued trigonometric polynomials with vanishing integral. But this expression is linear with respect to both u and v, so that it also extends to complex-valued functions. And finally, since the vector-valued Hilbert transform is the tensor extension of the scalar-valued one, the equation even holds for vector-valued trigonometric polynomials with vanishing integrals, whenever u and v take values in some Banach spaces X and Y, for which some bilinear product $\cdot : X \times Y \to Z$, with Z another Banach space, is defined. (Notice that the Hilbert transform of an X-valued trigonometric polynomial always exist, whether or not $\widetilde{H}$ acts boundedly on the full space $L^p_0(\mathbb{T};X)$.)

With this extended scope of validity, we repeat the magic identity with some terms moved over to the other side:

$$\widetilde{H}v \cdot \widetilde{H}u = v \cdot u + \widetilde{H}(v \cdot \widetilde{H}u + \widetilde{H}v \cdot u). \tag{5.42}$$

Proposition 5.4.2. *The Schatten spaces $\mathscr{C}^p$, $1 < p < \infty$, are UMD. In fact,*

$$\|\widetilde{H}\|_{\mathscr{L}(L^{2^n}_0(\mathbb{T};\mathscr{C}^{2^n}))} \leqslant \cot\frac{\pi}{2^{n+1}}, \quad n = 1, 2, 3, \dots$$

$$\|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};\mathscr{C}^p))} \leqslant 2\cot\frac{\pi}{2\cdot 2^{\lceil \log_2 p^*\rceil}} \leqslant \frac{8}{\pi}p^*, \quad 1 < p < \infty, \quad p^* := \max(p, p').$$

Actually, a factor 2 in the second bound is redundant, as we shall see in Corollary 5.7.8. Note that the last estimate in this bound simply follows from $\cot x \leqslant 1/x$ for $x \in (0, \pi/2)$ and $2^{\lceil \log_2 p^*\rceil} \leqslant 2p^*$.

Proof. Let us abbreviate $A_p := \|\widetilde{H}\|_{\mathscr{L}(L^p_0(\mathbb{T};\mathscr{C}^p))}$. Since $\mathscr{C}^2$ is a Hilbert space, we have $A_2 = 1$.

Suppose first that the Hilbert transform acts boundedly in some $L^p_0(\mathbb{T};\mathscr{C}^p)$, and let us see how it acts in $L^{2p}_0(\mathbb{T};\mathscr{C}^{2p})$. We first consider the subspace of $L^{2p}_0(\mathbb{T};\mathscr{C}^{2p})$ consisting of trigonometric polynomials of degree at most N, for some fixed N. It is easily seen that the Hilbert transform is bounded on this space, with a norm estimate *a priori* depending on N. Let x denote its (unknown) norm.

We apply (5.42) to the functions u from our test function space, taking $v(y) = u(y)^*$, its pointwise adjoint (remembering that $\mathscr{C}^p$ consists of bounded operators acting on some underlying Hilbert space). Since, by (D.6),

$$\|a\|_{\mathscr{C}^{2p}} = \|a^*\|_{\mathscr{C}^{2p}} = \|a^*a\|^{1/2}_{\mathscr{C}^p}$$

and $(\widetilde{H}u)^*(y) = \widetilde{H}(u^*)(y)$, it follows from (5.42) that

$$\begin{aligned}
&\|\widetilde{H}u\|^2_{L^{2p}_0(\mathbb{T};\mathscr{C}^{2p})} = \|\widetilde{H}(u^*)\widetilde{H}u\|_{L^p(\mathbb{T};\mathscr{C}^p)} \\
&\leqslant \|u^*u\|_{L^p(\mathbb{R};\mathscr{C}^p)} + A_p\big(\|u^* \cdot \widetilde{H}u\|_{L^p(\mathbb{R};\mathscr{C}^p)} + \|\widetilde{H}u^* \cdot u\|_{L^p(\mathbb{R};\mathscr{C}^p)}\big) \\
&\leqslant \|u\|^2_{L^{2p}_0(\mathbb{R};\mathscr{C}^p)} + 2A_p\|u\|_{L^{2p}(\mathbb{R};\mathscr{C}^{2p})}\|\widetilde{H}u\|_{L^{2p}(\mathbb{R};\mathscr{C}^{2p})} \\
&\leqslant \|u\|^2_{L^{2p}_0(\mathbb{R};\mathscr{C}^{2p})} + 2A_p x\|u\|^2_{L^{2p}_0(\mathbb{R};\mathscr{C}^{2p})}.
\end{aligned}$$

Taking the supremum over test functions with $\|u\|_{L^{2p}_0(\mathbb{R};\mathscr{C}^{2p})} \leqslant 1$, it follows that

$$x^2 \leqslant 1 + 2A_p x \quad \Rightarrow \quad x \leqslant A_p + \sqrt{A_p^2 + 1}.$$

This estimate is independent of our test function space, so that by density

$$A_{2p} \leqslant A_p + \sqrt{A_p^2 + 1}.$$

From the cotangent functional equation

$$\cot \frac{t}{2} = \cot t + \sqrt{\cot^2 t + 1}, \qquad t \in (0, \tfrac{1}{2}\pi),$$

and the initial condition $A_2 = 1 = \cot(\pi/4)$, it then follows by induction that

$$\|\widetilde{H}\|_{\mathscr{L}(L_0^{2^n}(\mathbb{T};\mathscr{C}^{2^n}))} = A_{2^n} \leqslant \cot \frac{\pi}{2^{n+1}}, \qquad n = 1, 2, \ldots,$$

and thus

$$\|\widetilde{H}\|_{\mathscr{L}(L^{2^n}(\mathbb{T};\mathscr{C}^{2^n}))} \leqslant 2A_{2^n} \leqslant 2\cot \frac{\pi}{2^{n+1}}, \qquad n = 1, 2, \ldots.$$

For $p \in (2^{n-1}, 2^n)$ we have, by Theorem C.2.6 and Proposition D.3.1,

$$L^p(\mathbb{T};\mathscr{C}^p) = [L^{2^{n-1}}(\mathbb{T};\mathscr{C}^{2^{n-1}}), L^{2^n}(\mathbb{T};\mathscr{C}^{2^n})]_\sigma$$

for $\sigma \in (0,1)$ determined by the condition $1/p = (1-\sigma)/2^{n-1} + \sigma/2^n$. Hence

$$\|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};\mathscr{C}^p))} \leqslant \|\widetilde{H}\|^{1-\sigma}_{\mathscr{L}(L^{2^{n-1}}(\mathbb{T};\mathscr{C}^{2^{n-1}}))} \|\widetilde{H}\|^{\sigma}_{\mathscr{L}(L^{2^n}(\mathbb{T};\mathscr{C}^{2^n}))} \leqslant 2\cot \frac{\pi}{2^{n+1}}.$$

This establishes the result for $p \in (2, \infty)$. The remaining case $p \in (1, 2)$ is handled by duality:

$$\|\widetilde{H}\|_{\mathscr{L}(L^{p'}(\mathbb{T};\mathscr{C}^{p'}))} = \|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};\mathscr{C}^p))},$$

as $(L^p(\mathbb{T};\mathscr{C}^p))^* = L^{p'}(\mathbb{T};\mathscr{C}^{p'})$ and $\widetilde{H}^* = -\widetilde{H}$. □

The previous proposition as such gives information about a singular integral operator acting on functions with values in $\mathscr{C}^p$. In the following section, we give some examples of how such estimates may be transferred to the analysis of operators on the space $\mathscr{C}^p$ itself.

5.4.b Schur multipliers and transference on $\mathscr{C}^p$

The *Schur product* of two matrices (m_{ij}) and (v_{ij}) is the entry-wise product $(m_{ij}v_{ij})$. Considering one of the matrices fixed and the other one variable, we obtain an operator $(v_{ij}) \mapsto (m_{ij}v_{ij})$ known as a *Schur multiplier*. This notion has the following extension in the realm of operators on an infinite-dimensional Hilbert space H.

We say that $(e_\lambda)_{\lambda\in\Lambda}$ is a *countable spectral resolution* of H if $\Lambda \subseteq \mathbb{R}$ is a countable set and:

(i) each e_λ is an orthogonal projection in H;
(ii) the e_λ are pairwise orthogonal, i.e., $e_\lambda e_\mu = 0$ for $\lambda \neq \mu$;
(iii) $\sum_{\lambda \in \Lambda} e_\lambda x = x$ for all $x \in H$.

The convergence of an orthogonal series is always unconditional, so that the last condition is well defined without specifying a particular order on the countable family Λ. A countable spectral resolution is nothing but an orthogonal Schauder decomposition of H, but in the present context its seems more natural to adhere to the present terminology, which is borrowed from the spectral theory of self-adjoint operators.

Given a countable spectral resolution $(e_\lambda)_{\lambda \in \Lambda}$ and a matrix $(m_{\lambda,\mu})_{\lambda,\mu \in \Lambda}$ with the same index set, the related Schur multiplier on operators of $v \in \mathscr{L}(H)$ is defined by the formal series

$$\sum_{\lambda,\mu \in \Lambda} m_{\lambda,\mu} e_\lambda v e_\mu .$$

We mainly restrict our attention to Schur multipliers of a particular form, where $m_{\lambda,\mu} = m(\lambda - \mu)$ depends on the difference of the two indices only.

Theorem 5.4.3. *Let $m \in C(\mathbb{R} \setminus \{0\})$, and $(e_\lambda)_{\lambda \in \Lambda}$ be a countable spectral resolution. If $T_m \in \mathscr{L}(L^2(\mathbb{R}; \mathscr{C}^p))$, then the series*

$$M_m^e v := \sum_{\substack{\lambda,\mu \in \Lambda \\ \lambda \neq \mu}} m(\lambda - \mu) e_\lambda v e_\mu := \lim_{n \to \infty} \sum_{\substack{\lambda,\mu \in \Lambda_n \\ \lambda \neq \mu}} m(\lambda - \mu) e_\lambda v e_\mu$$

converges in $\mathscr{C}^p$ for each $v \in \mathscr{C}^p$, where the last limit is along any sequence of finite subsets $\Lambda_n \uparrow \Lambda$, and satisfies

$$\|M_m^e\|_{\mathscr{L}(\mathscr{C}^p)} \leqslant 2 \|T_m\|_{\mathscr{L}(L^2(\mathbb{R};\mathscr{C}^p))}.$$

Remark 5.4.4. The sum over the diagonal $\lambda = \mu \in \Lambda$ can be estimated by more elementary means: We have

$$\Big\| \sum_{\lambda \in \Lambda} a_\lambda e_\lambda v e_\lambda \Big\|_{\mathscr{C}^p} \leqslant \|a\|_\infty \|v\|_{\mathscr{C}^p}, \qquad \forall a = (a_\lambda)_{\lambda \in \Lambda} \in \ell^\infty(\Lambda),$$

which applies in particular with $a_\lambda = m(\lambda - \lambda) \equiv m(0)$. To see this, just observe that (say, in the sense of weak operator convergence, using the unconditional convergence $\sum_{\lambda \in \Lambda} e_\lambda x = x$ on both sides of the operator)

$$\sum_{\lambda \in \Lambda} a_\lambda e_\lambda v e_\lambda = \mathbb{E}\Big(\sum_{\lambda \in \Lambda} \varepsilon_\lambda a_\lambda e_\lambda \Big) v \Big(\sum_{\mu \in \Lambda} \bar{\varepsilon}_\mu e_\mu \Big),$$

so that, by the ideal property (D.5) and an elementary estimate for diagonal operators on a Hilbert space,

$$\begin{aligned} \Big\| \sum_{\lambda \in \Lambda} a_\lambda e_\lambda v e_\lambda \Big\|_{\mathscr{C}^p} &\leqslant \mathbb{E} \Big\| \sum_{\lambda \in \Lambda} \varepsilon_\lambda a_\lambda e_\lambda \Big\|_{\mathscr{L}(H)} \|v\|_{\mathscr{C}^p} \Big\| \sum_{\mu \in \Lambda} \bar{\varepsilon}_\mu e_\mu \Big\|_{\mathscr{L}(H)} \\ &\leqslant \|a\|_\infty \cdot \|v\|_{\mathscr{C}^p} \cdot 1. \end{aligned}$$

Before going into the proof of Theorem 5.4.3, we record some interesting consequences.

Corollary 5.4.5 (Macaev's theorem on triangular truncations). *For any countable spectral resolution* $(e_\lambda)_{\lambda\in\Lambda}$*, the associated triangular truncations*

$$v \mapsto \sum_{\substack{\lambda,\mu\in\Lambda\\ \lambda>\mu}} e_\lambda v e_\mu \qquad \text{and} \qquad v \mapsto \sum_{\substack{\lambda,\mu\in\Lambda\\ \lambda\geqslant\mu}} e_\lambda v e_\mu$$

are bounded on $\mathscr{C}^p$ *for all* $p \in (1,\infty)$*, with operator norms bounded by a quantity depending only on* p*.*

Proof. The first case follows from Theorem 5.4.3 applied to $m = \mathbf{1}_{(0,\infty)}$, for which the boundedness of T_m is equivalent to the boundedness of the Hilbert transform on $L^2(\mathbb{R};\mathscr{C}^p)$. The second case is obtained by combining this with Remark 5.4.4. □

More generally, in combination with Mihlin's Multiplier Theorem 5.3.18 and its Corollary 5.3.19, we obtain:

Corollary 5.4.6. *For* m *and* e *as in Theorem 5.4.3, if in addition* m *belongs to the Mihlin class, then*

$$\|M_m^e\|_{\mathscr{L}(\mathscr{C}^p)} \leqslant c_p\|m\|_{\mathfrak{M}(\mathbb{R})},$$

and in particular

$$\|M_{m_s}^e\|_{\mathscr{L}(\mathscr{C}^p)} \leqslant c_p(1+|s|), \qquad m_s(\xi) = |\xi|^{is}, \quad s\in\mathbb{R}.$$

Let us now turn to:

Proof of Theorem 5.4.3. In order to avoid repeated writing of the condition $\lambda \neq \mu$, we may set $m(0) := 0$.

We first observe that, if Λ_n is a sequence of finite subsets of Λ which increases towards Λ, then $v_n := \sum_{\lambda,\mu\in\Lambda_n} e_\lambda v e_\mu \to v$ in $\mathscr{C}^p$. Namely, if $x \in H$ (the underlying Hilbert space), it is clear that $\sum_{\lambda\in\Lambda_n} e_\lambda x \to x$ in H, and from this it easily follows that $v_n \to v$ in $\mathscr{C}^p$ for any finite-rank v, just by writing out its finite expansion in terms of elementary tensors. Since finite-rank operators are dense in $\mathscr{C}^p$, the convergence easily extends to a general $v \in \mathscr{C}^p$ by approximation.

Now, it suffices to prove the theorem for v_n in place of v. Namely, from this it follows that the finite sums

$$M_m^e v_n = \sum_{\lambda,\mu\in\Lambda_n} m(\lambda-\mu) e_\lambda v e_\mu$$

form a Cauchy sequence, and we may define $M_m^e v$ as their limit, which is easily seen to be independent of the particular approximating sequence. Thus we concentrate on such finite sums in the sequel, but drop the subscript n to simplify writing.

The key trick behind the transference is to pre- and post-compose the operator $M^e_m v$ with a unitary $u_t := \sum_{\lambda\in\Lambda} e^{2\pi i\lambda t} e_\lambda$ (leaving the $\mathscr{C}^p$ norm invariant), to the result that

$$\begin{aligned}
\|M^e_m v\|_{\mathscr{C}^p} &= \|u_t(M^e_m v)u_{-t}\|_{\mathscr{C}^p} \\
&= \Big\| \sum_{\lambda,\mu\in\Lambda} m(\lambda-\mu)e^{2\pi i(\lambda-\mu)t} e_\lambda v e_\mu \Big\|_{\mathscr{C}^p} \qquad \forall t\in\mathbb{R} \\
&= \Big\| \sum_{\theta\in\Delta:=(\Lambda-\Lambda)\setminus\{0\}} m(\theta)e^{2\pi i\theta t} v_\theta \Big\|_{\mathscr{C}^p}, \qquad v_\theta := \sum_{\substack{\lambda,\mu\in\Lambda\\ \lambda-\mu=\theta}} e_\lambda v e_\mu \in \mathscr{C}^p,
\end{aligned}$$

where the convergence of the sums is trivial by their finiteness. (Alternatively, one can check that $\|v_\theta\|_{\mathscr{C}^p} \leqslant \|v\|_{\mathscr{C}^p}$ by an argument analogous to that in Remark 5.4.4.) Note that the same computation with $m\equiv 1$ shows that

$$\|v\|_{\mathscr{C}^p} = \Big\| v_0 + \sum_{\theta\in\Delta} e^{2\pi i\theta t} v_\theta \Big\|_{\mathscr{C}^p} \qquad \forall t\in\mathbb{R},$$

where v_0 is defined like the other v_θ, setting $\theta=0$.

Let us pick an auxiliary function $\phi\in L^2(\mathbb{R})$ of unit norm. Then

$$\|M^e_m v\|^2_{\mathscr{C}^p} = \int_{\mathbb{R}} \|M^e_m v\|^2_{\mathscr{C}^p}\cdot\varepsilon|\phi(\varepsilon t)|^2\,\mathrm{d}t = \int_{\mathbb{R}} \Big\|\sum_{\theta\in\Delta} m(\theta)e^{2\pi i\theta t} v_\theta\phi(\varepsilon t)\Big\|^2_{\mathscr{C}^p}\varepsilon\,\mathrm{d}t.$$

If we replace $m(\theta)$ by the Fourier multiplier operator T_m, the resulting expression is immediately bounded by

$$\begin{aligned}
&\int_{\mathbb{R}} \Big\|\sum_{\theta\in\Delta} T_m\Big(t\mapsto e^{2\pi i\theta t}v_\theta\phi(\varepsilon t)\Big)\Big\|^2_{\mathscr{C}^p}\varepsilon\,\mathrm{d}t \\
&\qquad\leqslant \|T_m\|^2_{\mathscr{L}(L^2(\mathbb{R};\mathscr{C}^p))}\int_{\mathbb{R}}\Big\|\sum_{\theta\in\Delta} e^{2\pi i\theta t}v_\theta\phi(\varepsilon t)\Big\|^2_{\mathscr{C}^p}\varepsilon\,\mathrm{d}t \\
&\qquad= \|T_m\|^2_{\mathscr{L}(L^2(\mathbb{R};\mathscr{C}^p))}\int_{\mathbb{R}}\Big\|\sum_{\theta\in\Delta} e^{2\pi i\theta t}v_\theta\Big\|^2_{\mathscr{C}^p}\varepsilon|\phi(\varepsilon t)|^2\,\mathrm{d}t \\
&\qquad\leqslant \|T_m\|^2_{\mathscr{L}(L^2(\mathbb{R};\mathscr{C}^p))}\int_{\mathbb{R}}(\|v\|_{\mathscr{C}^p}+\|v_0\|_{\mathscr{C}^p})^2\varepsilon|\phi(\varepsilon t)|^2\,\mathrm{d}t \\
&\qquad\leqslant \|T_m\|^2_{\mathscr{L}(L^2(\mathbb{R};\mathscr{C}^p))}\cdot 2^2\|v\|^2_{\mathscr{C}^p},
\end{aligned}$$

observing in the last step that v_0 is the diagonal projection considered in Remark 5.4.4.

Hence it only remains to consider the error term

$$\Big(\int_{\mathbb{R}}\Big\|\sum_{\theta\in\Delta}(m(\theta)-T_m)\Big(t\mapsto e^{2\pi i\theta t}v_\theta\phi(\varepsilon t)\Big)\Big\|^2_{\mathscr{C}^p}\varepsilon\,\mathrm{d}t\Big)^{1/2}$$

$$\leqslant \sum_{\theta\in\Delta}\Big(\int_{\mathbb{R}}|(m(\theta)-T_m)(t\mapsto e^{2\pi i\theta t}\phi(\varepsilon t))|^2\varepsilon\,\mathrm{d}t\Big)^{1/2}\|v_\theta\|_{\mathscr{C}^p}$$
$$=\sum_{\theta\in\Delta}\Big(\int_{\mathbb{R}}\Big|(m(\theta)-m(\xi))\widehat{\phi}\Big(\frac{\xi-\theta}{\varepsilon}\Big)\Big|^2\frac{\mathrm{d}\xi}{\varepsilon}\Big)^{1/2}\|v_\theta\|_{\mathscr{C}^p},$$

where, after extracting the operator component v_θ from the norm, Plancherel's theorem for scalar-valued L^2 functions became available in the last step.

With the change of variable $\eta := (\xi-\theta)/\varepsilon$, we find that

$$\int_{\mathbb{R}}\Big|(m(\theta)-m(\xi))\widehat{\phi}\Big(\frac{\xi-\theta}{\varepsilon}\Big)\Big|^2\frac{\mathrm{d}\xi}{\varepsilon}=\int_{\mathbb{R}}|m(\theta)-m(\theta+\varepsilon\eta)|^2|\widehat{\phi}(\eta)|^2\,\mathrm{d}\eta.$$

Since m is continuous at $\theta\in\Delta=(\Lambda-\Lambda)\setminus\{0\}$, the first factor tends to zero with ε at every fixed $\eta\in\mathbb{R}$. It is also dominated uniformly by $(2\|m\|_\infty)^2$, and hence the whole integral converges to zero by dominated convergence, recalling that $\widehat{\phi}\in L^2(\mathbb{R})$. Since the summation set $\Delta=(\Lambda-\Lambda)\setminus\{0\}$ is finite and each $\|v_\theta\|_{\mathscr{C}^p}<\infty$, the entire error term tends to zero, and this completes the proof. □

5.4.c Operator Lipschitz functions

The following theorem is one of the highlights of the applications of the vector-valued Littlewood–Paley theory to the Schatten spaces:

Theorem 5.4.7 (Potapov–Sukochev). *Let u, v be compact self-adjoint operators on a Hilbert space H, and $f:\mathbb{R}\to\mathbb{R}$ be a Lipschitz function. Then*

$$\|f(u)-f(v)\|_{\mathscr{C}^p}\leqslant c_p\|f\|_{\mathrm{Lip}}\|u-v\|_{\mathscr{C}^p},\qquad 1<p<\infty,$$

where

$$\|f\|_{\mathrm{Lip}}:=\sup_{\substack{\lambda,\mu\in\mathbb{R}\\ \lambda\neq\mu}}\frac{|f(\lambda)-f(\mu)|}{|\lambda-\mu|}.$$

Actually, the word "compact" may be dropped from the assumptions: the theorem is valid for possibly unbounded self-adjoint operators, as long as their difference lies in the Schatten class $\mathscr{C}^p$. In the general case, $f(u)$ and $f(v)$ are defined via the continuous spectral resolution

$$u=\int_{\mathbb{R}}\lambda\,\mathrm{d}e_\lambda,\qquad f(u)=\int_{\mathbb{R}}f(\lambda)\,\mathrm{d}e_\lambda,$$

and the analysis of the difference $f(u)-f(v)$ depends on the theory of so-called *double operator integrals*. We will content ourselves with the more restricted formulation involving the countable spectral resolution $(e_\lambda)_{\lambda\in\Lambda}$ of compact operators,

$$u = \sum_{\lambda\in\Lambda} \lambda e_\lambda, \qquad f(u) = \sum_{\lambda\in\Lambda} f(\lambda) e_\lambda.$$

This case is already interesting and non-trivial, as well as representative of the use of the vector-valued Littlewood–Paley theory to this type of problems.

The approach to Theorem 5.4.7 proceeds via the following commutator estimate:

Proposition 5.4.8. *Let the operators u and v be compact self-adjoint and bounded, respectively, and let $f : \mathbb{R} \to \mathbb{R}$ be a Lipschitz function. Then*

$$\|[f(u), v]\|_{\mathscr{C}^p} \leqslant c_p \|f\|_{\mathrm{Lip}} \|[u, v]\|_{\mathscr{C}^p}.$$

Proof of Theorem 5.4.7 assuming Proposition 5.4.8. Given compact and self-adjoint a, b on H, consider

$$u = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \qquad v = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$$

on $H \times H$. Then

$$\begin{aligned} \|[f(u), v]\|_{\mathscr{C}^p(H)} &= \Big\| \begin{bmatrix} 0 & f(a) - f(b) \\ f(b) - f(a) & 0 \end{bmatrix} \Big\|_{\mathscr{C}^p(H \times H)} \\ &= 2^{1/p} \|f(a) - f(b)\|_{\mathscr{C}^p(H)}, \end{aligned}$$

where we used the general identity that

$$\|\widetilde{s}\|_{\mathscr{C}^p(H\times H)} = 2^{1/p} \|s\|_{\mathscr{C}^p(H)}, \qquad \widetilde{s} := \begin{bmatrix} 0 & s \\ s & 0 \end{bmatrix},$$

since each singular value $\tau_k = a_k(s)$ of s (cf. (D.1) and Lemma D.1.1) appears twice in the singular value decomposition of $\widetilde{s}$, and hence

$$\|\widetilde{s}\|_{\mathscr{C}^p(H\times H)} = \Big(\sum_{k=1}^{\infty} a_k(\widetilde{s})^p \Big)^{1/p} = \Big(2 \sum_{k=1}^{\infty} a_k(s)^p \Big)^{1/p} = 2^{1/p} \|s\|_{\mathscr{C}^p(H)}.$$

Similarly we have $\|[u, v]\|_{\mathscr{C}^p} = 2^{1/p} \|a - b\|_{\mathscr{C}^p}$, so it is immediate that the Theorem follows from the Proposition. $\square$

Proof of Proposition 5.4.8. Since u is compact and self-adjoint, there is a countable spectral resolution $(e_\lambda)_{\lambda\in\Lambda}$ such that $u = \sum_{\lambda\in\Lambda} \lambda e_\lambda$. Since $f(u) = \sum_{\lambda\in\Lambda} f(\lambda) e_\lambda$ and $I = \sum_{\lambda\in\Lambda} e_\lambda$,

$$\begin{aligned} [f(u), v] &= \sum_{\lambda\in\Lambda} f(\lambda) e_\lambda \sum_{\mu\in\Lambda} v e_\mu - \sum_{\lambda\in\Lambda} e_\lambda v \sum_{\mu\in\Lambda} f(\mu) e_\mu \\ &= \sum_{\lambda,\mu\in\Lambda} (f(\lambda) - f(\mu)) e_\lambda v e_\mu, \end{aligned}$$

and likewise $[u,v]=\sum_{\lambda,\mu\in\Lambda}(\lambda-\mu)e_\lambda v e_\mu$. Thus

$$[f(u),v]=\Phi_f([u,v]),$$

where

$$\Phi_f w=\sum_{\substack{\lambda,\mu\in\Lambda\\ \lambda\neq\mu}}\phi_f(\lambda,\mu)e_\lambda w e_\mu,\qquad \phi_f(\lambda,\mu)=\frac{f(\lambda)-f(\mu)}{\lambda-\mu}$$

is a Schur multiplier, although not quite of the special form considered in Theorem 5.4.3. The core of the proof is the decomposition of this Schur multiplier into new ones of the form that we can estimate.

Before proceeding, we observe a useful reduction. Since both the left and the right side of our inequality behave in the same way under the replacement of f by $\alpha f+\beta$, we may without loss of generality assume that $\|f\|_{\mathrm{Lip}}=1$ and $f(0)=0$. Let us then consider $\widetilde{f}(t):=f(t)+2t$, which has $\|\widetilde{f}\|_{\mathrm{Lip}}\leqslant 3$ and the slope of $\widetilde{f}$ (i.e., the value of $\phi_{\widetilde{f}}$) stays between 1 and 3. If we can prove the proposition for $\widetilde{f}$, it also follows for f (with another constant), since $[f(u),v]=[\widetilde{f}(u)-2u,v]=[\widetilde{f}(u),v]-2[u,v]$. Note that $\widetilde{f}(0)=f(0)=0$. So we consider from this point on a function $\widetilde{f}$ with $\widetilde{f}(0)=0$ and slope between 1 and 3, but denote it simply by f.

For $x\in[1,3]$, we have $x=e^{2\pi t}$ with $t\in[0,(\log 3)/2\pi]\subseteq[0,1]$. Now pick some $\psi\in\mathscr{S}(\mathbb{R})$ that agrees with $e^{2\pi t}$ for $t\in[0,1]$. Thus, by the Fourier inversion formula,

$$x=e^{2\pi t}=\psi(t)=\int_{\mathbb{R}}\widehat{\psi}(s)e^{2\pi i s t}\,\mathrm{d}s=\int_{\mathbb{R}}\widehat{\psi}(s)x^{is}\,\mathrm{d}s,\qquad x\in[1,3].$$

Substituting $x=(f(\lambda)-f(\mu))/(\lambda-\mu)\in[1,3]$, we find that

$$\begin{aligned}\Phi_f w&=\int_{\mathbb{R}}\widehat{\psi}(s)\Big(\sum_{\substack{\lambda,\mu\in\Lambda\\ \lambda\neq\mu}}|f(\lambda)-f(\mu)|^{is}|\lambda-\mu|^{-is}e_\lambda w e_\mu\Big)\,\mathrm{d}s\\ &=:\int_{\mathbb{R}}\widehat{\psi}(s)M^{e,f}_{m_s}M^e_{m_{-s}}w\,\mathrm{d}s,\end{aligned}$$

where $M^e_{m_{-s}}$ (with $m_{-s}(\xi)=|\xi|^{-is}$) is an operator as in Theorem 5.4.3, and

$$M^{e,f}_m w:=\sum_{\substack{\lambda,\mu\in\Lambda\\ \lambda\neq\mu}}m(f(\lambda)-f(\mu))e_\lambda w e_\mu=\sum_{\substack{\alpha,\beta\in f(\Lambda)\\ \alpha\neq\beta}}m(\alpha-\beta)e_{f^{-1}(\alpha)}w e_{f^{-1}(\beta)}$$

is exactly of the same form with another (differently parametrised) spectral resolution $(e_{f^{-1}(\alpha)})_{\alpha\in f(\Lambda)}$. By Corollary 5.4.6, it follows that

$$\|\Phi_f w\|_{\mathscr{C}^p}\leqslant\int_{\mathbb{R}}|\widehat{\psi}(s)|\|M^{e,f}_{m_s}\|_{\mathscr{L}(\mathscr{C}^p)}\|M^e_{m_{-s}}\|_{\mathscr{L}(\mathscr{C}^p)}\|w\|_{\mathscr{C}^p}\,\mathrm{d}s$$

$$\leqslant c\int_{\mathbb{R}} |\widehat{\psi}(s)|\|T_{m_s}\|_{\mathscr{L}(L^2(\mathbb{R};\mathscr{C}^p)}\|T_{m_{-s}}\|_{\mathscr{L}(L^2(\mathbb{R};\mathscr{C}^p))}\|w\|_{\mathscr{C}^p}\,\mathrm{d}s$$
$$\leqslant c\int_{\mathbb{R}} |\widehat{\psi}(s)|\cdot c_p(1+|s|)\cdot c_p(1+|s|)\cdot\|w\|_{\mathscr{C}^p}\,\mathrm{d}s \leqslant c_p\|w\|_{\mathscr{C}^p},$$

observing the finiteness of the integral $\int_{\mathbb{R}} |\widehat{\psi}(s)|(1+|s|)^2\,\mathrm{d}s$ for $\psi\in\mathscr{S}(\mathbb{R})$ in the last step. □

5.5 Fourier multipliers on $\mathbb{R}^d$

We now return to the development of the multiplier theory in several variables. The foundations of this theory were already set up in Subsection 5.3.a, and the present aim is to obtain several results guaranteeing the $L^p(\mathbb{R}^d;X)$-boundedness of important (classes of) multipliers when X is a UMD space. We begin with some relatively simple constructions based on "lifting" the one-dimensional results that we already have, and then proceed to a full analogue of Mihlin's multiplier theorem and the Littlewood–Paley inequalities.

5.5.a Riesz transforms and other liftings from $\mathbb{R}$

We begin with a concrete example of multipliers on $\mathbb{R}^d$ that appear in multiple applications. For $1\leqslant j\leqslant n$, let the jth *Riesz transform* R_j be the Fourier multiplier with symbol $-i\frac{\xi_j}{|\xi|}$ for $\xi\neq 0$. Odd and purely imaginary, these multipliers fall in the scope of Proposition 5.3.11, so that they can be studied on real Banach spaces as well. The Riesz transforms are often referred to as the "analogues of the Hilbert transform in $\mathbb{R}^d$", and the following statement (and its proof) makes this analogy very precise:

Theorem 5.5.1. *Let X be a Banach space and let $p\in(1,\infty)$. Then the Riesz transform R_j is bounded on $L^p(\mathbb{R}^d;X)$ if and only if X is a UMD space. In this case*

$$\|R_j\|_{\mathscr{L}(L^p(\mathbb{R}^d;X))} = \|H\|_{\mathscr{L}(L^p(\mathbb{R};X))}.$$

The estimate "$\geqslant$", and thereby the necessity of UMD, is a special case of the following Proposition with $\lambda=0$. We record the slightly more general statement with $\lambda\geqslant 0$, since this is needed for some other applications.

Proposition 5.5.2. *Let X be a Banach space and let $p\in(1,\infty)$. Let*

$$m_{\lambda,j}(\xi) = \frac{\xi_j}{(\lambda+|\xi|^2)^{1/2}}$$

for some index $1\leqslant j\leqslant d$ and $\lambda\geqslant 0$. If $m_{\lambda,j}\in\mathfrak{M}L^p(\mathbb{R}^d;X)$, then X is a UMD space and

$$\|H\|_{\mathscr{L}(L^p(\mathbb{R};X))}\leqslant\|m_{\lambda,j}\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)}.$$

Proof. By the rotation invariance of $\mathfrak{M}L^p(\mathbb{R}^d;X)$, it is enough to consider $j=1$. Let $m_k(\xi) = m_{\lambda,1}(k\xi_1, \xi_2, \dots, \xi_d)$ be a dilation in the first coordinate. Then by Proposition 5.3.8,

$$\|m_k\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)} = \|m_{\lambda,1}\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)}.$$

Clearly $m_k \to \rho$ as $k \to \infty$, where $\rho(\xi) = \frac{\xi_1}{|\xi_1|}$ and hence from Proposition 5.3.16 we find $\rho \in \mathfrak{M}L^p(\mathbb{R}^d;X)$ with

$$\|\rho\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)} \leqslant \liminf_{k\to\infty} \|m_k\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)}.$$

By Fubini's theorem it is immediate that

$$\|H\|_{\mathscr{L}(L^p(\mathbb{R};X))} = \|\rho\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)}$$

and Theorem 5.1.1 yields that X has UMD. $\square$

For the other direction of Theorem 5.5.1, we use a version of the useful "method of rotations". Let σ denote the surface measure on the unit sphere S^{d-1} in $\mathbb{R}^d$. We have the following representation:

Lemma 5.5.3. *Let $c = \int_{S^{d-1}} |\theta_1| \, \mathrm{d}\sigma(\theta)$. Then for each $1 \leqslant j \leqslant n$,*

$$\frac{\xi_j}{|\xi|} = \frac{1}{c}\int_{S^{d-1}} \operatorname{sgn}(\xi\cdot\theta)\theta_j \, \mathrm{d}\sigma(\theta), \quad \xi \in \mathbb{R}^d \setminus \{0\}.$$

Proof. We may assume by scaling that ξ is a unit vector. It is then convenient to prove the slightly more general formula

$$\xi\cdot\eta = \frac{1}{c}\int_{S^{d-1}} \operatorname{sgn}(\xi\cdot\theta)(\eta\cdot\theta) \, \mathrm{d}\sigma(\theta),$$

where $\eta \in \mathbb{R}^d$ is another vector. Let us decompose $\eta = (\eta\cdot\xi)\xi + \eta_\perp$, where $\eta_\perp \cdot \xi = 0$. Then

$$\begin{aligned}
&\int_{S^{d-1}} \operatorname{sgn}(\xi\cdot\theta)(\eta\cdot\theta) \, \mathrm{d}\sigma(\theta) \\
&\quad = (\eta\cdot\xi)\int_{S^{d-1}} |\xi\cdot\theta| \, \mathrm{d}\sigma(\theta) + \int_{S^{d-1}} \operatorname{sgn}(\xi\cdot\theta)(\eta_\perp\cdot\theta) \, \mathrm{d}\sigma(\theta) \\
&\quad =: (\eta\cdot\xi) I + II.
\end{aligned}$$

In integral I, using the rotational invariance of the surface measure, we replace θ by $Q^t\theta$, where Q is an orthogonal transformation that takes ξ to e_1. Then $|\xi\cdot Q^t\theta| = |e_1\cdot\theta| = |\theta_1|$, so that $I = c$ by definition. In II, we also replace θ by $Q^t\theta$, where we now choose Q to be the orthogonal transformation that takes ξ to $-\xi$ and leaves the vectors perpendicular to ξ invariant. Then

$$\operatorname{sgn}(\xi\cdot Q^t\theta)(\eta_\perp\cdot Q^t\theta) = \operatorname{sgn}(-\xi\cdot\theta)(\eta_\perp\cdot\theta) = -\operatorname{sgn}(\xi\cdot\theta)(\eta_\perp\cdot\theta),$$

and this shows that $II = -II = 0$, completing the proof. $\square$

Proof of Theorem 5.5.1. It remains to prove the boundedness of the Riesz transforms assuming that X is a UMD space. Since by Theorem 5.1.1 the Hilbert transform is bounded, also ρ given by $\rho(\xi) = \operatorname{sgn}(\xi_1)$ is a Fourier multiplier on $L^p(\mathbb{R}^d; X)$ with $\|\rho\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)} = \|H\|_{\mathscr{L}(L^p(\mathbb{R};X))}$. We claim that

$$\|m_\theta\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)} = \|H\|_{\mathscr{L}(L^p(\mathbb{R};X))}, \tag{5.43}$$

where $m_\theta(\xi) = \operatorname{sgn}(\xi \cdot \theta)$. Indeed, let A be an orthogonal matrix such that $A\theta = e_1$. Then

$$m_\theta(\xi) = \operatorname{sgn}(\xi \cdot A^* e_1) = \operatorname{sgn}(A\xi \cdot e_1) = \rho \circ A(\xi)$$

and the claim follows.

Let $f \in \mathscr{S}(\mathbb{R}^d; X)$. By Lemma 5.5.3, Minkowski's inequality and the claim (5.43),

$$\begin{aligned}
c\|R_j f\|_{L^p(\mathbb{R}^d;X)} &= \Big\| \Big(\xi \mapsto \int_{S^{n-1}} m_\theta(\xi)\widehat{f}(\xi)\theta_j \,\mathrm{d}\sigma(\theta)\Big)^{\vee} \Big\|_{L^p(\mathbb{R}^d;X)} \\
&\leqslant \int_{S^{d-1}} \|(\xi \mapsto m_\theta(\xi)\widehat{f}(\xi))^{\vee}\|_{L^p(\mathbb{R}^d;X)} |\theta_j| \,\mathrm{d}\sigma(\theta) \\
&= \int_{S^{d-1}} \|T_{m_\theta} f\|_{L^p(\mathbb{R}^d;X)} |\theta_j| \,\mathrm{d}\sigma(\theta) \\
&\leqslant \|H\| \, \|f\|_{L^p(\mathbb{R}^d;X)} \int_{S^{d-1}} |\theta_j| \,\mathrm{d}\sigma(\theta) = c\|H\| \, \|f\|_{L^p(\mathbb{R}^d;X)}.
\end{aligned}$$

Upon dividing by c we infer that R_j is bounded with

$$\|R_j\|_{\mathscr{L}(L^p(\mathbb{R}^d;X))} \leqslant \|H\|_{\mathscr{L}(L^p(\mathbb{R};X))}.$$

□

An application to mixed derivatives

We briefly interrupt the main line of development to give a quick application of the foregoing result. Let X be an arbitrary Banach space. We recall from Definition 2.5.1 that a function $g \in L^1_{\mathrm{loc}}(\mathbb{R}^d; X)$ is said to be the *weak derivative of order* α of a function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d; X)$, where $\alpha \in \mathbb{N}^d$ is a multi-index, if

$$\int_{\mathbb{R}^d} f(x)\partial^\alpha \phi(x) \,\mathrm{d}x = (-1)^{|\alpha|} \int_{\mathbb{R}^d} g(x)\phi(x) \,\mathrm{d}x \quad \text{for all } \phi \in C_{\mathrm{c}}^\infty(\mathbb{R}^d).$$

We have seen in Proposition 2.5.2 that weak derivatives are necessarily unique and we write

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

Along similar lines we may call a function $g \in L^1_{\mathrm{loc}}(\mathbb{R}^d; X)$ the *weak Laplacian* of of the function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d; X)$ if

$$\int_{\mathbb{R}^d} f(x)\Delta\phi(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} g(x)\phi(x)\,\mathrm{d}x \quad \text{for all } \phi \in C_{\mathrm{c}}^\infty(\mathbb{R}^d). \tag{5.44}$$

A weak Laplacian, if it exists, is again unique and we may write $g =: \Delta f$.

For $p \in [1,\infty]$, we denote by $H^{2,p}(\mathbb{R}^d;X)$ the space of all $f \in L^p(\mathbb{R}^d;X)$ with a weak Laplacian in $L^p(\mathbb{R}^d;X)$. This is a Banach space with respect to the norm

$$\|f\|_{H^{2,p}(\mathbb{R}^d;X)} := \|f\|_p + \|\Delta f\|_p. \tag{5.45}$$

Indeed, if $(f_n)_{n\geqslant 1}$ is a Cauchy sequence in $H^{2,p}(\mathbb{R}^d;X)$, then both $(f_n)_{n\geqslant 1}$ and $(\Delta f_n)_{n\geqslant 1}$ are Cauchy sequences in $L^p(\mathbb{R}^d;X)$, say with limits f and g, and then it is immediate from (5.44) that f has weak Laplacian g, so $f \in H^{2,p}(\mathbb{R}^d;X)$ and $f_n \to f$ in $H^{2,p}(\mathbb{R}^d;X)$. Later on, the definition of $H^{2,p}(\mathbb{R}^d;X)$ will be replaced by a more general one (with a slightly different but equivalent norm), but for our present purposes the given definition is adequate, in that it leads to a streamlined presentation of the result that we have in mind.

The result we wish to prove here reads as follows.

Proposition 5.5.4. *Let X be a UMD space and let $p \in (1,\infty)$, and let $f \in L^p(\mathbb{R}^d;X)$. If f admits a weak Laplacian, then f admits weak derivatives of every order $\alpha \in \mathbb{N}^d$ with $|\alpha| = 2$, and*

$$\|\partial^\alpha f\|_p \leqslant \hbar_{p,X}^2 \|\Delta f\|_p.$$

Writing $\partial^\alpha = \partial_j\partial_k$, where $j = k$ is allowed, the idea is to prove the identity

$$\partial^\alpha f = R_j R_k \Delta f$$

and then to invoke the boundedness of the Riesz transforms R_j, whose norm by Theorem 5.5.1 satisfies $\|R_j\| \leqslant \hbar_{p,X}$.

Let us proceed to the rigorous details. First, as an immediate corollary to Lemma 2.5.5 we note:

Lemma 5.5.5. *$C_{\mathrm{c}}^\infty(\mathbb{R}^d;X)$ is dense in $H^{2,p}(\mathbb{R}^d;X)$ for every $p \in [1,\infty)$.*

Proof of Proposition 5.5.4. Let $\partial^\alpha = \partial_j\partial_k$. By taking Fourier transforms, the identity $\partial^\alpha f = R_j R_k \Delta f$ is obvious for Schwartz functions $f \in \mathscr{S}(\mathbb{R}^d;X)$. In the general case we use Lemma 5.5.5 to find a sequence of Schwartz functions $(f_n)_{n\geqslant 1}$ such that $f_n \to f$ in $H^{2,p}(\mathbb{R}^d;X)$. Then by the estimate which we already know to be true for the f_n we find that $(\partial^\alpha f_n)_{n\geqslant 1}$ is a Cauchy sequence in $L^p(\mathbb{R}^d;X)$, so $\partial^\alpha f_n \to g$ for some $g \in L^p(\mathbb{R}^d;X)$. It is easily checked that g is a weak derivative of order α of f, so that $\|\partial^\alpha f\|_p = \lim_{n\to\infty} \|\partial^\alpha f_n\|_p \leqslant \hbar_{p,X}^2 \lim_{n\to\infty} \|\Delta f_n\|_p = \hbar_{p,X}^2 \lim_{n\to\infty} \|\Delta f\|_p$. $\square$

Under the assumptions of the proposition, not only do all mixed derivatives of order $|\alpha| = 2$ exist, but so do those of order $|\alpha| = 1$. This can be proved by rather elementary means, but as it will also follow from a far-reaching generalisation of Proposition 5.5.4 proved in Subsection 5.5.b (see Theorem 5.6.11), we will not delve any deeper into this at present.

Tensor products and indicators of rectangles

Notwithstanding the importance of the Riesz transforms for many other applications, for our forthcoming considerations the role of the one-dimensional Hilbert transform H will rather be played by its d-fold tensor product

$$H^{\otimes d} := H \otimes \cdots \otimes H \quad (d \text{ times}).$$

It is immediate from Fubini's theorem that it satisfies

$$\|H^{\otimes d}\|_{\mathscr{L}(L^p(\mathbb{R}^d;X))} \leqslant \|H\|^d_{\mathscr{L}(L^p(\mathbb{R};X))} =: \hbar^d_{p,X}.$$

Just as the indicators of intervals played a part in the proof of Mihlin's Multiplier Theorem 5.3.18 on $\mathbb{R}$, their d-fold tensor products will have a similar importance in the extension to several variables. The Cartesian product of d intervals is a rectangle

$$R := [a,b) := [a_1,b_1) \times \cdots \times [a_d,b_d), \quad a,b \in \mathbb{R}^d, \tag{5.46}$$

and it is useful to introduce the following notation:

- $a(R) := a$ and $b(R) := b$ are the "lower left" and "upper right" corners.
- For each $\gamma \in \{0,1\}^d$, we denote

$$c(R,\gamma) := (c_1,\dots,c_d), \qquad c_j := \begin{cases} a_j, & \text{if } \gamma_j = 0, \\ b_j, & \text{if } \gamma_j = 1, \end{cases}$$

so that $\{c(R,\gamma) : \gamma \in \{0,1\}^d\}$ is an enumeration of all corners of R.

Lemma 5.5.6. *For any rectangle $R \subseteq \mathbb{R}^d$, we have*

$$\varDelta_R := T_{\mathbf{1}_R} = \Big(\frac{i}{2}\Big)^d \sum_{\gamma \in \{0,1\}^d} (-1)^{|\gamma|} M_{c(R,\gamma)} H^{\otimes d} M_{-c(R,\gamma)},$$

so that in particular

$$\|\varDelta_R\|_{\mathscr{L}(L^p(\mathbb{R}^d;X))} \leqslant \hbar^d_{p,X}.$$

As in the one-dimensional case, we will need the precise representation of $\varDelta_R$, rather than just the norm bound derived from it.

Proof. If R is as in (5.46), then $\varDelta_R$ is a (tensor) product of one-dimensional multipliers acting in each coordinate direction, more precisely

$$\varDelta_R = \bigotimes_{j=1}^d \varDelta^{(j)}_{[a_j,b_j)} = \bigotimes_{j=1}^d \Big(\frac{i}{2}\Big)[M^{(j)}_{a_j} H^{(j)} M^{(j)}_{-a_j} - M^{(j)}_{b_j} H^{(j)} M^{(j)}_{-b_j}],$$

where the superscript (j) indicates the action of a one-dimensional operator along the jth coordinate of $\mathbb{R}^d$. Multiplying out, this becomes $((i/2)^d$ times) a sum of 2^d terms of the form

$$(-1)^k M^{(1)}_{c_1} H^{(1)} M^{(1)}_{-c_1} \circ M^{(2)}_{c_2} H^{(2)} M^{(2)}_{-c_2} \circ \ldots \circ M^{(d)}_{-c_d} H^{(d)} M^{(d)}_{-c_d},$$

where $c_j \in \{a_j, b_j\}$ for each $j = 1, \ldots, d$, and k is the total number of j's such that $c_j = b_j$. Observing that any two operators acting in the different coordinates commute, we can arrange all $M^{(j)}_{c_j}$'s to the left, all $H^{(j)}$'s in the middle, and all $M^{(j)}_{-c_j}$'s to the right, arriving at

$$(-1)^k M_c H^{\otimes d} M_{-c} = (-1)^{|\gamma|} M_{c(R,\gamma)} H^{\otimes d} M_{-c(R,\gamma)}$$

for some $\gamma \in \{0,1\}^d$. Going through all choices of $c_j \in \{a_j, b_j\}$ for all j, we go through all choices of $\gamma \in \{0,1\}^d$ exactly once.

The norm bound is immediate, since we have an average of 2^d terms, each of norm $\hbar^d_{p,X}$. □

5.5.b Mihlin's multiplier theorem on $\mathbb{R}^d$

The main result of this section is an extension of Theorem 5.3.18 to multipliers of several variables defined on $\mathbb{R}^d$. This is based on the same ideas as the one-dimensional version, and it is essentially just the notation that becomes somewhat more complicated. We begin with the relevant decomposition of the domain $\mathbb{R}^d \setminus \{0\}$ into appropriate cubes, playing the role of the intervals $\mathscr{I}$ on $\mathbb{R}$. It is convenient to take our decomposition as a subset of the standard dyadic cubes

$$\mathscr{D}^0 := \{2^{-j}([0,1)^d + m) : \ j \in \mathbb{Z}, \ m \in \mathbb{Z}^d\}$$

of the domain $\mathbb{R}^d$.

Definition 5.5.7. *Let*

$$\mathscr{I}^d := \big\{Q \in \mathscr{D}^0 : \mathrm{dist}_\infty(Q,0) = \ell(Q)\big\},$$
$$\mathscr{I}^d_j := \{Q \in \mathscr{I}^d : \ell(Q) = 2^{-j}\}, \qquad j \in \mathbb{Z},$$

where $\ell(Q)$ is the side-length of Q and dist_∞ *refers to the distance in the ℓ^∞-metric of $\mathbb{R}^d$.*

This definition is illustrated by Figure 5.2. Equivalently, we could have set

$$\mathscr{I}^d := \{\text{maximal } Q \in \mathscr{D}^0 : \mathrm{dist}_\infty(Q,0) \geqslant \ell(Q)\},$$

from which we recognise $\mathscr{I}^d$ as a *Whitney decomposition* of the open set $\mathbb{R}^d \setminus \{0\}$. This reflects the fact that we wish to deal with multipliers exhibiting singular behaviour in the neighbourhood of the origin only. We leave it to the reader to verify the basic properties of these cubes recorded in the following lemma.

Lemma 5.5.8. *The following properties hold:*

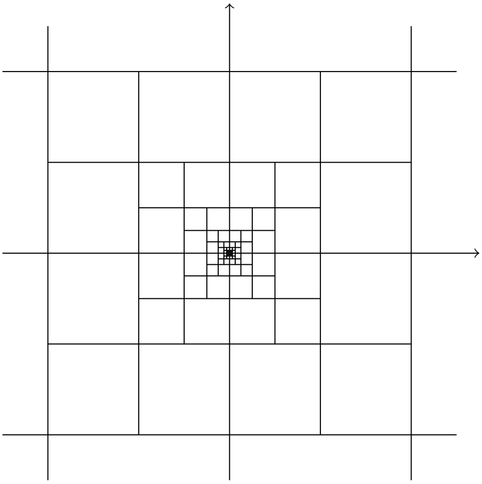

Fig. 5.2: Some of the cubes (squares) of $\mathscr{I}^2$ in $\mathbb{R}^2$. The scale of the picture is irrelevant, as the same pattern repeats at every length scale. The number of squares of each fixed size is $2^2(2^2-1)=12$.

(i) *$\mathscr{I}^d$ is a partition of $\mathbb{R}^d \setminus \{0\}$.*
(ii) *$\mathscr{I}^d_j = 2^{-j}\mathscr{I}^d_0 := \{2^{-j}Q : Q \in \mathscr{I}^d_0\}$ for each $j \in \mathbb{Z}$.*
(iii) *$\#\mathscr{I}^d_0 = 2^d(2^d-1)$.*

The d-dimensional version of Mihlin's multiplier class is defined as follows:

Definition 5.5.9 (Mihlin's multiplier class $\mathfrak{M}(\mathbb{R}^d; X, Y)$). *We say that $m \in L^\infty(\mathbb{R}^d; \mathscr{L}(X,Y))$ belongs to Mihlin's class $\mathfrak{M}(\mathbb{R}^d; X, Y)$ if the following conditions hold:*

(i) *The function m is continuous in the interior Q° of each $Q \in \mathscr{I}^d$, and it has continuous mixed partial derivatives $\partial^\alpha m$, $\alpha \in \{0,1\}^d$, away from finitely many hyperplanes of the form $\{\xi \in Q^\circ : \xi_i = a\}$, in each Q°.*
(ii) *The following Mihlin norm is finite:*

$$\|m\|_{\mathfrak{M}_p(\mathbb{R}^d;X,Y)} := \sum_{\alpha\in\{0,1\}^d} \sup_{|\xi|_\infty\in(1,2)} \mathscr{R}_p\Big(\Big\{|2^k\xi|_\infty^{|\alpha|}(\partial^\alpha m)(2^k\xi)\Big\}_{k\in\mathbb{Z}}\Big). \tag{5.47}$$

We write $\mathfrak{M}(\mathbb{R}^d) := \mathfrak{M}(\mathbb{R}^d; \mathbb{C}, \mathbb{C})$ for the scalar-valued case.

Remarks similar to those following the one-dimensional Definition 5.3.17 could be made here as well, but will not be repeated in all detail. As in the one-dimensional cases, allowing discontinuities over the boundaries of the cubes $\mathscr{I}^d$ will simplify some considerations, even if one is just interested in the case of continuous multipliers in the end. From the assumptions it easily follows that $\partial^\alpha m$ has a one-sided limit on the boundary along any of the non-differentiated coordinates ξ_i with $\alpha_i = 0$.

The fact that we use the ℓ^∞ norm on $\mathbb{R}^d$ in (5.47) is of course completely irrelevant, since all norms on $\mathbb{R}^d$ are equivalent anyway. However, this is the natural norm for considerations involving cubes, and allows us to avoid some $\sqrt{d}$ factors in the estimates.

Theorem 5.5.10 (Mihlin's multiplier theorem on $\mathbb{R}^d$). *Let X and Y be complex UMD spaces and let $p \in (1,\infty)$. Then each $m \in \mathfrak{M}(\mathbb{R}^d; X, Y)$ is a Fourier multiplier for $L^p(\mathbb{R}^d; X)$, and*

$$\|m\|_{\mathfrak{M}L^p(\mathbb{R}^d;X,Y)} \leqslant 100 \cdot 8^d \cdot \min(\hbar_{p,X}, \hbar_{p,Y})^d \cdot \beta_{p,X}\beta_{p,Y} \cdot \|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)}.$$

For scalar-valued m with $\overline{m(-\xi)} = m(\xi)$, the result is valid for real UMD spaces as well.

As in dimension one, our first example of the range of Mihlin's theorem is the following important class of multipliers:

Corollary 5.5.11. *Let X be a complex UMD space and let $p \in (1,\infty)$. Then $m_s(\xi) = |\xi|^{is}$ is a Fourier multiplier for $L^p(\mathbb{R}^d; X)$ for every $s \in \mathbb{R}$, and*

$$\|T_{m_s}\|_{\mathscr{L}(L^p(\mathbb{R}^d;X))} \leqslant 100 \cdot 8^d \cdot \hbar_{p,X}^d \cdot \beta_{p,X}^2 \cdot (|s| + d)^d.$$

Proof. We need to estimate the Mihlin norm $\|m_s\|_{\mathfrak{M}(\mathbb{R}^d)}$. One checks by induction on $|\alpha|$ that for all $\alpha \in \{0,1\}^d$

$$\partial^\alpha m_s(\xi) = \prod_{k=0}^{|\alpha|-1} (is - 2k)|\xi|^{is-2|\alpha|}\xi^\alpha,$$

and thus, estimating by the inequality between geometric and arithmetic means,

$$\begin{aligned} |\xi|_\infty^{|\alpha|}|\partial^\alpha m_s(\xi)| &\leqslant \prod_{k=0}^{|\alpha|-1} (|s| + 2k) \\ &\leqslant \Big(\frac{1}{|\alpha|}\sum_{k=0}^{|\alpha|-1} (|s| + 2k)\Big)^{|\alpha|} = (|s| + |\alpha| - 1)^{|\alpha|}. \end{aligned}$$

Summing over $\alpha \in \{0,1\}^d$, we get

$$\|m_s\|_{\mathfrak{M}(\mathbb{R}^d)} \leqslant \sum_{\alpha\in\{0,1\}^d} (|s| + |\alpha| - 1)^{|\alpha|} \leqslant \sum_{k=0}^{d} \binom{d}{k} (|s| + d - 1)^k = (|s| + d)^d,$$

and the result follows from Theorem 5.5.10. $\square$

The proof of Mihlin's theorem proceeds in parallel with the one-dimensional case, and starts from the basic invariances of the multiplier class:

Proposition 5.5.12 (Invariances of the Mihlin class on $\mathbb{R}^d$). *Let $m, m_1, m_2 \in \mathfrak{M}(\mathbb{R}^d; X, Y)$, $m_3 \in \mathfrak{M}(\mathbb{R}^d; Y, Z)$ and $m_Q \in \mathfrak{M}(\mathbb{R}^d)$. Then the following multipliers also belong to the Mihlin class, with the indicated estimates for their norms:*

(1) *The pointwise sum and product $m_1 + m_2$ and $m_3 m_2$ with*

$$\begin{aligned}\|m_1 + m_2\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} &\leqslant \|m_1\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} + \|m_2\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)},\\ \|m_3 m_2\|_{\mathfrak{M}(\mathbb{R}^d;X,Z)} &\leqslant \|m_3\|_{\mathfrak{M}(\mathbb{R}^d;Y,Z)}\|m_2\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)}.\end{aligned}$$

(2) *The dyadic dilations and independent reflections in each coordinate:*

$$\|m(\eta_1 2^j \cdot, \cdots, \eta_d 2^j \cdot)\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} = \|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} \quad (j \in \mathbb{Z},\ \eta \in \{-1,1\}^d).$$

(3) *The sum of dyadic pieces $\sum_{Q\in\mathscr{I}} \mathbf{1}_Q m_Q$ of scalar-valued multipliers m_Q of uniformly bounded Mihlin norm:*

$$\Big\| \sum_{Q\in\mathscr{I}} \mathbf{1}_Q m_Q \Big\|_{\mathfrak{M}(\mathbb{R}^d)} \leqslant \sup_{Q\in\mathscr{Q}} \|m_Q\|_{\mathfrak{M}(\mathbb{R}^d)}.$$

Proof. Most claims are straightforward and analogous to the one-dimensional case. Concerning products of multipliers, we have

$$\begin{aligned}&\mathscr{R}_p\Big(|2^k\xi|_\infty^{|\alpha|}\partial^\alpha(m_1 m_2)(2^k\xi) : k \in \mathbb{Z}\Big)\\ &\leqslant \sum_{\substack{\theta,\beta\in\{0,1\}^d\\ \theta+\beta=\alpha}} \mathscr{R}_p\Big(|2^k\xi|_\infty^{|\theta|}\partial^\theta m_1(2^k\xi) : k \in \mathbb{Z}\Big)\mathscr{R}_p\Big(|2^k\xi|_\infty^{|\beta|}\partial^\beta m_2(2^k\xi) : k \in \mathbb{Z}\Big).\end{aligned}$$

We then take the supremum over $|\xi|_\infty \in (1,2)$, estimate up by moving this supremum inside the sum, and sum over $\alpha \in \{0,1\}^d$, concluding with the observation that

$$\sum_{\alpha\in\{0,1\}^d} \sum_{\substack{\theta,\beta\in\{0,1\}^d\\ \theta+\beta=\alpha}} = \sum_{\substack{\theta,\beta\in\{0,1\}^d\\ \theta+\beta\in\{0,1\}^d}} \leqslant \sum_{\theta,\beta\in\{0,1\}^d},$$

which allows us to split the double sum into a product of two single sums, giving precisely the estimate

$$\|m_3 m_2\|_{\mathfrak{M}(\mathbb{R}^d;X,Z)} \leqslant \|m_3\|_{\mathfrak{M}(\mathbb{R}^d;Y,Z)}\|m_2\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)}.$$

□

Proof of Mihlin's theorem

By Proposition 5.3.11 again, it suffices to consider the complex case. Before we get properly started, we still need some additional technical notation. For $\alpha \in \{0,1\}^d$, we denote by

$$\mathbb{R}^\alpha := ((\xi_i)_{i:\alpha_i=1} : \xi_i \in \mathbb{R}\}$$

the space $\mathbb{R}^{|\alpha|}$ with an indexing of the components derived from α, and for any $\xi \in \mathbb{R}^d$ and $Q = I_1 \times \cdots \times I_d \subseteq \mathbb{R}^d$, we let

$$\xi_\alpha := (\xi_i)_{i:\alpha_i=1} \in \mathbb{R}^\alpha, \qquad Q_\alpha := \prod_{i:\alpha_i=1} I_i \subseteq \mathbb{R}^\alpha$$

be their natural projections onto $\mathbb{R}^\alpha$. In particular, we will often use the splittings $\xi = (\xi_\alpha, \xi_{\mathbf{1}-\alpha}) \in \mathbb{R}^\alpha \times \mathbb{R}^{\mathbf{1}-\alpha}$ and $Q = Q_\alpha \times Q_{\mathbf{1}-\alpha}$, where $\mathbf{1} := (1,\ldots,1)$.

The "fundamental theorem of calculus" formula (5.36) now takes the form

$$\begin{aligned}
\mathbf{1}_{[a,b)}(\xi)m(\xi) &= \mathbf{1}_{[a,b)}(\xi) \sum_{\alpha\in\{0,1\}^d} \int_{[a,\xi)_\alpha} \partial^\alpha m(a_{\mathbf{1}-\alpha}, \eta_\alpha)\,\mathrm{d}\eta_\alpha \\
&= \sum_{\alpha\in\{0,1\}^d} \int_{[a,b)_\alpha} \mathbf{1}_{[\eta,b)_\alpha\times[a,b)_{\mathbf{1}-\alpha}}(\xi)\partial^\alpha m(a_{\mathbf{1}-\alpha}, \eta_\alpha)\,\mathrm{d}\eta_\alpha.
\end{aligned} \tag{5.48}$$

For concreteness, we spell out the two-dimensional case:

$$\begin{aligned}
\mathbf{1}_{[a,b)}(\xi)m(\xi) = \mathbf{1}_{[a,b)}(\xi)m(a) &+ \int_{a_1}^{b_1} \mathbf{1}_{[\eta_1,b_1)\times[a_1,b_1)}(\xi)\partial_1 m(\eta_1,a_2)\,\mathrm{d}\eta_1 \\
&+ \int_{a_2}^{b_2} \mathbf{1}_{[a_1,b_1)\times[\eta_2,b_2)}(\xi)\partial_2 m(a_1,\eta_2)\,\mathrm{d}\eta_2 \\
&+ \int_{a_1}^{b_1}\int_{a_2}^{b_2} \mathbf{1}_{[\eta_1,b_1)\times[\eta_2,b_2)}(\xi)\partial_{12} m(\eta_1,\eta_2)\,\mathrm{d}\eta_1\,\mathrm{d}\eta_2.
\end{aligned}$$

We hope that this will convince the reader of the need of a reasonably efficient multi-index notation, as in (5.48), in order to avoid formulae getting out of hand in the multi-dimensional situation.

We also rewrite (5.48) as follows.

Lemma 5.5.13. *For a cube $Q \subseteq \mathbb{R}^d$, we have*

$$\mathbf{1}_Q(\xi)m(\xi) = \sum_{\alpha\in\{0,1\}^d} \int_{[0,1)^\alpha} \mathbf{1}_{Q[t_\alpha]}(\xi)\partial^\alpha m(a(Q[t_\alpha]))\,\mathrm{d}t_\alpha,$$

where

$$Q[t_\alpha] := \{\xi \in Q : a(Q) + t_\alpha \ell(Q) \leqslant \xi < b(Q)\}.$$

Proof. This is an immediate reformulation of (5.48) after the change of variable $\eta = a(Q) + t\ell(Q)$, $t \in [0,1)$. $\square$

Now Lemma 5.3.21 has an almost exact analogue on $\mathbb{R}^d$:

Lemma 5.5.14. *Let X or Y be a UMD space and let $p \in (1,\infty)$. If $m \in \mathfrak{M}(\mathbb{R}^d; X, Y)$ and $f_k \in L^p(\mathbb{R}^d; X)$ for each $k \in \mathbb{Z}$, then for each $Q \in \mathscr{I}_0^d$, we have*

$$\Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k T_{\mathbf{1}_{2^k Q} m} f_k \Big\|_{L^p(\Omega\times\mathbb{R}^d;Y)} \leqslant 2^d \|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} \min(\hbar_{p,X}, \hbar_{p,Y})^d \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k f_k \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)}.$$

Proof. A combination of Lemmas 5.5.13 and 5.5.6 gives the identity

$$\begin{aligned} T_{\mathbf{1}_{2^k Q} m} f_k = \sum_{\alpha,\gamma\in\{0,1\}^d} \Big(\frac{i}{2}\Big)^d (-1)^{|\gamma|} \int_{[0,1)^\alpha} & 2^{k|\alpha|} \partial^\alpha m(a(2^k Q[t_\alpha])) \\ & \times M_{c(2^k Q[t_\alpha],\gamma)} H^{\otimes d} M_{-c(2^k Q[t_\alpha],\gamma)} f_Q \, \mathrm{d}t_\alpha, \end{aligned}$$

where we note that

$$\begin{aligned} &\mathscr{R}_p\Big(2^{k|\alpha|}\partial^\alpha m(a(2^k Q[t_\alpha])) : Q \in \mathscr{I}^d\Big) \\ &= \mathscr{R}_p\Big(\Big[\frac{2^k}{|a(2^k Q[t_\alpha])|_\infty}\Big]^{|\alpha|} |a(2^k Q[t_\alpha])|_\infty^{|\alpha|} \partial^\alpha m(a(2^k Q[t_\alpha])) : Q \in \mathscr{I}^d\Big) \\ &\leqslant \|m\|_{\mathfrak{M}_p(\mathbb{R}^d;X,Y)}, \end{aligned}$$

since $a(2^k Q[t_\alpha]) = 2^k a(Q[t_\alpha]) \in 2^k Q$ satisfies

$$|a(2^k Q[t_\alpha])|_\infty \geqslant \mathrm{dist}_\infty(2^k Q, 0) = 2^k \ell(Q) = 2^k.$$

Thus, from the definition of R-boundedness, the contraction principle and the boundedness of $H^{\otimes d}$, we have

$$\begin{aligned} &\Big\| \sum_{k\in\mathbb{Z}} \varepsilon_Q 2^{k|\alpha|} \partial^\alpha m(a(2^k Q[t_\alpha])) M_{c(2^k Q[t_\alpha],\gamma)} H^{\otimes d} M_{-c(2^k Q[t_\alpha],\gamma)} f_k \Big\|_{L^p(\Omega\times\mathbb{R}^d;Y)} \\ &\leqslant \|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k M_{c(2^k Q[t_\alpha],\gamma)} H^{\otimes d} M_{-c(2^k Q[t_\alpha],\gamma)} f_k \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \\ &\leqslant \|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k H^{\otimes d} M_{-c(2^k Q[t_\alpha],\gamma)} f_k \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \\ &\leqslant \|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} \hbar_{p,X}^d \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k M_{-c(2^k Q[t_\alpha],\gamma)} f_k \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \\ &\leqslant \|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} \hbar_{p,X}^d \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k f_k \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)}. \end{aligned}$$

Since $2^{k|\alpha|}\partial^\alpha m(a(2^k Q[t_\alpha]))$ commutes with $M_{c(2^kQ[t_\alpha],\gamma)}H^{\otimes d}M_{-c(2^kQ[t_\alpha],\gamma)}$, we could also have performed the estimates in a different order, using the boundedness of $H^{\otimes d}$ on $L^p(\mathbb{R}^d;Y)$ instead, and giving the factor $\hbar_{p,Y}^d$ instead of $\hbar_{p,X}^d$. Taking the minimum of the two possible estimates, we hence conclude that

$$\begin{aligned}
&\Big\|\sum_{k\in\mathbb{Z}}\varepsilon_k T_{\mathbf{1}_{2^kQ}m} f_k\Big\|_{L^p(\mathbb{R}^d;Y)}\\
&\leqslant \sum_{\alpha,\gamma\in\{0,1\}^d}\frac{1}{2^d}\int_{[0,1)^\alpha}\|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)}\min(\hbar_{p,X},\hbar_{p,Y})^d\Big\|\sum_{k\in\mathbb{Z}}\varepsilon_k f_k\Big\|_{L^p(\Omega\times\mathbb{R}^d;X)}\,\mathrm{d}t_\alpha\\
&= 2^d\|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)}\min(\hbar_{p,X},\hbar_{p,Y})^d\Big\|\sum_{k\in\mathbb{Z}}\varepsilon_k f_k\Big\|_{L^p(\Omega\times\mathbb{R}^d;X)}.
\end{aligned}$$

□

Lemma 5.5.15. *Let ϕ and ϕ_I be defined as in (5.38). For each $Q = I_1\times\cdots\times I_d\in\mathscr{I}^d$, let*

$$\tau(Q) := \min\{j\in\{1,\dots,d\} : 0\notin\partial I_j\}$$

and

$$\phi_Q(x) := \phi_{I_{\tau(Q)}}(x_{\tau(Q)}),\qquad \widehat{\phi}_Q(\xi) := \widehat{\phi}_{I_{\tau(Q)}}(\xi_{\tau(Q)}).$$

Then

$$\Big\|\sum_{Q\in\mathscr{I}^d}\frac{\mathbf{1}_Q}{\widehat{\phi}_Q}\Big\|_{\mathfrak{M}(\mathbb{R}^d)} \leqslant \sup_{\xi\in[\frac12,1]}\Big|\frac{1}{\widehat{\phi}(\xi)}\Big| + \sup_{\xi\in[\frac12,1]}\Big|\frac{d}{d\xi}\Big(\frac{1}{\widehat{\phi}(\xi)}\Big)\Big| \leqslant 9.$$

Although we identify the one-variable functions ϕ_Q and $\widehat{\phi}_Q$ with functions of $\mathbb{R}^d$ that are constant with respect to the other variables, note that $\widehat{\phi}_Q$ is the Fourier transform of ϕ_Q viewed as a function of one variable, not as a function on $\mathbb{R}^d$. This is hopefully sufficiently clear, so that we do not indicate it explicitly, in order not to overburden the notation.

Proof. Fix a $Q = I_1\times\cdots\times I_d\in\mathscr{I}^d$, and let $j := \tau(Q)$ and $I := I_j$. For $\xi\in Q^\circ$, the function $\mathbf{1}_Q(\xi)/\widehat{\phi}_Q(\xi) = 1/\widehat{\phi}_I(\xi_j)$ depends only on the coordinate ξ_j, and hence all its partial derivatives ∂^α, $\alpha\in\{0,1\}^d$, vanish, except for $\alpha\in\{0,e_j\}$. So for the Mihlin bound we only need to compute

$$\sup_{\xi\in Q^\circ}\Big|1/\widehat{\phi}_I(\xi_j)\Big| + \sup_{\xi\in Q^\circ}|\xi|_\infty\Big|(1/\widehat{\phi}_I)'(\xi_j)\Big|,$$

where $|\xi|_\infty\leqslant 2\ell(Q) = 2\ell(I)$ by the properties of $Q\in\mathscr{I}^d$. Then the above expression depends on I only, and by scaling and reflection invariance we may take $I = [\frac12,1)$. But then $2\ell(I) = 1$, leading to the asserted bound. The final numerical estimate was obtained in Lemma 5.3.23. □

Lemma 5.5.16. *For each $Q \in \mathscr{I}^d$, we have*

$$\Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k \phi_{2^k Q} *_{\tau(Q)} f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \leqslant \beta_{p,X}^{+} \|f\|_{L^p(\mathbb{R}^d;X)},$$

where $*_{\tau(Q)}$ indicates convolution with respect to the variable $x_{\tau(Q)}$.

Proof. It is straightforward from the definition that $\tau(2^k Q) = \tau(Q) =: j$, and thus $\phi_{2^k Q} = \phi^{(j)}_{2^k I_j}$, where I_j, the projection of Q on the jth coordinate, is a dyadic interval of $\mathbb{R}$ at the distance $\ell(I^{(j)})$ from the origin. As $k \in \mathbb{Z}$, the interval $2^k I_j$ runs through either $\mathscr{I}_-$ or $\mathscr{I}_+$, depending on whether I_j lies on the negative or the positive half-line. Thus the claim is equivalent to

$$\Big\| \sum_{I\in\mathscr{I}_\pm} \varepsilon_I \phi_I *_j f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \leqslant \beta_{p,X}^{+} \|f\|_{L^p(\mathbb{R};X)},$$

which is immediate from Lemma 5.3.22 by Fubini's theorem. □

Lemma 5.5.17. *Let X be a UMD space, Y a Banach space, and let $p \in (1,\infty)$. If $m \in \mathfrak{M}(\mathbb{R}^d;X,Y)$, $Q \in \mathscr{I}^d$ and $f \in L^p(\mathbb{R}^d;X)$, then we have the following estimate for all $a \in \mathbb{N}$:*

$$\begin{aligned} &\Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k T_{\mathbf{1}_{2^k Q} m\cdot(\widehat{\phi}_{2^k Q})^{-a}} f \Big\|_{L^p(\Omega\times\mathbb{R}^d;Y)} \\ &\qquad \leqslant 9^{a+1}\cdot 2^d \|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} \min(\hbar_{p,X},\hbar_{p,Y})^d \beta_{p,X}^{+} \|f\|_{L^p(\mathbb{R}^d;X)}. \end{aligned}$$

Proof. Let $j := \tau(Q) = \tau(2^k Q)$ for all $k \in \mathbb{Z}$. Then we compute

$$\begin{aligned} &\Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k T_{\mathbf{1}_{2^k Q} m\cdot(\widehat{\phi}_{2^k Q})^{-a}} f \Big\|_{L^p(\Omega\times\mathbb{R}^d;Y)} \\ &= \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k T_{\mathbf{1}_{2^k Q} m\cdot(\widehat{\phi}_{2^k Q})^{-a-1}} (\phi_{2^k Q} *_j f) \Big\|_{L^p(\Omega\times\mathbb{R}^d;Y)} \\ &\leqslant 2^d \Big\| \sum_{Q\in\mathscr{I}^d} \frac{\mathbf{1}_Q m}{\widehat{\phi}_Q^{a+1}} \Big\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} \min(\hbar_{p,X},\hbar_{p,Y})^d \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k \phi_{2^k Q} *_j f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \\ &\qquad \text{by Lemma 5.5.14} \\ &\leqslant 2^d \|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} \Big\| \sum_{Q\in\mathscr{I}^d} \frac{\mathbf{1}_Q}{\widehat{\phi}_Q} \Big\|_{\mathfrak{M}^1(\mathbb{R}^d)}^{a+1} \min(\hbar_{p,X},\hbar_{p,Y})^d \beta_{p,X}^{+} \|f\|_{L^p(\mathbb{R}^d;X)} \\ &\qquad \text{by Proposition 5.5.12 and Lemma 5.5.16} \\ &\leqslant 2^d \|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)} \cdot 9^{a+1} \cdot \min(\hbar_{p,X},\hbar_{p,Y})^d \beta_{p,X}^{+} \|f\|_{L^p(\mathbb{R}^d;X)} \\ &\qquad \text{by Lemma 5.5.15,} \end{aligned}$$

and this proves the claim. □

Proof of Mihlin's Multiplier Theorem 5.5.10. Fix a function $f \in \mathscr{S}(\mathbb{R}^d; X) \subseteq \check{L}^1(\mathbb{R}^d; X)$. Then $T_m f \in \check{L}^1(\mathbb{R}^d; Y)$. We shall estimate its $L^p(\mathbb{R}^d; Y)$ norm by dualising with $g \in \mathscr{S}(\mathbb{R}^d; Y^*)$, a norming subspace of the dual space. Then, expanding the pairing over the cubes $Q \in \mathscr{I}^d$ in the Fourier domain, we have

$$\begin{aligned}
|\langle T_m f, g\rangle| &= \Big| \sum_{Q \in \mathscr{I}^d} \langle T_{\mathbf{1}_Q m} f, g\rangle \Big| \\
&= \Big| \sum_{Q \in \mathscr{I}_0^d} \sum_{k\in\mathbb{Z}} \langle T_{\mathbf{1}_{2^kQ} m/\check{\phi}_{2^kQ}} f, \phi_{2^kQ} *_{\tau(Q)} g\rangle \Big| \\
&= \Big| \sum_{Q \in \mathscr{I}_0^d} \mathbb{E}\Big\langle \sum_{k\in\mathbb{Z}} \varepsilon_k T_{\mathbf{1}_{2^kQ} m/\check{\phi}_{2^kQ}} f, \sum_{k\in\mathbb{Z}} \varepsilon_k \phi_{2^kQ} *_{\tau(Q)} g \Big\rangle \Big| \\
&\leqslant \sum_{Q \in \mathscr{I}_0^d} \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k T_{\mathbf{1}_{2^kQ} m/\check{\phi}_{2^kQ}} f \Big\|_{L^p(\Omega\times\mathbb{R}^d; Y)} \\
&\qquad\qquad \times \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k \phi_{2^kQ} *_{\tau(Q)} g \Big\|_{L^{p'}(\Omega\times\mathbb{R}^d; Y^*)} =: \sum_{Q \in \mathscr{I}_0^d} A_Q \cdot B_Q.
\end{aligned}$$

Here

$$A_Q \leqslant 9^2 \cdot 2^d \|m\|_{\mathfrak{M}(\mathbb{R}^d; X, Y)} \min(\hbar_{p,X}, \hbar_{p,Y})^d \beta^+_{p,X} \|f\|_{L^p(\mathbb{R}; X)}$$

by Lemma 5.5.17, and

$$B_Q \leqslant \beta^+_{p',Y^*} \|g\|_{L^{p'}(\mathbb{R}^d; Y^*)},$$

for each $Q \in \mathscr{I}_0^d$. Since the number of these cubes is $2^d(2^d - 1)$, altogether we have

$$\begin{aligned}
|\langle T_m f, g\rangle| &\leqslant \sum_{Q \in \mathscr{I}_0^d} A_Q \cdot B_Q \\
&\leqslant 2^d(2^d-1) \cdot 9^2 \cdot 2^d \min(\hbar_{p,X}, \hbar_{p,Y})^d \beta_{p,X}\beta_{p,Y} \|f\|_{L^p(\mathbb{R}; X)} \|g\|_{L^{p'}(\mathbb{R}^d; Y^*)} \\
&\leqslant 100 \cdot 8^d \cdot \min(\hbar_{p,X}, \hbar_{p,Y})^d \beta_{p,X}\beta_{p,Y} \|f\|_{L^p(\mathbb{R}; X)} \|g\|_{L^{p'}(\mathbb{R}^d; Y^*)},
\end{aligned}$$

which is the asserted bound of the Theorem. □

5.5.c Littlewood–Paley inequalities on $\mathbb{R}^d$

As in dimension one, the technique of proof of the multiplier theorem may also be utilised to provide an equivalent 'randomised' norms on the space $L^p(\mathbb{R}^d; X)$, as detailed in the following:

Theorem 5.5.18 (Littlewood–Paley inequality with sharp cut-offs). *Let X be a complex UMD space and let $p \in (1, \infty)$. Then, with*

$$C := C_{d,p,X} = 10 \cdot 8^d \hbar^d_{p,X} \beta_{p,X},$$

we have the following estimates for all $f \in L^p(\mathbb{R}^d; X)$:

$$\frac{1}{C}\|f\|_{L^p(\mathbb{R};X)} \leqslant \Big\| \sum_{Q\in\mathscr{I}^d} \varepsilon_Q \Delta_Q f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \leqslant C\|f\|_{L^p(\mathbb{R}^d;X)}, \tag{5.49}$$

$$\frac{1}{C}\|f\|_{L^p(\mathbb{R};X)} \leqslant \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k \sum_{Q\in\mathscr{I}^d_k} \Delta_Q f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \leqslant C\|f\|_{L^p(\mathbb{R}^d;X)}. \tag{5.50}$$

The last two-sided estimate (5.50) is meaningful and valid for real UMD spaces X as well.

In the second formulation, observe that the kth term is simply

$$\sum_{Q\in\mathscr{I}^d_k} \Delta_Q f = \Delta_{[-2^{1-k},2^{1-k})^d\setminus[-2^{-k},2^{-k})^d} f, \tag{5.51}$$

which is the restriction of f to a rectangular annulus of scale 2^{-k} in the Fourier side. In the first formulation, this annulus is further decomposed into the cubes $Q \in \mathscr{I}^d_k$, each of which are equipped with their 'own' random coefficient ε_Q, while in the second formulation, a random coefficient ε_k is only assigned to each length scale. As one or the other variant may be better suited for a particular purpose, we present them both with an essentially unified proof.

Proof. Writing $\mathscr{I}^d_k = \{2^k Q : \ Q \in \mathscr{I}^d_0, \ k \in \mathbb{Z}\}$, we find that the norms in the middle of both claimed estimates are dominated by

$$\begin{aligned}
\sum_{Q\in\mathscr{I}^d_0} \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_{2^kQ} \Delta_{2^kQ} f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \\
&\leqslant \sum_{Q\in\mathscr{I}^d_0} 9\cdot 2^d \hbar^d_{p,X} \beta^+_{p,X} \|f\|_{L^p(\mathbb{R}^d;X)} \\
&\leqslant 2^d(2^d-1)\cdot 9\cdot 2^d \hbar^d_{p,X} \beta^+_{p,X} \|f\|_{L^p(\mathbb{R}^d;X)},
\end{aligned} \tag{5.52}$$

using Lemma 5.5.17 with $a = 0$ in the first step and counting the number of cubes $Q \in \mathscr{I}^d_0$ in the second.

The reverse bounds are obtained by a duality argument, which we detail in the first case as

$$\begin{aligned}
|\langle f, g\rangle| &= \Big| \mathbb{E} \Big\langle \sum_{Q\in\mathscr{I}^d} \varepsilon_Q \Delta_Q f, \sum_{Q\in\mathscr{I}^d} \bar{\varepsilon}_Q \Delta_Q g \Big\rangle \Big| \\
&\leqslant \Big\| \sum_{Q\in\mathscr{I}^d} \varepsilon_Q \Delta_Q f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \Big\| \sum_{Q\in\mathscr{I}^d} \bar{\varepsilon}_Q \Delta_Q g \Big\|_{L^{p'}(\Omega\times\mathbb{R}^d;X^*)} \\
&\leqslant \Big\| \sum_{Q\in\mathscr{I}^d} \varepsilon_Q \Delta_Q f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \cdot C_{d,p',X^*} \|g\|_{L^{p'}(\mathbb{R}^d;X^*)},
\end{aligned}$$

the second one being a simple modification. Finally, observe that $C_{d,p',X^*} = C_{d,p,X}$, since $\hbar_{p',X^*} = \hbar_{p,X}$ and $\beta_{p',X^*} = \beta_{p,X}$.

Concerning the claim about real UMD spaces, note that the multiplier appearing in (5.50), which is re-expressed in (5.51), is real-valued and even, and hence in the scope of Proposition 5.3.11. To be precise, since different measure spaces $\mathbb{R}^d$ and $\mathbb{R}^d \times \Omega$ appear in (5.50), one should directly apply Proposition 2.1.12 to the operator

$$T : f \mapsto \sum_{k\in\mathbb{Z}} \varepsilon_k \sum_{Q\in\mathscr{I}_k^d} \Delta_Q f, \quad L^p(\mathbb{R}^d; X) \to L^p(\mathbb{R}^d \times \Omega; X)$$

to deduce the right hand side of (5.50). The left hand side then follows by repeating the same duality argument as in the complex case. □

As in the one-dimensional case, the following variant is a straightforward consequence:

Corollary 5.5.19. *Let X be a UMD space and let $p \in (1,\infty)$. Then, with the constant $C := C_{d,p,X} = 10 \cdot 8^d \cdot \hbar_{p,X}^d \beta_{p,X}$, we have the following estimate for all $f \in L^p(\mathbb{R}^d; X)$:*

$$\begin{aligned}\frac{1}{C}\|f\|_{L^p(\mathbb{R};X)} &\leqslant \Big\| \varepsilon_0 \Delta_{[-1,1)^d} f + \sum_{k=1}^{\infty} \varepsilon_k \Delta_{[-2^k,2^k)^d \setminus [-2^{k-1},2^{k+1})^d} f \Big\|_{L^p(\Omega\times\mathbb{R}^d)} \\ &\leqslant C\|f\|_{L^p(\mathbb{R}^d;X)}.\end{aligned}$$

Proof. Abbreviating

$$S_0 := [-1,1)^d, \qquad S_k := [-2^k, 2^k)^d \setminus [-2^{k-1}, 2^{k+1})^d, \quad k \geqslant 1,$$

the norm in the middle is estimated by

$$\begin{aligned}\Big\| \sum_{k=0}^{\infty} \varepsilon_k \Delta_{S_k} f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} &\leqslant \|\Delta_{S_0} f\|_{L^p(\mathbb{R}^d;X)} + \Big\| \sum_{k=1}^{\infty} \varepsilon_k \Delta_{S_k} f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \\ &=: A + B,\end{aligned}$$

where $A \leqslant \hbar_{p,X}^d \|f\|_{L^p(\mathbb{R}^d;X)}$, according to Lemma 5.5.6. The second term, on the other hand, can be dominated by the contraction principle with

$$B \leqslant \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k \Delta_{[-2^k,2^k)^d \setminus [-2^{k-1},2^{k-1})^d} f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)},$$

which is just a re-indexed version in the second variant of Theorem 5.5.18. (Note that a re-numeration of a Rademacher sequence is again a Rademacher sequence.) Thus we have

$$B \leqslant 9 \cdot 8^d \hbar_{p,X}^d \beta_{p,X} \|f\|_{L^p(\mathbb{R}^d;X)}$$

by (5.52) from the proof of Theorem 5.5.18, and the proof of the right side of the claim is completed by estimating $1 + 9 \cdot 8^d \beta_{p,X} \leqslant 10 \cdot 8^d \beta_{p,X}$. The left side follows by a duality argument completely analogous to that in the proof of Theorem 5.5.18. □

We now turn to a "smooth" version of Theorem 5.5.18, and for this we introduce the following definition:

Definition 5.5.20. *A function $\psi \in \mathscr{S}(\mathbb{R}^d)$ is a* smooth Littlewood–Paley function *if:*

(i) *$\widehat{\psi}$ is smooth, non-negative, and supported in $\{\xi \in \mathbb{R}^d : \frac{1}{2} \leqslant |\xi| \leqslant 2\}$;*
(ii) *$\displaystyle\sum_{k\in\mathbb{Z}} \widehat{\psi}(2^{-k}\xi) = 1$ for all $\xi \in \mathbb{R}^d \setminus \{0\}$.*

In the following, we provide the standard construction of such functions. It can be easily adapted to enforce additional conditions on the function ψ, if needed in a particular application.

Lemma 5.5.21. *Smooth Littlewood–Paley functions exist.*

Proof. Let $\widehat{\varphi} \in \mathscr{D}(\mathbb{R}^d)$ satisfy $\mathbf{1}_{B(0,1)} \leqslant \widehat{\varphi} \leqslant \mathbf{1}_{B(0,2)}$; thus $\mathbf{1}_{B(0,\frac{1}{2})} \leqslant \widehat{\varphi}(2\cdot) \leqslant \mathbf{1}_{B(0,1)}$. We define

$$\widehat{\psi}(\xi) := \widehat{\varphi}(\xi) - \widehat{\varphi}(2\xi),$$

so that

$$0 = \mathbf{1}_{B(0,1)} - \mathbf{1}_{B(0,1)} \leqslant \widehat{\psi} \leqslant \mathbf{1}_{B(0,2)} - \mathbf{1}_{B(0,\frac{1}{2})} \leqslant \mathbf{1}_{\{\xi:\frac{1}{2}\leqslant|\xi|\leqslant 2\}},$$

and we have proved (i).

Moreover, for any fixed $\xi \in \mathbb{R}^d$, we have

$$\sum_{k\in\mathbb{Z}} \widehat{\psi}(2^{-k}\xi) = \sum_{k=a}^{b} [\widehat{\varphi}(2^{-k}\xi) - \widehat{\varphi}(2^{-(k-1)}\xi)] = \widehat{\varphi}(2^{-b}\xi) - \widehat{\varphi}(2^{-(a-1)}\xi)$$

for any large enough b and small enough a, since the remaining terms vanish. But then $\widehat{\varphi}(2^{-b}\xi) = 1$ and $\widehat{\varphi}(2^{-(a-1)}\xi) = 0$, and (ii) is proved.

Finally, ψ is a Schwartz function, since the inverse Fourier transforms of a Schwartz function is a Schwartz function again. □

Theorem 5.5.22 (Littlewood–Paley inequality on $\mathbb{R}^d$ with smooth cut-offs). *Let X be a UMD space and let $p \in (1,\infty)$. If ψ is a smooth Littlewood–Paley function, then, with $C := C_{\psi,d,p,X} = c_{d,\psi} \hbar_{p,X}^d \beta_{p,X}$,*

$$\frac{1}{C}\|f\|_{L^p(\mathbb{R}^d;X)} \leqslant \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k \psi_{2^{-k}} * f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \leqslant C\|f\|_{L^p(\mathbb{R}^d;X)},$$

where, as always, $\psi_t = t^{-1}\psi(t^{-1}x)$.

Proof. Since $|\xi|_\infty \leqslant |\xi| \leqslant \sqrt{d}|\xi|_\infty$, we observe that

$$\operatorname{supp}\widehat{\psi} \subseteq \{\xi : \tfrac{1}{2} \leqslant |\xi| \leqslant 2\} \subseteq \{\xi : \tfrac{1}{2\sqrt{d}} \leqslant |\xi|_\infty \leqslant 2\}$$

is contained in the $\bigcup_{j=0}^{J(d)} \bigcup_{Q\in\mathscr{I}_j^d} Q$ for $J(d) = \lceil \log_2(2\sqrt{d})\rceil$. Note that, in the one-dimensional case, this simply amounts to

$$\operatorname{supp}\widehat{\psi} \subseteq [-2,-1] \cup [-1,-\tfrac{1}{2}] \cup [\tfrac{1}{2},1] \cup [1,2], \qquad d=1.$$

Thus

$$\begin{aligned}
&\Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k \psi_{2^{-k}} * f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \\
&\leqslant \sum_{j=0}^{J(d)} \sum_{Q\in\mathscr{I}_j^d} \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k T_{\mathbf{1}_{2^k Q}\widehat{\psi}(2^{-k}\cdot)} f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \\
&\leqslant \sum_{j=0}^{J(d)} \sum_{Q\in\mathscr{I}_j^d} 10\cdot 2^d \hbar_{p,X}^d \beta_{p,X} \Big\| \sum_{k\in\mathbb{Z}} \mathbf{1}_{2^k Q}\widehat{\psi}(2^{-k}\cdot) \Big\|_{\mathfrak{M}(\mathbb{R}^d)} \|f\|_{L^p(\mathbb{R}^d;X)},
\end{aligned}$$

and

$$\Big\| \sum_{k\in\mathbb{Z}} \mathbf{1}_{2^k Q}\widehat{\psi}(2^{-k}\cdot) \Big\|_{\mathfrak{M}(\mathbb{R}^d)} = \sup_{k\in\mathbb{Z}} \|\mathbf{1}_{2^k Q}\widehat{\psi}(2^{-k}\cdot)\|_{\mathfrak{M}(\mathbb{R}^d)} = \|\mathbf{1}_Q\widehat{\psi}\|_{\mathfrak{M}(\mathbb{R}^d)}.$$

From the smoothness of $\widehat{\psi}$, it is immediate that the finite sum

$$\sum_{j=0}^{J(d)} \sum_{Q\in\mathscr{I}_j^d} 10\cdot 2^d \|\mathbf{1}_Q\widehat{\psi}\|_{\mathfrak{M}(\mathbb{R}^d)}$$

is some number $c'_{d,\psi}$ depending only on d and ψ.

For the reverse inequality, we observe that

$$\sum_{u=-1}^{1} \widehat{\psi}(2^{-k-u}\cdot) \equiv 1 \text{ on the support of } \widehat{\psi}(2^{-k}\cdot).$$

Thus

$$\begin{aligned}
|\langle f,g\rangle| &= \Big| \Big\langle \sum_{k\in\mathbb{Z}} \psi_{2^{-k}} * f, g \Big\rangle \Big| \\
&= \Big| \sum_{u=-1}^{1} \sum_{k\in\mathbb{Z}} \langle \psi_{2^{-k}} * f, \psi_{2^{-k-u}} * g\rangle \Big| \\
&= \Big| \sum_{u=-1}^{1} \mathbb{E}\Big\langle \sum_{k\in\mathbb{Z}} \varepsilon_k \psi_{2^{-k}} * f, \sum_{k\in\mathbb{Z}} \bar{\varepsilon}_k \psi_{2^{-k-u}} * g \Big\rangle \Big|
\end{aligned}$$

$$
\begin{aligned}
&\leqslant \sum_{u=-1}^{1} \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k \psi_{2^{-k}} * f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} \Big\| \sum_{k\in\mathbb{Z}} \bar{\varepsilon}_k \psi_{2^{-k-u}} * g \Big\|_{L^{p'}(\Omega\times\mathbb{R}^d;X^*)} \\
&\leqslant \sum_{u=-1}^{1} \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k \psi_{2^{-k}} * f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} c'_{d,\psi} \hbar^d_{p',X^*} \beta_{p',X^*} \|g\|_{L^{p'}(\mathbb{R}^d;X^*)} \\
&\leqslant 3 \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k \psi_{2^{-k}} * f \Big\|_{L^p(\Omega\times\mathbb{R}^d;X)} c'_{d,\psi} \hbar^d_{p,X} \beta_{p,X} \|g\|_{L^{p'}(\mathbb{R}^d;X^*)},
\end{aligned}
$$

and the theorem follows with $c_{d,\psi} = 3c'_{d,\psi}$. □

5.6 Applications to Sobolev spaces

In this section we shall apply some of the multiplier results obtained so far to characterise, for UMD spaces X and $p \in (1,\infty)$, the complex interpolation spaces between $L^p(\mathbb{R}^d;X)$ and the Sobolev spaces $W^{k,p}(\mathbb{R}^d;X)$ as being the Bessel potential spaces $H^{s,p}(\mathbb{R}^d;X)$. More precisely we shall prove:

Theorem 5.6.1 (Seeley). *Let X be a UMD space and let $p \in (1,\infty)$. Then for all integers $k \geqslant 1$ and all $0 < \theta < 1$ we have*

$$[L^p(\mathbb{R}^d;X), W^{k,p}(\mathbb{R}^d;X)]_\theta = H^{\theta k,p}(\mathbb{R}^d;X) \quad \textit{with equivalent morms.}$$

Recall from Section 2.5.b that for $k \in \mathbb{N}$ and $p \in [1,\infty]$, the Sobolev space $W^{k,p}(\mathbb{R}^d;X)$ has been defined as the space of all $f \in L^p(\mathbb{R}^d;X)$ whose weak derivatives of all orders $|\alpha| \leqslant k$ exist and belong to $L^p(\mathbb{R}^d;X)$. The Bessel potential spaces $H^{s,p}(\mathbb{R}^d;X)$ will be defined in terms of Fourier multipliers below.

The plan of the proof consist of two steps, both of which apply Mihlin's multiplier theorem and thus assume that X be UMD and $p \in (1,\infty)$:

1. To prove that the spaces $H^{s,p}(\mathbb{R}^d;X)$ interpolate by the complex method;
2. To prove that $W^{k,p}(\mathbb{R}^d;X) = H^{k,p}(\mathbb{R}^d;X)$ for all integers $k \geqslant 1$.

5.6.a Bessel potential spaces

Formally, the Bessel potential spaces $H^{s,p}(\mathbb{R}^d;X)$ constitute the maximal domains on which the operators $(I-\Delta)^{s/2}$ act as mappings into $L^p(\mathbb{R}^d;X)$. The rigorous definition is in terms of Fourier multipliers.

The *Bessel potential operators* are the operators J_s, $s \in \mathbb{R}$, acting on the space of tempered distributions $\mathscr{S}'(\mathbb{R}^d;X)$, defined by

$$J_s u := ((1+4\pi^2|\cdot|^2)^{s/2}\widehat{u})^{\vee}, \quad u \in \mathscr{S}'(\mathbb{R}^d;X).$$

It is easily seen that multiplication by $(1+4\pi^2|\cdot|^2)^{s/2}$ is continuous on $\mathscr{S}(\mathbb{R}^d;X)$ and therefore the operators J_s are continuous on $\mathscr{S}'(\mathbb{R}^d;X)$. They clearly satisfy the identities

$$J_0 = I, \qquad J_{s_1+s_2} = J_{s_1} \circ J_{s_2},$$

and for even integers $s = 2k$ we have

$$J_{2k} = (1 - \Delta)^k$$

in the distributional sense (cf. Example 2.4.27).

Since the Fourier transform is a continuous bijection on $\mathscr{S}(\mathbb{R}^d; X)$ with continuous inverse, and since multiplication by $(1 + 4\pi^2|\cdot|^2)^{s/2}$ is continuous on $\mathscr{S}(\mathbb{R}^d; X)$, the operators J_s restrict to continuous bijections on $\mathscr{S}(\mathbb{R}^d; X)$, with inverse J_{-s}.

Definition 5.6.2. *Let $s \in \mathbb{R}$ and $p \in [1, \infty]$. The* Bessel potential space *$H^{s,p}(\mathbb{R}^d; X)$ is the space of all $u \in \mathscr{S}'(\mathbb{R}^d; X)$ for which $J_s u$ belongs to $L^p(\mathbb{R}^d; X)$.*

Endowed with the norm

$$\|u\|_{H^{s,p}(\mathbb{R}^d;X)} := \|J_s u\|_{L^p(\mathbb{R}^d;X)},$$

$H^{s,p}(\mathbb{R}^d; X)$ is a Banach space. To see this, let us first observe that this norm is well defined: for if $J_s u_1 = J_s u_2$ in $L^p(\mathbb{R}; X)$, then also $J_s u_1 = J_s u_2$ in $\mathscr{S}'(\mathbb{R}; X)$ (see Example 2.4.26), and therefore $u_1 = J_{-s} J_s u_1 = J_{-s} J_s u_2 = u_2$ in $\mathscr{S}'(\mathbb{R}^d; X)$. Now if $(u_n)_{n\geqslant 1}$ is a Cauchy sequence in $H^{s,p}(\mathbb{R}^d; X)$, then $(J_s u_n)_{n\geqslant 1}$ is a Cauchy sequence in $L^p(\mathbb{R}^d; X)$. Denoting by $f \in L^p(\mathbb{R}^d; X)$ its limit and identifying it with an element in $\mathscr{S}'(\mathbb{R}^d; X)$, we may define $u \in \mathscr{S}'(\mathbb{R}^d; X)$ by $u := J_{-s} f$. Then $J_s u = f$ belongs to $L^p(\mathbb{R}^d; X)$ and therefore $u \in H^{s,p}(\mathbb{R}^d; X)$ and

$$\|u_n - u\|_{H^{s,p}(\mathbb{R}^d;X)} = \|J_s u_n - f\|_{L^p(\mathbb{R}^d;X)} = 0.$$

This proves completeness.

The reader may easily check that for $s = 2$ the definition of $H^{2,p}(\mathbb{R}^d; X)$ given above coincides with the provisional one given in (5.45).

Proposition 5.6.3. *For all $s \in \mathbb{R}$ and $p \in [1, \infty)$, J_s is an isometry from $H^{s,p}(\mathbb{R}^d; X)$ onto $L^p(\mathbb{R}^d; X)$, with inverse J_{-s}. More generally, J_s is an isometry from $H^{r+s,p}(\mathbb{R}^d; X)$ onto $H^{r,p}(\mathbb{R}^d; X)$ for all $r \in \mathbb{R}$, with inverse J_{-s}.*

Proof. It is clear from the definitions that J_s maps $H^{s,p}(\mathbb{R}^d; X)$ to $L^p(\mathbb{R}^d; X)$ isometrically. Also, $J_s \circ J_{-s} = I$ on $\mathscr{S}'(\mathbb{R}^d; X)$. It follows from these facts that $J_s : H^{s,p}(\mathbb{R}^d; X) \to L^p(\mathbb{R}^d; X)$ is surjective and that $J_{-s}|_{L^p(\mathbb{R}^d;X)} : L^p(\mathbb{R}^d; X) \to H^{s,p}(\mathbb{R}^d; X)$ is its inverse. This proves the first assertion. The second assertion is obtained from it via the identification $J_{r+s} = J_s \circ J_r$. $\square$

Proposition 5.6.4. *For all $s \in \mathbb{R}$ and $p \in [1, \infty)$, we have the following dense and continuous embeddings:*

$$\mathscr{S}(\mathbb{R}^d; X) \hookrightarrow H^{s,p}(\mathbb{R}^d; X) \hookrightarrow \mathscr{S}'(\mathbb{R}^d; X).$$

Proof. For $u \in \mathscr{S}(\mathbb{R}^d; X)$ and $s_0 \in \mathbb{R}$ we may write $u = J_{-s} \circ j \circ J_s u$, where we view J_s as a continuous mapping on $\mathscr{S}(\mathbb{R}^d; X)$, J_{-s} as an isometry from $L^p(\mathbb{R}^d; X)$ to $H^{s,p}(\mathbb{R}^d; X)$, and $j : \mathscr{S}(\mathbb{R}^d; X) \to L^p(\mathbb{R}^d; X)$ is the inclusion mapping (which is continuous by Proposition 2.4.23). This factorisation shows that $\mathscr{S}(\mathbb{R}^d; X)$ is contained in $H^{s,p}(\mathbb{R}^d; X)$ and that the inclusion mapping is continuous. Since j has dense range and both J_s and J_s are bijective, this argument also gives the density of $\mathscr{S}(\mathbb{R}^d; X)$ in $H^{s,p}(\mathbb{R}^d; X)$.

The inclusion $H^{s,p}(\mathbb{R}^d; X) \subseteq \mathscr{S}'(\mathbb{R}^d; X)$ holds by the very definition of $H^{s,p}(\mathbb{R}^d; X)$. Its continuity follows from the continuity of $L^p(\mathbb{R}^d; X) \hookrightarrow \mathscr{S}'(\mathbb{R}^d; x)$ and the continuity of J_s and its inverse $J_s^{-1} = J_{-s}$ on $\mathscr{S}'(\mathbb{R}^d; X)$. The density of $H^{s,p}(\mathbb{R}^d; X)$ in $\mathscr{S}'(\mathbb{R}^d; X)$ follows from the density of $\mathscr{S}(\mathbb{R}^d; X)$ in $\mathscr{S}'(\mathbb{R}^d; X)$ (see Proposition 2.4.33). □

Our next aim is to show that for $s > 0$ the operators J_{-s} can be represented as convolution operators with a positive kernel, obtained from the Gaussian kernel by an integral transform.

Lemma 5.6.5. *For all $s > 0$, the function*

$$G_s(x) := \frac{1}{(4\pi)^{d/2}\Gamma(s/2)} \int_0^\infty e^{-t} e^{-|x|^2/4t} t^{(s-d)/2} \frac{\mathrm{d}t}{t}, \quad x \in \mathbb{R}^d,$$

belongs to $L^1(\mathbb{R}^d)$ and has norm one, and it satisfies

$$\widehat{G_s}(\xi) = (1 + 4\pi^2|\xi|^2)^{-s/2}, \quad \xi \in \mathbb{R}^d.$$

Proof. The fact that G_s is integrable is a simple consequence of Fubini's theorem.

As we have seen in Lemma 2.4.4, the Fourier transform of $e^{-|x|^2/4t}$ equals $(4t\pi)^{d/2} e^{-4t\pi^2|\xi|^2}$. Therefore, by Fubini's theorem, writing $\lambda = \lambda(\xi) = 1 + 4\pi^2|\xi|^2$ for brevity,

$$\begin{aligned}
\Big(\int_0^\infty e^{-t} e^{-|x|^2/4t} t^{(s-d)/2} \frac{\mathrm{d}t}{t}\Big)\widehat{\ }(\xi) &= \int_0^\infty (4\pi t)^{d/2} e^{-\lambda t} t^{(s-d)/2} \frac{\mathrm{d}t}{t} \\
&= \lambda^{-s/2} (4\pi)^{d/2} \int_0^\infty e^{-t} t^{s/2} \frac{\mathrm{d}t}{t} \\
&= \lambda^{-s/2} (4\pi)^{d/2} \Gamma(s/2).
\end{aligned}$$

This establishes the formula for $\widehat{G_s}$. Finally, from the fact that $G \geqslant 0$ we see that $\|G\|_{L^1(\mathbb{R}^d)} = \widehat{G_s}(0) = 1$. □

Proposition 5.6.6. *Let $p \in [1, \infty)$. For $s > 0$ the operator $J_{-s} : L^p(\mathbb{R}^d; X) \to H^{s,p}(\mathbb{R}^d; X)$ maps $L^p(\mathbb{R}^d; X)$ into itself. As an operator on $L^p(\mathbb{R}^d; X)$, it is a positive contraction given by convolution with G_s:*

$$J_{-s} f = G_s * f, \quad f \in L^p(\mathbb{R}^d; X).$$

As a consequence, for all $s_0 \geqslant s_1$ we have a continuous inclusion mapping $H^{s_0,p}(\mathbb{R}^d; X) \hookrightarrow H^{s_1,p}(\mathbb{R}^d; X)$, and this mapping is in fact contractive.

Proof. We have already seen that $J_{-s} = J_s^{-1}$ is an isometry from $L^p(\mathbb{R}^d; X)$ onto $H^{s,p}(\mathbb{R}^d; X)$. For functions f belonging to the dense subspace $\mathscr{S}(\mathbb{R}^d; X)$ of $L^p(\mathbb{R}^d; X)$ we have

$$J_{-s}f = ((1 + 4\pi^2|\cdot|^2)^{-s/2}\widehat{f})^{\smallsmile} = G_s * f.$$

By density, the identity $J_{-s}f = G_s * f$ extends to arbitrary $f \in L^p(\mathbb{R}^d; X)$. This proves that J_{-s} maps $L^p(\mathbb{R}^d; X)$ into itself and is a positive contraction on this space. In particular, we find $\|J_{s_1-s_0}f\|_{L^p(\mathbb{R}^d;X)} \leqslant \|f\|_{L^p(\mathbb{R}^d;X)}$ for $f \in L^p(\mathbb{R}^d; X)$. Now the required embedding result follows from Proposition 5.6.3. $\square$

Duality

We pause for a brief intermezzo on duality. The result proved below will not be needed in what follows, and the reader who is primarily interested in the proof of Theorem 5.6.1 may proceed to the next paragraph.

For $u \in H^{s,p}(\mathbb{R}^d; X)$ and $v \in H^{-s,p'}(\mathbb{R}^d; X^*)$, with $p \in [1, \infty)$, we can define a duality pairing by

$$\langle u, v\rangle_{s,p} := \langle J_s u, J_{-s} v\rangle,$$

using the pairing of $L^p(\mathbb{R}^d; X)$ and $L^{p'}(\mathbb{R}^d; X^*)$ on the right-hand side. Clearly,

$$\big|\langle u, v\rangle_{s,p}\big| \leqslant \|u\|_{H^{s,p}(\mathbb{R}^d;X)}\|v\|_{H^{-s,p'}(\mathbb{R}^d;X^*)}$$

which implies that we have a natural inclusion

$$H^{-s,p}(\mathbb{R}^d; X^*) \hookrightarrow H^{s,p}(\mathbb{R}^d; X)^*.$$

It is easily seen to be injective and contractive, and Proposition 1.3.1 tells us that under this identification $H^{-s,p}(\mathbb{R}^d; X^*)$ is norming for $H^{s,p}(\mathbb{R}^d; X)$. As an application of Theorem 1.3.10, we shall now prove that the identification is an isometric isomorphism when X^* has the Radon-Nikodým property (we recall that this is for example the case when X is reflexive, and in particular when X is UMD; see Theorems 1.3.21 and 4.3.3).

Proposition 5.6.7 (Duality). *If X is a Banach space whose dual X^* has the Radon-Nikodým property, then for all $p \in [1, \infty)$ and $s \in \mathbb{R}$ the above embedding establishes an isometric isomorphism*

$$H^{s,p}(\mathbb{R}^d; X)^* \eqsim H^{-s,p'}(\mathbb{R}^d; X^*).$$

Proof. After what has already been said, it remains to be shown that every functional $\varphi \in H^{s,p}(\mathbb{R}^d; X)^*$ can be represented by a tempered distribution v in $H^{-s,p'}(\mathbb{R}^d; X^*)$ of the same norm.

Since J_s is an isometry from $H^{s,p}(\mathbb{R}^d; X)$ onto $L^p(\mathbb{R}^d; X)$ we may define $\varphi_s \in L^p(\mathbb{R}^d; X)^*$ by $\varphi_s(u) := \varphi(J_{-s}u)$. By Theorem 1.3.10 there exists a function $g_s \in L^{p'}(\mathbb{R}^d;X^*)$ of norm $\|g_s\|_{L^{p'}(\mathbb{R}^d;X^*)} = \|\varphi_s\|_{L^p(\mathbb{R}^d;X)^*}$ such that $\varphi_s(f) = \langle f, g_s\rangle$ for all $f \in L^p(\mathbb{R}^d;X)$. Let $v \in H^{-s,p}(\mathbb{R}^d;X^*)$ be given by $v = J_s g_s$. Then v has the desired properties: for all $u \in H^{s,p}(\mathbb{R}^d;X)$ we find

$$\varphi(u) = \varphi_s(J_s u) = \langle J_s u, g_s\rangle = \langle J_s u, J_{-s}v\rangle = \langle u, v\rangle_{s,p}$$

and $\|v\|_{H^{-s,p'}(\mathbb{R}^d;X^*)} = \|g_s\|_{L^{p'}(\mathbb{R}^d;X^*)} = \|\varphi\|_{H^{s,p}(\mathbb{R}^d;X)^*}$. □

5.6.b Complex interpolation of Bessel potential spaces

Throughout this subsection we shall assume that X is a UMD space. We fix an exponent $p \in (1,\infty)$. Under these assumptions, for $\Re z \geqslant 0$ we may consider the Fourier multiplier operator on $L^p(\mathbb{R}^d;X)$ corresponding to the multiplier

$$m_{-z}(\xi) := (1 + 4\pi^2|\xi|^2)^{-z/2}, \quad \xi \in \mathbb{R}^d.$$

It is straightforward to verify that the Mihlin conditions are satisfied for this function. In fact, writing $z = s + it$, the multiplier $T_{m_{-s}}$ coincides with the operator J_{-s} when the latter is viewed as a bounded operator of norm $\leqslant 1$ on $L^p(\mathbb{R}^d;X)$ (cf. Proposition 5.6.3). Furthermore, in a similar way as in Corollary 5.5.11 one sees that $T_{m_{-it}} =: J_{-it}$ is bounded on $L^p(\mathbb{R}^d;X)$, with bound $\|J_{-it}\| = \|m_{-it}\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)} \leqslant C(1+|t|)$. Therefore $T_{m_{-s}} \circ T_{m_{-it}} = T_{m_{-z}} =: J_{-z}$ is bounded on $L^p(\mathbb{R}^d;X)$ as well, with

$$\|J_{-z}\| \leqslant C(1 + |\Im z|), \quad \Re z \geqslant 0. \tag{5.53}$$

Also note that

$$J_{-z_1} \circ J_{-z_2} = J_{-z_1-z_2}, \quad \Re z_1, \Re z_2 \geqslant 0.$$

Lemma 5.6.8. *Let X be a UMD space and let $p \in (1,\infty)$. For all $f \in L^p(\mathbb{R}^d;X)$ the mapping $z \mapsto J_{-z}f$ is continuous on $\{\Re z \geqslant 0\}$ and holomorphic on $\{\Re z > 0\}$ with respect to the norm of $L^p(\mathbb{R}^d;X)$.*

Proof. For Schwartz functions f, this is immediate from the properties of the Fourier transform, and the general case then follows by an approximation argument, using the bounds for J_{-z} just proved. □

We are now ready for the first main step towards the proof of Theorem 5.6.1:

Theorem 5.6.9 (Seeley). *Let X be a UMD space, and let $p \in (1,\infty)$ and $-\infty < s_0 < s_1 < \infty$. For all $0 < \theta < 1$ we have*

$$[H^{s_0,p}(\mathbb{R}^d;X), H^{s_1,p}(\mathbb{R}^d;X)]_\theta = H^{s_\theta,p}(\mathbb{R}^d;X)$$

with equivalent norms, with $s_\theta = (1-\theta)s_0 + \theta s_1$. In particular, for all $s \geqslant 0$ and $0 < \theta < 1$ we have

$$[L^p(\mathbb{R}^d;X), H^{s,p}(\mathbb{R}^d;X)]_\theta = H^{\theta s,p}(\mathbb{R}^d;X).$$

Proof. By using the isomorphism $J_{s_0} : H^{s_0,p}(\mathbb{R}^d; X) \to L^p(\mathbb{R}^d; X)$ it suffices to prove the second assertion.

Let us write $A_z := J_{-sz}$, and $L^p := L^p(\mathbb{R}^d; X)$ for brevity. Keeping in mind Proposition 5.6.3, the range space $A_1(L^p) = \{A_1 f : f \in L^p\}$ is precisely $H^{s,p}(\mathbb{R}^d; X)$ and similarly $A_\theta(L^p) = H^{\theta s,p}(\mathbb{R}^d; X)$. The definition of the $H^{s,p}$-norm translates into the suggestive identity $\|A_\theta f\|_{A_\theta(L^p)} = \|f\|_{L^p}$. With these notations, we wish to show that $[L^p, A_1(L^p)]_\theta = A_\theta(L^p)$ for all $0 < \theta < 1$.

Consider the strip $\mathbb{S} := \{z \in \mathbb{C} : \ 0 < \Re z < 1\}$. Fixing $f \in L^p$, we define $F : \overline{\mathbb{S}} \to L^p$ by $F(z) := e^{z^2-\theta^2} A_z f$. By Lemma 5.6.8, this function is continuous on $\overline{\mathbb{S}}$ and holomorphic in $\mathbb{S}$. Furthermore, by (5.53),

$$\begin{aligned} \|F(iy)\|_{L^p} &= e^{-y^2-\theta^2}\|A_{iy}f\|_{L^p} \leqslant c_s\|f\|_{L^p}, \\ \|F(1+iy)\|_{A_1(L^p)} &= e^{1-y^2-\theta^2}\|A_{1+iy}f\|_{A_1(L^p)} \\ &= e^{1-y^2-\theta^2}\|A_{iy}f\|_{L^p} \leqslant ec_s\|f\|_{L^p}, \end{aligned}$$

where $c_s := C\sup_{y\in\mathbb{R}}(1+s|y|)e^{-y^2}$, with C the constant in (5.53). By the definition of the complex interpolation spaces, this means that $F(\theta) = A_\theta f$ belongs to $[L^p, A_1(L^p)]_\theta$ with bound $\|A_\theta f\|_{[L^p,A_1(L^p)]_\theta} \leqslant ec_s\|f\|_{L^p}$. This gives the continuous inclusion $A_\theta(L^p) \subseteq [L^p, A_1(L^p)]_\theta$.

For the proof of the reverse inclusion suppose that $f \in [L^p, A_1(L^p)]_\theta$. This means that there exists a continuous function $F : \overline{\mathbb{S}} \to L^p$, holomorphic on $\mathbb{S}$, such that $F(\theta) = f$ and

$$\|F(iy)\|_{L^p} \leqslant 2\|f\|_{A_\theta(L^p)}, \quad \|F(1+iy)\|_{A_1(L^p)} \leqslant 2\|f\|_{A_\theta(L^p)}.$$

Define $G : \mathbb{S} \to L^p$ by $G(z) := e^{z^2-\theta^2} A_{1-z}F(z)$. Then, with c_s as before,

$$\begin{aligned} \|G(iy)\|_{A_1(L^p)} &= e^{-y^2-\theta^2}\|A_{iy}F(iy)\|_{L^p} \leqslant 2c_s\|f\|_{A_\theta(L^p)}, \\ \|G(1+iy)\|_{A_1(L^p)} &= e^{1-y^2-\theta^2}\|A_{iy}F(1+iy)\|_{A_1(L^p)} \\ &= e^{1-y^2-\theta^2}\|A_{iy}A_{-1}F(1+iy)\|_{L^p} \\ &\leqslant ec_s\|A_{-1}F(1+iy)\|_{L^p} \\ &\leqslant ec_s\|F(1+iy)\|_{A_1(L^p)} \\ &\leqslant 2ec_s\|f\|_{A_\theta(L^p)}. \end{aligned}$$

Hence, by the three lines lemma (Lemma 2.2.2),

$$\|f\|_{A_\theta(L^p)} = \||A_{-\theta}f\|_{L^p} = \|A_{1-\theta}f\|_{A_1(L^p)} = \|G(\theta)\|_{A_1(L^p)} \leqslant 2ec_s\|f\|_{A_\theta(L^p)}.$$

□

5.6.c Coincidence of Sobolev and Bessel potential spaces

We now turn to the second step of our programme, namely the proof that for UMD spaces X and $p \in (1,\infty)$ we have $W^{k,p}(\mathbb{R}^d; X) = H^{k,p}(\mathbb{R}^d; X)$ with equivalent norms.

We start with a preliminary observation. For even integers $k = 2\ell \geqslant 0$ and Schwartz functions $f \in \mathscr{S}(\mathbb{R}^d; X)$ we have

$$((1 + 4\pi^2|\cdot|^2)^{k/2}\widehat{f})^{\smile} = (1 - \Delta)^\ell f = \sum_{|\alpha| \leqslant k} c_\alpha \partial^\alpha f$$

with suitable coefficients $c_\alpha = c_{\alpha,k}$. Since the Schwartz functions are dense in both $H^{2\ell,p}(\mathbb{R}^d; X)$ and $W^{2\ell,p}(\mathbb{R}^d; X)$ for $p \in [1, \infty)$, this results in a continuous inclusion

$$W^{2\ell,p}(\mathbb{R}^d; X) \subseteq H^{2\ell,p}(\mathbb{R}^d; X), \quad p \in [1, \infty).$$

No UMD assumption on X is required here. In dimension $d = 1$, due to the absence of mixed partial derivatives the converse embedding also holds.

Proposition 5.6.10. *Let X be an arbitrary Banach space and let $p \in [1, \infty)$. For all even integers $k = 2\ell$ with $\ell \geqslant 1$ we have*

$$W^{2\ell,p}(\mathbb{R}; X) = H^{2\ell,p}(\mathbb{R}; X) \quad \textit{with equivalent norms.}$$

Proof. It remains to be shown that $H^{2\ell,p}(\mathbb{R}; X)$ is contained in $W^{2\ell,p}(\mathbb{R}; X)$ continuously. By density (see Proposition 5.6.4) it suffices to prove that for all $j \in \{0, \dots, 2\ell\}$

$$\|f^{(j)}\|_p \leqslant C\|f\|_{H^{2\ell,p}(\mathbb{R};X)}, \quad f \in \mathscr{S}(\mathbb{R}; X). \tag{5.54}$$

First we check (5.54) for even $j = 2n$. Since $-\Delta J_{-2} = I - J_{-2}$,

$$(-\Delta)^n J_{-2\ell} = [(-\Delta)J_{-2}]^n J_{-(2\ell-2n)} = (I - J_{-2})^n J_{-(2\ell-2n)}.$$

Therefore, by Proposition 5.6.6,

$$\|f^{(2n)}\|_p = \|(-\Delta)^n f\|_p \leqslant 2^n \|J_{2\ell} f\|_p = 2^n \|f\|_{H^{2\ell,p}(\mathbb{R};X)}.$$

For odd $j = 2n + 1 \leqslant 2\ell$, the reader may check the formula

$$f^{2n+1}(x) = -\int_0^\infty \zeta(t) f^{2n+2}(x + t)\,\mathrm{d}t + \int_0^\infty \zeta''(t) f^{2n}(x + t)\,\mathrm{d}t,$$

where $\zeta \in C_{\mathrm{c}}^\infty(\mathbb{R})$ is any test function satisfying $\zeta(0) = 1$ and $\zeta'(0) = 0$. Therefore,

$$\|f^{(2n+1)}\|_p \leqslant \|\zeta\|_1 \|f^{(2n+2)}\|_p + \|\zeta''\|_1 \|f^{(2n)}\|_p \leqslant C\|f\|_{H^{2\ell,p}(\mathbb{R};X)},$$

with $C := 2^{n+1}\|\zeta\|_1 + 2^n\|\zeta''\|_1$. □

It is a non-trivial fact that for UMD spaces X and $p \in (1, \infty)$, this equality extends to dimensions $d \geqslant 2$ and arbitrary integers $k \geqslant 1$:

Theorem 5.6.11. *Let X be a UMD space. For all integers $k \geqslant 1$ and $p \in (1,\infty)$ we have*

$$W^{k,p}(\mathbb{R}^d;X) = H^{k,p}(\mathbb{R}^d;X) \quad \textit{with equivalent norms.}$$

Together with Theorem 5.6.9, Theorem 5.6.11 provides a proof of Theorem 5.6.1. A partial converse to Theorem 5.6.11 will be proved in Theorem 5.6.12.

Proof. Fix k and p in the indicated ranges. We begin by proving the inequality

$$\|f\|_{W^{k,p}(\mathbb{R}^d;X)} \leqslant C\|f\|_{H^{k,p}(\mathbb{R}^d;X)}$$

for functions $f \in \mathscr{S}(\mathbb{R}^d;X)$. To this end, it suffices to prove that

$$\|\partial^\alpha f\|_{L^p(\mathbb{R}^d;X)} \leqslant C_\alpha\|f\|_{H^{k,p}(\mathbb{R}^d;X)}$$

for each multi-index $|\alpha| \leqslant k$. In terms of $g := ((1+4\pi^2|\cdot|^2)^{k/2}\widehat{f})^{\vee}$, this inequality can be restated as

$$\big\|((-2\pi i\xi)^\alpha(1+4\pi^2|\xi|^2)^{-k/2}\widehat{g}(\xi))^{\widecheck{}}\big\|_{L^p(\mathbb{R}^d;X)} \leqslant C_\alpha\|g\|_{L^p(\mathbb{R}^d;X)}.$$

This will follow once we know that the multiplier operator T_m is bounded on $L^p(\mathbb{R}^d;X)$, where

$$m(\xi) := M_{\alpha,k}(\xi) := \xi^\alpha/(1+4\pi^2|\xi|^2)^{k/2}.$$

To verify the condition of the Mihlin multiplier theorem for this function, we check by induction on $|\beta|$ that for any multi-index β, the function $\partial^\beta m(\xi)$ is a linear combination of functions of the form $\xi^\theta(1+4\pi^2|\xi|^2)^{-n/2}$ for some multi-index θ and number n such that $n-|\theta| = k-|\alpha|+|\beta|$. From this it follows that $|\xi|^{|\beta|}|\partial^\beta m(\xi)|$ is uniformly bounded, as needed for Mihlin's conditions.

To prove the reverse estimate

$$\|f\|_{H^{k,p}(\mathbb{R}^d;X)} \leqslant C\|f\|_{W^{k,p}(\mathbb{R}^d;X)}$$

we again apply the Mihlin theorem. Let $\rho \in C_{\mathrm{c}}^\infty(\mathbb{R})$ be a function satisfying $\rho(t) \geqslant 0$ for $t \geqslant 0$, $\rho(t) = 1$ for $t \geqslant 1$, $\rho(t) = 0$ for $0 \leqslant t \leqslant \frac{1}{2}$, and $\rho(-t) = -\rho(t)$ for all $t \in \mathbb{R}$. Put

$$m(\xi) := \frac{(1+4\pi^2|\xi|^2)^{k/2}}{1+\sum_{j=1}^d(2\pi\rho(\xi_j)\xi_j)^k}.$$

Clearly, $\rho(\xi_j)\xi_j \geqslant 0$, so that the denominator, which we shall call $D(\xi)$, satisfies $D(\xi) \geqslant 1$. It is easy to check that $m \in L^\infty(\mathbb{R}^d) \cap C^\infty(\mathbb{R}^d)$. For non-zero multi-indices β, $\xi^\beta\partial^\beta m(\xi)$ is a linear combination of functions of the form

$$\frac{(1+4\pi^2|\xi|^2)^{\frac{k}{2}-a}\phi(\xi)}{D(\xi)} \quad \text{and} \quad \frac{(1+4\pi^2|\xi|^2)^{b/2}\psi(\xi)}{D(\xi)^c}, \tag{5.55}$$

where a and c are non-negative integers, $b \in \mathbb{Z}$, $\phi \in C^\infty(\mathbb{R}^d \setminus \{0\})$ homogeneous of order $2a$, and $\psi \in C_c^\infty(\mathbb{R}^d)$. From this it is again evident that Mihlin's conditions are satisfied by m.

Let $g \in \mathscr{S}(\mathbb{R}^d; X)$ be the function whose Fourier transform is given as

$$\widehat{g}(\xi) := \Big(1 + \sum_{j=1}^d (2\pi\rho(\xi_j)\xi_j)^k\Big)\widehat{f}(\xi) = \widehat{f}(\xi) + i^k \sum_{j=1}^d (\rho(\xi_j))^k \widehat{\partial_j^k f}(\xi).$$

Then $(1 + 4\pi^2|\xi|^2)^{k/2}\widehat{f}(\xi) = m(\xi)\widehat{g}(\xi)$ and therefore

$$\begin{aligned}\|f\|_{H^{k,p}(\mathbb{R}^d;X)} &= \|T_m g\|_{L^p(\mathbb{R}^d;X)} \leqslant C\|g\|_{L^p(\mathbb{R}^d;X)} \\ &\leqslant C\|f\|_{L^p(\mathbb{R}^d;X)} + CC' \sum_{j=1}^d \|T_{j,\rho^k}(\partial_j^k f)\|_{L^p(\mathbb{R}^d;X)},\end{aligned}$$

where T_{j,ρ^k} is the Fourier multiplier in the jth variable with multiplier ρ^k. It is simple to check that ρ^k satisfies the Mihlin condition. It follows that

$$\|T_{j,\rho^k}(\partial_j^k f)\|_{L^p(\mathbb{R}^d;X)} \leqslant C''\|\partial_j^k f\|_{L^p(\mathbb{R}^d;X)}.$$

Collecting the various estimates, this completes the proof. $\square$

Theorem 5.6.11 admits the following partial converse.

Theorem 5.6.12. *Let X be a Banach space. Suppose that for some $k \geqslant 1$ and $p \in (1,\infty)$ there is a constant $C > 0$ such that for all $f \in \mathscr{S}(\mathbb{R}^d; X)$ the estimate*

$$\|f\|_{W^{k,p}(\mathbb{R}^d;X)} \leqslant C\|f\|_{H^{k,p}(\mathbb{R}^d;X)} \tag{5.56}$$

holds, then the following statements hold:

(1) *If k is odd, then X is a UMD space and its Hilbert transform constant satisfies $\hbar_{p,X} \leqslant c_{d,k}C$, where $c_{d,1} = 1$.*
(2) *If k is even and $d \geqslant 2$, then the second order Riesz transform $(R_1)^2$ satisfies*
$$\|(R_1)^2\|_{\mathscr{L}(L^p(\mathbb{R}^d;X))} \leqslant c_{d,k}C,$$
where $c_{d,2} = 1$.

In the second case it also follows that X is a UMD space. We have already seen that the boundedness of the Hilbert transform implies the UMD property; that the boundedness of $(R_1)^2$ works equally well is discussed in the Notes.

Proof. First assume that k is odd, say $k = 2\ell+1$ with $\ell \geqslant 0$. Writing $\partial_1(I-\Delta)^\ell$ as a sum of mixed derivatives of order $\leqslant k$, the assumption implies that

$$\|\partial_1(1-\Delta)^\ell g\|_{L^p(\mathbb{R}^d;X)} \leqslant c_{d,k}\|g\|_{W^{k,p}(\mathbb{R}^d;X)} \leqslant c_{d,k}C\|g\|_{H^{k,p}(\mathbb{R}^d;X)} \tag{5.57}$$

for all $g \in \mathscr{S}(\mathbb{R}^d; X)$, where C is the constant of (5.56). Note that for $k = 1$, we have $\ell = 0$ and $c_{d,k} = 1$.

Let $f \in \mathscr{S}(\mathbb{R}^d; X)$ be arbitrary. For $g := ((1 + 4\pi^2|\cdot|^2)^{-k/2}\widehat{f})^{\smile}$, (5.57) can be restated as

$$\|T_m f\|_{L^p(\mathbb{R}^d;X)} \leqslant c_{d,k} C \|f\|_{L^p(\mathbb{R}^d;X)}$$

for the multiplier $m(\xi) := 2\pi\xi_1/(1 + 4\pi^2|\xi|^2)^{1/2}$. Thus $m \in \mathfrak{M}L^p(\mathbb{R}^d; X)$ with norm $\|m\|_{\mathfrak{M}L^p(\mathbb{R}^d;X)} \leqslant c_{d,k} C$. We now apply the result of Proposition 5.5.2 to arrive at (1).

If k is even, as in the previous case one obtains

$$\|T_m f\|_{L^p(\mathbb{R}^d;X)} \leqslant c_{d,k} C \|f\|_{L^p(\mathbb{R}^d;X)},$$

where $c_{d,2} = 1$, this time using the multiplier $m(\xi) := 4\pi^2\xi_1^2/(1 + 4\pi^2|\xi|^2)$. As $N \to \infty$, the dilation $m(N\cdot)$ (which are multipliers by Proposition 5.3.8) tend pointwise to the multiplier $\xi_1^2/|\xi|^2$ of $(R_1)^2$, and thus (2) follows. □

5.7 Transference and Fourier multipliers on $\mathbb{T}^d$

Pursuing the same analogy of Fourier transforms and Fourier series that led from the Hilbert transform H to its periodic analogue $\widetilde{H}$, we may also consider periodic Fourier multipliers given by

$$\widetilde{T}_m f = \sum_{k \in \mathbb{Z}^d} m(k) \widehat{f}(k) e_k$$

for any X-valued trigonometric polynomial $f = \sum_{k \in \mathbb{Z}^d} \widehat{f}(k) e_k$, and inquire about the inequality

$$\|\widetilde{T}_m f\|_{L^p(\mathbb{T}^d;Y)} \leqslant C \|f\|_{L^p(\mathbb{T}^d;X)},$$

which then allows the extension of $\widetilde{T}_m$ as a bounded linear operator from $L^p(\mathbb{T}^d; X)$ to $L^p(\mathbb{T}^d; Y)$. To address this problem, we could follow the procedure already familiar from the case of $\mathbb{R}^d$, going through parallel and analogous intermediate steps. Rather than boring the reader with such repetition, we choose a different approach of some independent interest by setting up a general machine by which results already achieved on $\mathbb{R}^d$ may be *transferred* to $\mathbb{T}^d$, and also vice versa to some extent. This has the benefit of applying not only to the multiplier theorems that we have handled in the present treatment, but also many other similar results that the readers may find in the literature or prove on their own: rather than building a parallel theory on $\mathbb{R}^d$ and $\mathbb{T}^d$ every time, it is often enough to treat just one of them by hand and leave the rest for the transference.

Given $m = (m(k))_{k \in \mathbb{Z}^d} \in \ell^\infty(\mathbb{Z}^d; \mathscr{L}(X, Y))$, we define

$$\widetilde{T}_m : f \mapsto \sum_{k\in\mathbb{Z}^d} m(k)\widehat{f}(k)e_k$$

from trigonometric polynomials with X-valued coefficients to those with Y-valued coefficients. If $\|\widetilde{T}_m f\|_{L^p(\mathbb{T}^d;Y)} \leqslant C\|f\|_{L^p(\mathbb{T}^d;X)}$ for all trigonometric polynomials $f \in L^p(\mathbb{T}^d;X)$, we say that m is a(n operator-valued) Fourier multiplier from $L^p(\mathbb{T}^d;X)$ into $L^p(\mathbb{T}^d;Y)$. We also make a similar definition with L^p replaced by L^p_0 throughout; for this latter case, the value of $m(0)$ is of course immaterial, and we may also say that $(m(k))_{k\in\mathbb{Z}^d\setminus\{0\}}$ is a multiplier.

We also introduce the multiplier norms

$$\begin{aligned}\|m\|_{\mathfrak{M}L^p(\mathbb{T}^d;X,Y)} &:= \|\widetilde{T}_m\|_{\mathscr{L}(L^p(\mathbb{T}^d;X),L^p(\mathbb{T}^d;Y))},\\ \|m\|_{\mathfrak{M}L^p_0(\mathbb{T}^d;X,Y)} &:= \|\widetilde{T}_m\|_{\mathscr{L}(L^p_0(\mathbb{T}^d;X),L^p_0(\mathbb{T}^d;Y))}.\end{aligned}$$

5.7.a Transference from $\mathbb{R}^d$ to $\mathbb{T}^d$

We now turn to the details of the promised transference machine, starting with the direction from $\mathbb{R}^d$ to $\mathbb{T}^d$.

Proposition 5.7.1 (Transference from $\mathbb{R}^d$ to $\mathbb{T}^d$). *Let X and Y be Banach spaces and let $p \in (1,\infty)$. Let $m \in L^\infty(\mathbb{R}^d;\mathscr{L}(X,Y))$ be a Fourier multiplier from $L^p(\mathbb{R}^d;X)$ to $L^p(\mathbb{R}^d;Y)$. Suppose that for all $x \in X$ the point $k \in \mathbb{Z}^d$ is a Lebesgue point of $\xi \mapsto m(\xi)x$, and set $m_k x := m(k)x$. Then $(m_k)_{k\in\mathbb{Z}^d}$ is a Fourier multiplier from $L^p(\mathbb{T}^d;X)$ to $L^p(\mathbb{T}^d;Y)$, and in fact*

$$\|\widetilde{T}_{(m_k)_{k\in\mathbb{Z}^d}}\|_{\mathscr{L}(L^p(\mathbb{T}^d;X),L^p(\mathbb{T}^d;Y))} \leqslant \|T_m\|_{\mathscr{L}(L^p(\mathbb{R}^d;X),L^p(\mathbb{R}^d;Y))}.$$

The proof rests on the following lemma relating the action of the periodic and non-periodic multipliers.

Lemma 5.7.2. *Let $\phi,\psi \in \mathscr{S}(\mathbb{R}^d)$ be radial functions which satisfy*

$$\int_{\mathbb{R}^d} \widehat{\phi}\widecheck{\psi}\,\mathrm{d}\xi = 1.$$

Then, in the situation of Proposition 5.7.1, the equality

$$\langle T_{(m_k)_{k\in\mathbb{Z}^d}} f, g\rangle = \lim_{\varepsilon\downarrow 0} \varepsilon^d \langle T_m\big(\phi(\varepsilon\cdot)f\big), \psi(\varepsilon\cdot)g\rangle$$

holds for all trigonometric polynomials $f : \mathbb{T}^d \to X$, $g : \mathbb{T}^d \to Y^*$.

Proof. We first check that

$$\lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon^d}\int_{\mathbb{R}^d} \widehat{\phi}\Big(\frac{\xi-k}{\varepsilon}\Big)\widecheck{\psi}\Big(\frac{\xi-\ell}{\varepsilon}\Big)m(\xi)x\,\mathrm{d}\xi = \delta_{k\ell}m_k x, \quad x \in X. \tag{5.58}$$

For $k = \ell$, this is immediate from Theorem 2.3.8. If $k \neq \ell$, then

$$\frac{1}{\varepsilon^d}\int_{\mathbb{R}^d}\widehat{\phi}(\frac{\xi-k}{\varepsilon})\widecheck{\psi}(\frac{\xi-\ell}{\varepsilon})m(\xi)x\,\mathrm{d}\xi=\int_{\mathbb{R}^d}\widehat{\phi}(\eta)\widecheck{\psi}(\eta+\frac{k-\ell}{\varepsilon})m(k+\varepsilon\eta)x\,\mathrm{d}\eta,$$

where the first factor integrable, the second tends to zero as $\varepsilon\downarrow 0$ and is dominated by $\|\widecheck{\psi}\|_\infty$, and the third is dominated by $\|m(\cdot)x\|_\infty$. Hence the integral converges to zero as $\varepsilon\downarrow 0$ by dominated convergence, and this completes the verification of (5.58).

Now let $f:=\sum_{k\in\mathbb{Z}^d}a_k e_k$ and $g:=\sum_{\ell\in\mathbb{Z}^d}b_\ell e_\ell$, with $e_k(t)=e^{2\pi i k\cdot t}$, $a_k\in X$ and $b_\ell\in X^*$ be trigonometric polynomials, with finitely many non-zero a_k and b_ℓ only. Then

$$\begin{aligned}\langle\widetilde{T}_{(m_k)_{k\in\mathbb{Z}^d}}f,g\rangle&=\sum_{k,\ell\in\mathbb{Z}^d}\langle m_k a_k,b_\ell\rangle\int_{\mathbb{T}^d}e_k e_\ell\,\mathrm{d}t=\sum_{k\in\mathbb{Z}^d}\langle m_k a_k,b_{-k}\rangle\\&=\sum_{k,\ell\in\mathbb{Z}^d}\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon^d}\int_{\mathbb{R}^d}\langle m(\xi)a_k,b_{-\ell}\rangle\widehat{\phi}(\frac{\xi-k}{\varepsilon})\widecheck{\psi}(\frac{\xi-\ell}{\varepsilon})\,\mathrm{d}\xi\\&=\lim_{\varepsilon\downarrow 0}\varepsilon^d\int_{\mathbb{R}^d}\Big\langle m(\xi)\sum_{k\in\mathbb{Z}^d}a_k\frac{1}{\varepsilon^d}\widehat{\phi}(\frac{\xi-k}{\varepsilon}),\sum_{\ell\in\mathbb{Z}^d}b_{-\ell}\frac{1}{\varepsilon^d}\widecheck{\psi}(\frac{\xi-\ell}{\varepsilon})\Big\rangle\,\mathrm{d}\xi.\end{aligned}$$

Observing that

$$\frac{1}{\varepsilon^d}\widehat{\phi}(\frac{\cdot-k}{\varepsilon})=\big(\phi(\varepsilon\cdot)e_k\big)\widehat{\ },\qquad\frac{1}{\varepsilon^d}\widecheck{\psi}(\frac{\cdot-\ell}{\varepsilon})=\big(\psi(\varepsilon\cdot)e_{-\ell}\big)\widecheck{\ },$$

we have shown that

$$\begin{aligned}\langle\widetilde{T}_{(m_k)_{k\in\mathbb{Z}^d}}f,g\rangle&=\lim_{\varepsilon\downarrow 0}\varepsilon^d\Big\langle m\Big(\phi(\varepsilon\cdot)\sum_{k\in\mathbb{Z}^d}a_k e_k\Big)\widehat{\ },\Big(\psi(\varepsilon\cdot)\sum_{\ell\in\mathbb{Z}^d}b_{-\ell}e_{-\ell}\Big)\widecheck{\ }\Big\rangle\\&=\lim_{\varepsilon\downarrow 0}\varepsilon^d\Big\langle m(\phi(\varepsilon\cdot)f)\widehat{\ },(\psi(\varepsilon\cdot)g)\widecheck{\ }\Big\rangle\\&=\lim_{\varepsilon\downarrow 0}\varepsilon^d\Big\langle T_m(\phi(\varepsilon\cdot)f),\psi(\varepsilon\cdot)g\Big\rangle,\end{aligned}$$

which was the claim. □

A very simple relation between norms on $\mathbb{R}^d$ and $\mathbb{T}^d$ is the following. Indeed, it is essentially just a restatement of Lemma 5.2.6 with some superscript-d's added, but due to its shortness we may just as well reproduce the full detail here:

Lemma 5.7.3. *For $\phi\in\mathscr{S}(\mathbb{R}^d)$ and $f\in L^p(\mathbb{T}^d;X)$,*

$$\lim_{\varepsilon\downarrow 0}\varepsilon^{d/p}\|\phi(\varepsilon\cdot)f\|_{L^p(\mathbb{R}^d;X)}=\|\phi\|_{L^p(\mathbb{R}^d)}\|f\|_{L^p(\mathbb{T}^d;X)}.$$

Proof. Rearranging and using the periodicity of f, we have

$$\varepsilon^d\int_{\mathbb{R}^d}\|\phi(\varepsilon x)f(x)\|^p\,\mathrm{d}x=\int_{[0,1)^d}\Big(\varepsilon^d\sum_{k\in\mathbb{Z}^d}|\phi(\varepsilon(x+k))|^p\Big)\|f(x)\|^p\,\mathrm{d}x,$$

where the quantity in parentheses is a Riemann sum of the function $|\phi(\cdot)|^p$. It is also bounded uniformly in x and ε. Hence the assertion follows from the existence of the Riemann integral and the dominated convergence theorem. □

Now we are ready for the proof of Proposition 5.7.1:

Proof of Proposition 5.7.1. We use the auxiliary functions $\phi(x) := e^{-\pi|x|^2/p}$, $\psi(x) := e^{-\pi|x|^2/p'}$, which clearly satisfy

$$1 = \|\phi\|_p = \|\psi\|_{p'} = \int_{\mathbb{R}^d} \phi\psi \,\mathrm{d}x = \int_{\mathbb{R}^d} \widehat{\phi}\widecheck{\psi} \,\mathrm{d}\xi.$$

By Lemmas 5.7.2 and 5.7.3, we know that

$$\begin{aligned}
|\langle \widetilde{T}_{(m_k)_{k\in\mathbb{Z}^d}} f, g\rangle| &= \lim_{\varepsilon\downarrow 0} \varepsilon^d |\langle T_m\big(\phi(\varepsilon\cdot)f\big), \psi(\varepsilon\cdot)g\rangle| \\
&\leqslant \lim_{\varepsilon\downarrow 0} \|T_m\| \varepsilon^{d/p} \|\phi(\varepsilon\cdot)f\|_{L^p(\mathbb{R}^d;X)} \varepsilon^{d/p'} \|\psi(\varepsilon\cdot)g\|_{L^{p'}(\mathbb{R}^d;Y^*)} \\
&= \|T_m\| \|\phi\|_{L^p(\mathbb{R}^d)} \|\psi\|_{L^{p'}(\mathbb{R}^d)} \|f\|_{L^p(\mathbb{T}^d;X)} \|g\|_{L^{p'}(\mathbb{T}^d;Y^*)},
\end{aligned}$$

where

$$\|T_m\| := \|T_m\|_{\mathscr{L}(L^p(\mathbb{R}^d;X),L^p(\mathbb{R}^d;Y))}.$$

Since $\|\phi\|_p = \|\psi\|_{p'} = 1$, Proposition 5.7.1 has been established. □

5.7.b Transference from $\mathbb{T}^d$ to $\mathbb{R}^d$

The transference results in this direction will be slightly more restricted, essentially by necessity, since it is clear that the restriction to lattice points, $(m(k))_{k\in\mathbb{Z}}$, cannot possibly contain enough information to recover the boundedness properties of an arbitrary multiplier $m \in L^\infty(\mathbb{R}^d; \mathscr{L}(X,Y))$. Nevertheless, the statements that we prove will be quite sufficient for our needs. On the one hand, we have already given direct proof of several multiplier results on $\mathbb{R}^d$, so that our main application of transference is going from there to $\mathbb{T}^d$. On the other hand, the reverse direction of transference is mostly used for proving the sharpness of some estimates, and such considerations are typically restricted to multiplier operators with some additional structure and regularity.

It is here in order to recall that

$$L^p_0(\mathbb{T}^d; X) := \Big\{ f \in L^p(\mathbb{T}^d; X) : \int_{\mathbb{T}^d} f \,\mathrm{d}t = 0 \Big\}.$$

Proposition 5.7.4 (Transference from $\mathbb{T}^d$ to $\mathbb{R}^d$).

$$m \in C(\mathbb{R}^d \setminus \{0\}; \mathscr{L}(X,Y)),$$

and suppose that $(m(\varepsilon k))_{k\in\mathbb{Z}^d\setminus\{0\}}$ *is a Fourier multiplier from* $L^p_0(\mathbb{T}^d;X)$ *to* $L^p_0(\mathbb{T}^d;Y)$*, having uniformly bounded multiplier norm for some sequence of numbers* $\varepsilon\downarrow 0$*. Then* m *is a Fourier multiplier for* $L^p(\mathbb{R}^d;X)$*, and in fact*

$$\|T_m\|_{\mathscr{L}(L^p(\mathbb{R}^d;X),L^p(\mathbb{R}^d;Y))} \leqslant \liminf_{\varepsilon\downarrow 0} \|\widetilde{T}_{(m(\varepsilon k))_{k\in\mathbb{Z}^d\setminus\{0\}}}\|_{\mathscr{L}(L^p_0(\mathbb{T}^d;X),\mathscr{L}(L^p_0(\mathbb{T}^d;Y))}.$$

This is based on the following approximation result.

Lemma 5.7.5. *For* $f\in\mathscr{S}(\mathbb{R}^d;X)$*,*

$$\lim_{\varepsilon\downarrow 0}\varepsilon^{d/p'}\Big\|\sum_{k\in\mathbb{Z}^d}\widehat{f}(\varepsilon k)e_k\Big\|_{L^p(\mathbb{T}^d;X)} = \|f\|_{L^p(\mathbb{R}^d;X)}.$$

Proof. We begin with the Poisson summation formula

$$\sum_{k\in\mathbb{Z}^d}\widehat{f}(\varepsilon k)e_k(t) = \sum_{k\in\mathbb{Z}^d}\widehat{f}_\varepsilon(k)e_k(t) = \sum_{k\in\mathbb{Z}^d} f_\varepsilon(t+k) = \sum_{k\in\mathbb{Z}^d}\frac{1}{\varepsilon^d}f\big(\frac{t}{\varepsilon}+\frac{k}{\varepsilon}\big).$$

Then, making the change of variable $u := t/\varepsilon$,

$$\begin{aligned}
\varepsilon^{d/p'}\Big\|\sum_{k\in\mathbb{Z}^d}\widehat{f}(\varepsilon k)e_k\Big\|_{L^p(\mathbb{T}^d;X)} &= \varepsilon^{-d/p}\Big\|\sum_{k\in\mathbb{Z}^d} f\big(\frac{\cdot}{\varepsilon}+\frac{k}{\varepsilon}\big)\Big\|_{L^p(\mathbb{T}^d;X)}\\
&= \Big\|\sum_{k\in\mathbb{Z}^d} f\big(\,\cdot+\frac{k}{\varepsilon}\big)\Big\|_{L^p(\varepsilon^{-1}\mathbb{T}^d;X)}\\
&= \|f\|_{L^p(\varepsilon^{-1}\mathbb{T}^d;X)} + O\Big(\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\Big\|f\big(\,\cdot+\frac{k}{\varepsilon}\big)\Big\|_{L^p(\varepsilon^{-1}\mathbb{T}^d;X)}\Big).
\end{aligned}$$

Clearly the first term tends to $\|f\|_{L^p(\mathbb{R}^d;X)}$ as $\varepsilon\downarrow 0$, and it suffices to check that the error term goes to zero.

For $k\neq 0$, we have $|f(u+k/\varepsilon)|\leqslant C(|k|/\varepsilon)^{-N}$ for $u\in\varepsilon^{-1}\mathbb{T}^d$, and hence $\|f(\cdot+k/\varepsilon)\|_{L^p(\varepsilon^{-1}\mathbb{T}^d;X)}\leqslant C\varepsilon^{N-d/p}|k|^{-N}$. With $N>d>d/p$, it follows that

$$\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\Big\|f\big(\,\cdot+\frac{k}{\varepsilon}\big)\Big\|_{L^p(\varepsilon^{-1}\mathbb{T}^d;X)} \leqslant C\varepsilon^{N-d/p}\sum_{k\in\mathbb{Z}^d\setminus\{0\}}|k|^{-N}\leqslant C\varepsilon^{N-d/p}\to 0,$$

as required. □

Proof of Proposition 5.7.4. With the help of Lemma 5.7.5, this goes fast. Let $f\in\check{\mathscr{D}}(\mathbb{R}^d\setminus\{0\};X)$, $g\in\check{\mathscr{D}}(\mathbb{R}^d\setminus\{0\};Y^*)$, recalling that such functions are dense in $L^p(\mathbb{R}^d;X)\cap\check{L}^1(\mathbb{R}^d;X)$ and $L^{p'}(\mathbb{R}^d;Y^*)$, respectively (Proposition 2.4.23). The function $\langle m(\cdot)\widehat{f}(\cdot),\breve{g}(\cdot)\rangle$, being continuous and compactly supported, is Riemann integrable, and hence

$$\langle g,T_m f\rangle = \int_{\mathbb{R}^d}\langle m(\xi)\widehat{f}(\xi),\breve{g}(\xi)\rangle\,\mathrm{d}\xi$$

$$
\begin{aligned}
&= \lim_{\varepsilon \downarrow 0} \varepsilon^d \sum_{k \in \mathbb{Z}^d} \langle m(\varepsilon k) \widehat{f}(\varepsilon k), \breve{g}(\varepsilon k) \rangle \\
&= \lim_{\varepsilon \downarrow 0} \varepsilon^d \sum_{k, \ell \in \mathbb{Z}^d} \langle m(\varepsilon k) \widehat{f}(\varepsilon k), \widehat{g}(\varepsilon \ell) \rangle \int_{\mathbb{T}^d} e_k e_\ell \, \mathrm{d}t \\
&= \lim_{\varepsilon \downarrow 0} \varepsilon^d \Big\langle \widetilde{T}_{(m(\varepsilon k))_{k \in \mathbb{Z}^d \setminus \{0\}}} \sum_{k \in \mathbb{Z}^d} \widehat{f}(\varepsilon k) e_k, \sum_{\ell \in \mathbb{Z}^d} \widehat{g}(\varepsilon \ell) e_\ell \Big\rangle.
\end{aligned}
$$

Thus, by Lemma 5.7.5,

$$
\begin{aligned}
|\langle T_m f, g \rangle| \leqslant \liminf_{\varepsilon \downarrow 0} & \big\| \widetilde{T}_{(m(\varepsilon k))_{k \in \mathbb{Z}^d \setminus \{0\}}} \big\| \\
& \times \varepsilon^{d/p} \Big\| \sum_{k \in \mathbb{Z}^d} \widehat{f}(\varepsilon k) e_k \Big\|_{L^p(\mathbb{T}^d; X)} \times \varepsilon^{d/p'} \Big\| \sum_{\ell \in \mathbb{Z}^d} \widehat{g}(\varepsilon \ell) e_\ell \Big\|_{L^{p'}(\mathbb{T}^d; Y^*)} \\
= \liminf_{\varepsilon \downarrow 0} & \big\| \widetilde{T}_{(m(\varepsilon k))_{k \in \mathbb{Z}^d \setminus \{0\}}} \big\| \|f\|_{L^p(\mathbb{R}^d; X)} \|g\|_{L^{p'}(\mathbb{R}^d; X^*)},
\end{aligned}
$$

where

$$
\|\widetilde{T}_{(m(\varepsilon k))_{k \in \mathbb{Z}^d \setminus \{0\}}}\| := \|\widetilde{T}_{(m(\varepsilon k))_{k \in \mathbb{Z}^d \setminus \{0\}}}\|_{\mathscr{L}(L^p_0(\mathbb{T}^d; X), L^p_0(\mathbb{T}^d; Y))},
$$

and we observed that $\sum_{k \in \mathbb{Z}^d} \widehat{f}(\varepsilon k) e_k \in L^p_0(\mathbb{T}^d; X)$, since $\widehat{f}(0) = 0$. $\square$

A combination of the transference results in both directions yields particularly nice corollaries for homogeneous multipliers of order zero, i.e., ones that satisfy

$$
m(t\xi) = m(\xi) \qquad \forall t > 0, \ \forall \xi \in \mathbb{R}^d \setminus \{0\}.
$$

We record some of them:

Corollary 5.7.6 (Homogeneous multipliers of order zero). *Let $m \in C(\mathbb{R}^d \setminus \{0\}; \mathscr{L}(X, Y))$ be a homogeneous function of order zero, and set*

$$
m(0) := \fint_{S(0,1)} m(\xi) \, \mathrm{d}\sigma(\xi),
$$

where σ is the normalised Haar measure on the unit sphere $S(0,1)$ of $\mathbb{R}^d$ and $\fint$ denotes the average. Then the following conditions are equivalent:

(1) *m is a Fourier multiplier from $L^p(\mathbb{R}^d; X)$ to $L^p(\mathbb{R}^d; Y)$;*
(2) *$(m(k))_{k \in \mathbb{Z}^d \setminus \{0\}}$ is a Fourier multiplier from $L^p_0(\mathbb{T}^d; X)$ to $L^p_0(\mathbb{T}^d; Y)$;*
(3) *$(m(k))_{k \in \mathbb{Z}^d}$ is a Fourier multiplier from $L^p(\mathbb{T}^d; X)$ to $L^p(\mathbb{T}^d; Y)$.*

If these equivalent conditions hold, then

$$
\begin{aligned}
\|T_m\|_{\mathscr{L}(L^p(\mathbb{R}^d; X), L^p(\mathbb{R}^d; Y))} &= \|\widetilde{T}_{(m(k))_{k \in \mathbb{Z}^d \setminus \{0\}}}\|_{\mathscr{L}(L^p_0(\mathbb{T}^d; X), L^p_0(\mathbb{T}^d; Y))} \\
&= \|\widetilde{T}_{(m(k))_{k \in \mathbb{Z}^d}}\|_{\mathscr{L}(L^p(\mathbb{T}^d; X), L^p(\mathbb{T}^d; Y))}.
\end{aligned}
$$

Proof. Denote the three norms by I, II, and III. By homogeneity,

$$m(0) = \fint_{S(0,r)} m(\xi)\,\mathrm{d}\sigma(\xi) = \fint_{B(0,r)} m(\xi)\,\mathrm{d}\xi$$

for any $r > 0$, and Proposition 5.7.4 applies to show that $I \leqslant II$. Obviously $II \leqslant III$, and $III \leqslant I$ by Proposition 5.7.1. □

Corollary 5.7.7. *Let X be a Banach space and let $p \in (1, \infty)$. Then the Hilbert transform H and its periodic version $\widetilde{H}$ satisfy*

$$\|H\|_{\mathscr{L}(L^p(\mathbb{R};X))} = \|\widetilde{H}\|_{\mathscr{L}(L^p_0(\mathbb{T};X))} = \|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};X))}.$$

Proof. The multiplier $m(\xi) = -i\,\mathrm{sgn}(\xi)$ is clearly in the scope of Corollary 5.7.6. □

Corollary 5.7.8. *The Schatten spaces $\mathscr{C}^p$, $1 < p < \infty$, satisfy*

$$\|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};\mathscr{C}^p))} \leqslant \cot \frac{\pi}{2 \cdot 2^{\lceil \log_2 p^* \rceil}} \leqslant \frac{4}{\pi} p^*, \qquad p^* = \max(p, p').$$

Proof. We has shown in Proposition 5.4.2 that

$$\|\widetilde{H}\|_{\mathscr{L}(L_0^{2^n}(\mathbb{T};\mathscr{C}^{2^n}))} \leqslant \cot \frac{\pi}{2^{n+1}}, \qquad \forall n = 1, 2, 3, \ldots.$$

By Corollary 5.7.7, the same is true with L^{2^n} in place of $L_0^{2^n}$, and the rest is essentially a repetition of the proof of Proposition 5.4.2:

For $p \in (2^{n-1}, 2^n)$ we have

$$L^p(\mathbb{T}; \mathscr{C}^p) = [L^{2^{n-1}}(\mathbb{T}; \mathscr{C}^{2^{n-1}}), L^{2^n}(\mathbb{T}; \mathscr{C}^{2^n})]_\sigma$$

for $\sigma \in (0, 1)$ determined by $1/p = (1 - \sigma)/2^{n-1} + \sigma/2^n$, and hence

$$\|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};\mathscr{C}^p))} \leqslant \|\widetilde{H}\|^{1-\sigma}_{\mathscr{L}(L^{2^{n-1}}(\mathbb{T};\mathscr{C}^{2^{n-1}}))} \|\widetilde{H}\|^{\sigma}_{\mathscr{L}(L^{2^n}(\mathbb{T};\mathscr{C}^{2^n}))} \leqslant \cot \frac{\pi}{2^{n+1}}.$$

For $p \in (1, 2)$, we have

$$\|\widetilde{H}\|_{\mathscr{L}(L^{p'}(\mathbb{T};\mathscr{C}^{p'}))} = \|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};\mathscr{C}^p))},$$

since $(L^p(\mathbb{T}; \mathscr{C}^p))^* = L^{p'}(\mathbb{T}; \mathscr{C}^{p'})$ and $H^* = -H$. □

5.7.c Periodic multiplier theorems

We now make systematic use of the transference machine to derive periodic analogues of several of the Euclidean multiplier theorems treated in the earlier sections. Their proofs feature a couple of useful techniques that the reader may find useful in prospective applications of the transference technique to other situations. We begin with a simple version of Mihlin's multiplier, whose proof allows a rather 'clean' application of transference, a luxury that is not always available.

Theorem 5.7.9 (Mihlin's multiplier theorem on $\mathbb{T}$). *Let X and Y be a UMD space and let $p \in (1,\infty)$. Then each sequence $m = (m(k))_{k\in\mathbb{Z}} \in \ell^\infty(\mathbb{Z};\mathscr{L}(X,Y))$ such that*

$$\|m\|_{\mathfrak{M}_p(\mathbb{Z};X,Y)} := \mathscr{R}_p(\{m(k)\}_{k\in\mathbb{Z}}) + \mathscr{R}_p(\{k[m(k)-m(k-\operatorname{sgn}(k))]\}_{k\in\mathbb{Z}\setminus\{0\}}) \tag{5.59}$$

is finite, is a Fourier multiplier from $L^p(\mathbb{T};X)$ to $L^p(\mathbb{T};Y)$, and

$$\|\widetilde{T}_m\|_{\mathscr{L}(L^p(\mathbb{T};X),L^p(\mathbb{T};Y))} \leqslant 800\min\{\hbar_{p,X},\hbar_{p,Y}\}\beta_{p,X}\beta_{p,Y}\|m\|_{\mathfrak{M}(\mathbb{Z};X,Y)}.$$

Proof of Theorem 5.7.9. We extend m to the piecewise affine function on $\mathbb{R}$ that interpolates between the given values at the lattice points $k \in \mathbb{Z}$. Thus

$$m(k-\sigma u) = (1-u)m(k) + um(k-\sigma) \text{ for } k \in \mathbb{Z}\setminus\{0\},\ \sigma = \operatorname{sgn}(k),\ u \in (0,1),$$

and it follows that

$$\begin{aligned}\mathscr{R}_p(m(\xi):\xi\in\mathbb{R}) \leqslant \mathscr{R}_p&((1-u)m(k):k\in\mathbb{Z},u\in[0,1]) \\ &+ \mathscr{R}_p(um(k):k\in\mathbb{Z},u\in[0,1]) \leqslant 2\mathscr{R}_p(m(k):k\in\mathbb{Z}).\end{aligned}$$

Concerning the derivative, we have

$$(k-\sigma u)m'(k-\sigma u) = -(1-\frac{\sigma u}{k})k(m(k)-m(k-\sigma)),$$

which belongs to the product of the R-bounded sets $[-2,2]\cdot I_Y$ (of R-bound 2) and $\{k(m(k)-m(k-\operatorname{sgn}(k)))\}_{k\in\mathbb{Z}\setminus\{0\}}$.

Thus $\|m\|_{\mathfrak{M}(\mathbb{R};X,Y)} \leqslant 2\|m\|_{\mathfrak{M}(\mathbb{Z};X,Y)}$, and the theorem follows from

$$\begin{aligned}\|\widetilde{T}_m\|_{\mathscr{L}(L^p(\mathbb{T};X),L^p(\mathbb{T};Y))} &\leqslant \|T_m\|_{\mathscr{L}(L^p(\mathbb{R};X),L^p(\mathbb{T};Y))} \\ &\leqslant 400\min(\hbar_{p,X},\hbar_{p,Y})\beta_{p,X}\beta_{p,Y}\|m\|_{\mathfrak{M}(\mathbb{R};X,Y)},\end{aligned}$$

i.e., a combination of Proposition 5.7.1 (the transference from $\mathbb{R}$ to $\mathbb{T}$) and Mihlin's Theorem 5.3.18. □

We proceed with a periodic version of the Littlewood–Paley inequalities. Here, a slight obstruction occurs, in that the dyadic lattice points fail to be Lebesgue points (which was assumed for the transference in Proposition 5.7.1) for the indicators $\mathbf{1}_{[2^k,2^{k+1})}$ inherent in this multiplier theorem. As we shall see, this issue is easily circumvented by a simple dilation argument, which is also useful in other situations dealing with transference of discontinuous multipliers.

To set the stage for the theorem, we define the projections

$$\widetilde{\Delta}_k f := \sum_{j\in\widetilde{I}_k} \widehat{f}(j)e_j, \qquad \widetilde{I}_k := \begin{cases}[2^{k-1},2^k)\cap\mathbb{Z}, & \text{if } k>0,\\ \{0\}, & \text{if } k=0,\\ (-2^{|k|},-2^{|k|-1}]\cap\mathbb{Z}, & \text{if } k<0.\end{cases}$$

Theorem 5.7.10 (Littlewood–Paley inequalities on $\mathbb{T}$). *Let X and Y be a UMD space, $p \in (1,\infty)$, and let $f : \mathbb{T} \to X$ be a trigonometric polynomial.*

(1) *For all R-bounded sequences $(\lambda_k)_{k\in\mathbb{Z}} \subseteq \mathscr{L}(X,Y)$ we have*

$$\Big\| \sum_{k\in\mathbb{Z}} \lambda_k \widetilde{\Delta}_k f \Big\|_{L^p(\mathbb{T};X)} \leqslant 400 \min(\hbar_{p,X}, \hbar_{p,Y}) \beta_{p,X} \beta_{p,Y} \mathscr{R}_p((\lambda_k)_{k\in\mathbb{Z}}) \|f\|_{L^p(\mathbb{T};X)}.$$

(2) *If $(\varepsilon_k)_{k\in\mathbb{Z}}$ is a Rademacher sequence, then*

$$\frac{1}{40\hbar_{p,X}\beta_{p,X}} \|f\|_{L^p(\mathbb{T};X)} \leqslant \Big\| \sum_{k\in\mathbb{Z}} \varepsilon_k \widetilde{\Delta}_k f \Big\|_{L^p(\Omega\times\mathbb{T};X)} \leqslant 40\hbar_{p,X}\beta_{p,X} \|f\|_{L^p(\mathbb{T};X)}.$$

Note that the first statement is the boundedness of $\widetilde{T}_m$, where $m(j) = \lambda_k$ for $j \in \widetilde{I}_k$. We have included this case in the above theorem, since it is not covered by the simplified Mihlin condition (5.59) of Theorem 5.7.9, which was most amenable to a clean transference argument. (The problem is that (5.59) demands that the jumps $|m(j) - m(j - \operatorname{sgn} j)|$ decay also along the dyadic powers $j = \pm 2^k$, while the piecewise constant multipliers may have big jumps all the way to infinity, as long as they remain uniformly bounded.)

Proof. The first estimate is the statement of the boundedness of the periodic multiplier $\widetilde{T}_m$ with $m(j) = \lambda_k$ for $j \in \widetilde{I}_k$. We may extend m to all $\mathbb{R}$ by setting

$$m(\xi) := \lambda_k \text{ for } \xi \in I_k := \begin{cases} [2^{k-1}, 2^k), & \text{if } k > 0, \\ (-1,1), & \text{if } k = 0, \\ (-2^{|k|}, -2^{|k|-1}], & \text{if } k < 0. \end{cases} \tag{5.60}$$

This function is constant over all intervals $I \in \mathscr{I}$, so it is in the Mihlin class with $\|m\|_{\mathfrak{M}(\mathbb{R};X,Y)} = \mathscr{R}_p((\lambda_k)_{k\in\mathbb{Z}})$, and in particular it induces a bounded multiplier operator from $L^p(\mathbb{R};X)$ into $L^p(\mathbb{R};Y)$. Although this agrees with $m(j)$ at the integer points, an initial problem in applying the transference is that m may have different one-sided limits at the points $\pm 2^k$, which prevents them from being Lebesgue points. However, this issue is easily resolved with a simple dilation trick:

Pick some irrational $\rho > 1$. Then $m(\rho\xi)$ is continuous at all integer points $\xi = j$, and hence by Theorem 5.3.18

$$\begin{aligned} \|m(\rho\cdot)\|_{\mathfrak{M}L^p(\mathbb{T};X,Y)} &\leqslant \|m(\rho\cdot)\|_{\mathfrak{M}L^p(\mathbb{R};X,Y))} = \|m\|_{\mathfrak{M}L^p(\mathbb{R};X,Y))} \\ &\leqslant 400 \min(\hbar_{p,X}, \hbar_{p,Y}) \beta_{p,X}\beta_{p,Y} \|m\|_{\mathfrak{M}_p(\mathbb{R};X,Y)}, \end{aligned} \tag{5.61}$$

where $\|m\|_{\mathfrak{M}_p(\mathbb{R};X,Y)} = \mathscr{R}_p((\lambda_k)_{k\in\mathbb{Z}})$.

Let us now consider an arbitrary trigonometric polynomial $f \in L^p(\mathbb{T};X)$. We would like to pick a ρ such that $m(\rho j) = m(j)$ for all $j \in \operatorname{supp} \widehat{f}$. This is certainly true for $j = 0$. Since m is constant on $\pm[|j|, |j|+1)$, for other values of j we need that $|j| \leqslant \rho|j| < |j| + 1$, where the first half is automatic from

$\rho > 1$, and the second one requires that $\rho < 1 + |j|^{-1}$. Since $\operatorname{supp} \widehat{f}$ is finite, we can find a $\rho = \rho_f > 1$ that satisfies this condition for all $j \in \operatorname{supp} \widehat{f}$. But then

$$\begin{aligned} \|\widetilde{T}_m f\|_{L^p(\mathbb{T};Y)} &= \|\widetilde{T}_{m(\rho_f \cdot)} f\|_{L^p(\mathbb{T};Y)} \\ &\leqslant \|m(\rho_f \cdot)\|_{\mathfrak{M}L^p(\mathbb{T};X,Y))} \|f\|_{L^p(\mathbb{T};X)} \\ &\leqslant \|m\|_{\mathfrak{M}L^p(\mathbb{R};X,Y)} \|f\|_{L^p(\mathbb{T};X)}. \end{aligned} \tag{5.62}$$

The final bound is uniform in the choice of trigonometric polynomial f, and shows that $\|m\|_{\mathfrak{M}L^p(\mathbb{T};X,Y)} \leqslant \|m\|_{\mathfrak{M}L^p(\mathbb{R};X,Y)}$. In combination with (5.61), this completes the proof of the first assertion of the Theorem.

We turn to the right-hand bound of the second assertion. Formally, this has the same form as the case already treated, except that the constant λ_k is now replaced by a random variable ε_k. We interpret this as an operator $\lambda_k : x \in X \mapsto \varepsilon_k \otimes x \in L^p(\Omega; X) =: Y$, and define the m exactly as in (5.60). This special multiplier satisfies the estimate

$$\|m\|_{\mathfrak{M}L^p(\mathbb{R};X,Y)} \leqslant 40\hbar_{p,X}\beta_{p,X},$$

which is simply a restatement of the Littlewood–Paley inequality on $\mathbb{R}$ as given in Corollary 5.3.25. The rest of the proof proceeds in complete analogy with the previous case, using an auxiliary dilation and the computation (5.62) to show that

$$\|\widetilde{T}_m f\|_{L^p(\mathbb{T};L^p(\Omega;X))} \leqslant \|m\|_{\mathfrak{M}L^p(\mathbb{R};X,Y)} \|f\|_{L^p(\mathbb{T};X)}$$

for all trigonometric polynomials $f \in L^p(\mathbb{T}; X)$, and hence all $f \in L^p(\mathbb{T}; X)$ by density. Finally, the left-hand bound of the second assertion is achieved by a duality argument completely analogous to that in the proof of Theorem 5.3.24. This completes the proof. □

Periodic multiplier theorems in several variables

As in the continuous case, there is a multi-dimensional version of Mihlin's Theorem 5.7.9, which is similar but technically more complicated. We provide a detailed statement and proof, which is also instructive in illustrating the use of 'partial difference operators' as a substitute for partial derivatives in discrete multivariate analysis. To this end, let us define

$$\tau_i m(k) := m(k - \operatorname{sgn}(k_i)e_i), \qquad \Delta_i m(k) := m(k) - \tau_i m(k),$$

as well as

$$\tau^\alpha m(k) := \Big(\prod_{i:\alpha_i=1} \tau_i\Big) m(k), \qquad \Delta^\alpha m(k) := \Big(\prod_{i:\alpha_i=1} \Delta_i\Big) m(k).$$

(We take the convention that $\operatorname{sgn}(0) = 0$, so that $\tau_i m(k) = m(k)$ and $\Delta_i m(k) = 0$ if $k_i = 0$.) We can now formulate the multiplier theorem as:

Theorem 5.7.11 (Mihlin's multiplier theorem on $\mathbb{T}^d$). *Let X and Y be UMD spaces and let $p \in (1, \infty)$. Then each sequence $m = (m(k))_{k\in\mathbb{Z}^d} \in \ell^\infty(\mathbb{Z}^d; \mathscr{L}(X, Y))$ such that*

$$\|m\|_{\mathfrak{M}_p(\mathbb{Z}^d;X,Y)} := \sum_{\alpha\in\{0,1\}^d} \mathscr{R}_p(\{|k|^{|\alpha|}\Delta^\alpha m(k) : k \in \mathbb{Z}^d\}) < \infty$$

is a Fourier multiplier of $L^p(\mathbb{T}^d; X)$, and

$$\|m\|_{\mathfrak{M}L^p(\mathbb{T}^d;X,Y)} \leqslant 100 \cdot 16^d \cdot \min(\hbar_{p,X}, \hbar_{p,Y})^d \cdot \beta_{p,X}\beta_{p,Y} \cdot \|m\|_{\mathfrak{M}(\mathbb{Z}^d;X,Y)}.$$

Proof. We proceed in a similar way as in the one-dimensional theorem, extending

$$m(\xi) := \sum_{k\in\mathbb{Z}^d} m(k)\phi_d(\xi - k),$$

where

$$\phi_d(\xi) := \prod_{i=1}^d \phi(\xi_i), \qquad \phi(\xi_i) = (1 - |\xi_i|)_+,$$

is a tensor product of the auxiliary functions from the one-dimensional proof. It is immediate that this satisfies the qualitative continuity and differentiability criteria for the membership in the Mihlin class $\mathfrak{M}(\mathbb{R}^d; X, Y)$, as defined in Definition 5.5.9.

To estimate the Mihlin norm of $m(\xi)$, it suffices to consider $\xi \in (\mathbb{R}\setminus\mathbb{Z})^d$. In order not to overburden the notation with repeated occurrences of the signum function, we restrict ourselves by symmetry to the case that $\xi \in (\mathbb{R}_+)^d$, and write it as $\xi = k - u$, where $k \in \mathbb{Z}^d_+$ and $u \in [0, 1)^d$. Then

$$\partial^\alpha m(\xi) = \sum_{j\in k-\{0,1\}^d} m(j)\partial^\alpha\phi_d(\xi - j) = \sum_{\beta\in\{0,1\}^d} m(k - \beta)\partial^\alpha\phi_d(\beta - u).$$

Since $\tau_i = I - \Delta_i$, it follows that

$$\tau^\beta = \prod_{i:\beta_i=1} (I - \Delta_i) = \sum_{\delta\leqslant\beta} (-1)^{|\delta|}\Delta^\delta,$$

and then

$$\begin{aligned}\sum_{\beta\in\{0,1\}^d} m(k - \beta)\partial^\alpha\phi_d(\beta - u) &= \sum_{\beta\leqslant\mathbf{1}} \tau^\beta m(k)\partial^\alpha\phi_d(\beta - u)\\ &= \sum_{\delta\leqslant\mathbf{1}} (-1)^{|\delta|}\Delta^\delta m(k) \sum_{\beta:\delta\leqslant\beta\leqslant\mathbf{1}} \partial^\alpha\phi_d(\beta - u).\end{aligned} \tag{5.63}$$

In the innermost sum, recall that $\partial^\alpha\phi_d(\beta - u)$ is a product of factors

$$\phi^{(\alpha_i)}(\beta_i - u_i) = \begin{cases} \phi'(\beta_i - u_i), & \text{if } \alpha_i = 1,\\ \phi(\beta_i - u_i), & \text{if } \alpha_i = 0.\end{cases}$$

If $\delta \not\geqslant \alpha$, then there is at least one i with $\delta_i = 0$ and $\alpha_i = 1$. For this i, the summation variable β_i goes through both values 0 and 1, so that the factor

$$\phi'(-u_i) + \phi'(1-u_i) = 0$$

appears, making the entire inner sum vanish. Thus only the summands with $\delta \geqslant \alpha$ survive in (5.63), and in this case the inner sum is estimated by

$$\begin{aligned}\Big|\sum_{\beta\geqslant\delta} \partial^\alpha \phi_d(\beta - u)\Big| &= \Big| \prod_{i:\delta_i=1} \phi^{(\alpha_i)}(1-u_i) \times \prod_{i:\delta_i=0} \Big(\sum_{\beta_i=0}^{1} \phi(-u) + \phi(1-u)\Big)\Big| \\ &= \prod_{i:\delta_i=1} |\phi^{(\alpha_i)}(1-u_i)| \times \prod_{i:\delta_i=0} 1 \leqslant 1,\end{aligned}$$

since $\|\phi'\|_\infty = \|\phi\|_\infty = 1$. (When $d > 1$, it seems difficult to exploit the more precise information that $\phi(1-u_i) = u_i$, although omitting it is partially responsible for the loss of the numerical factor 2^d in the final bound.)

Substituting back, we have

$$\partial^\alpha m(\xi) = \sum_{\delta\geqslant\alpha} (-1)^{|\delta|} \Delta^\delta m(k) \sum_{\beta\geqslant\delta} \partial^\alpha \phi_d(\beta-u),$$

and hence, using the just established (R-)boundedness of the scalar factors $\sum_{\beta\geqslant\delta} \partial^\alpha\phi_d(\beta-u)$, as well as $|\xi|^{|\alpha|} \leqslant |k|^{|\alpha|} \leqslant |k|^{|\delta|}$ since $|k| \geqslant \max\{|\xi|, 1\}$, we have

$$\mathscr{R}_p(|\xi|^{|\alpha|}\partial^\alpha m(\xi) : \xi \in (\mathbb{R}\setminus\mathbb{Z})^d) \leqslant \sum_{\delta\geqslant\alpha} \mathscr{R}_p(|k|^{|\delta|}\Delta^\delta m(k) : k \in \mathbb{Z}^d).$$

Summing over $\alpha \in \{0,1\}^d$, this gives

$$\|m\|_{\mathfrak{M}L^p(\mathbb{R}^d;X,Y)} \leqslant 2^d \|m\|_{\mathfrak{M}L^p(\mathbb{Z}^d;X,Y)}.$$

As in the case $d = 1$, the theorem then follows from

$$\begin{aligned}\|m\|_{\mathfrak{M}L^p(\mathbb{T}^d;X,Y)} &\leqslant \|m\|_{\mathfrak{M}L^p(\mathbb{R}^d;X,Y)} \\ &\leqslant 100 \cdot 8^d \min(\hbar_{p,X}, \hbar_{p,Y})^d \beta_{p,X}\beta_{p,Y} \|m\|_{\mathfrak{M}(\mathbb{R}^d;X,Y)},\end{aligned}$$

i.e., a combination of Proposition 5.7.1 (the transference from $\mathbb{R}^d$ to $\mathbb{T}^d$) and Mihlin's Theorem 5.5.10. □

We conclude with a version of the periodic Littlewood–Paley inequality in d variables:

Theorem 5.7.12 (Littlewood–Paley inequality on $\mathbb{T}^d$). *Let X be a UMD space and let $p \in (1,\infty)$. With the constant*

$$C := C_{d,p,X} = 10 \cdot 8^d \cdot \hbar_{p,X}^d \beta_{p,X},$$

we have the following estimate for all $f \in L^p(\mathbb{T}^d; X)$:

$$\frac{1}{C}\|f\|_{L^p(\mathbb{T};X)} \leqslant \Big\| \sum_{k=0}^{\infty} \varepsilon_k \widetilde{\Delta}_{[k]} f \Big\|_{L^p(\Omega\times\mathbb{T};X)} \leqslant C\|f\|_{L^p(\mathbb{T};X)},$$

where

$$\widetilde{\Delta}_{[k]} f := \sum_{j\in\widetilde{S}_k} \widehat{f}(j) e_j,$$

and

$$\widetilde{S}_k := \begin{cases} \{0\}, & \text{if } k = 0, \\ [(-2^k, 2^k)^d \setminus (-2^{k-1}, 2^{k-1})^d] \cap \mathbb{Z}^d, & \text{if } k > 0. \end{cases}$$

Proof. The deduction of this theorem by transference from Corollary 5.5.19 is completely analogous to the deduction of the one-variable version in Theorem 5.7.10 from Corollary 5.3.25. □

5.8 Notes

Section 5.1

The scalar-valued L^p-boundedness, $1 < p < \infty$, of the Hilbert transform is a classical result first announced in Riesz [1924], with a detailed proof presented in Riesz [1928]. The accompanying weak type estimate is due to Kolmogorov [1925].

The equivalence of the UMD property with the Hilbert transform boundedness in the vector-valued context, Theorem 5.1.1, was first established by Burkholder [1983] and Bourgain [1983]. Burkholder [1983] originally proved the sufficiency of the UMD condition for the Hilbert transform boundedness in the periodic setting of $L^p(\mathbb{T}; X)$ with the help of Brownian motion in the unit disk. The basic philosophy of this approach is the same as ours, but the advantage of using a continuous stochastic model is that it avoids the numerical factor produced by our discrete scheme, yielding the best available bound

$$\hbar_{p,X} \leqslant \beta_{p,X}^2.$$

A proof in spirit to the one presented above, where the Hilbert transform is recovered from elementary transformations of the Haar basis, was first devised by Figiel [1990], based on his earlier analysis (Figiel [1988]) of the mentioned elementary transformations. The approach of Figiel [1990] is much more general than the present one, and yields the $L^p(\mathbb{R}^d; X)$-boundedness of all Calderón–Zygmund singular integrals

$$Tf(x) = \int_{\mathbb{R}^d} K(x, y) f(y)\,\mathrm{d}y, \qquad x \notin \operatorname{supp} f, \tag{5.64}$$

assuming (say) the scalar-valued $L^2(\mathbb{R}^d)$-boundedness of the operator T, and the "standard estimates"

$$\begin{aligned} |K(x,y)| &\leqslant \frac{C}{|x-y|^d}, \\ |K(x,y)-K(x',y)|+|K(y,x)-K(y,x')| &\leqslant C\frac{|x-x'|^\alpha}{|x-y|^{d+\alpha}} \end{aligned} \tag{5.65}$$

for some $\alpha \in (0,1]$, whenever $|x-x'| < \frac{1}{2}|x-y|$. The core of this approach is the following representation theorem, which we formulate only in dimension $d = 1$ and in a special case for the sake of simplicity:

Theorem 5.8.1 (Figiel [1990]). *Suppose that $T \in \mathscr{L}(L^2(\mathbb{R}))$ has an integral representation with the standard estimates as above, and moreover that T and its adjoint T^* map constant functions to zero (in a suitable sense). Then*

$$T = \sum_{m\in\mathbb{Z}} (T_m \Lambda_m^0 + U_m \Lambda_m^1 + \Lambda_m^2 U_m^*), \tag{5.66}$$

where we have Haar multipliers

$$\Lambda_m^i : h_I \mapsto \lambda_{m,I}^i h_I, \qquad |\lambda_{m,I}^i| \leqslant C(1+|m|)^{-2},$$

and two kinds of Haar transformations

$$T_m : h_I \mapsto h_{I+m\ell(I)}, \qquad U_m : h_I \mapsto |I|^{-1/2}(\mathbf{1}_{I+m\ell(I)} - \mathbf{1}_I).$$

Without the annihilation of constants assumption, two additional terms—so called *dyadic paraproducts*—would need to be added to the representation formula (5.66). The unconditionality of the Haar system immediately implies that

$$\|\Lambda_m^i\|_{\mathscr{L}(L^p(\mathbb{R};X))} \leqslant C(1+|m|)^{-2}\beta_{p,X},$$

so that the summability of (5.66) in $\mathscr{L}(L^p(\mathbb{R};X))$ follows from the delicate combinatorial bounds

$$\|T_m\|_{\mathscr{L}(L^p(\mathbb{R};X))} + \|U_m\|_{\mathscr{L}(L^p(\mathbb{R};X))} \leqslant C_{p,X}(1+\log_+|m|)$$

established by Figiel [1988] for any UMD space X and $p \in (1,\infty)$.

While the Hilbert transform is certainly contained in this class, the argument of Figiel [1990] does not gain any particular simplification from the specialisation to this particular case, and requires the consideration of the countable family of Haar transformations T_m and U_m even in this case.

Instead, we have given a proof of Burkholder's theorem essentially due to Petermichl and Pott [2003], which is based on a special representation of the Hilbert transform discovered by Petermichl [2000]: at the cost of using translated and dilated dyadic systems rather than a fixed one, it achieves a representation of the Hilbert transform with a single Haar transformation, the dyadic shift

$$S : h_I \mapsto \frac{1}{\sqrt{2}}(h_{I_{\text{left}}} - h_{I_{\text{right}}}).$$

On the level of details, we have used a slight modification of the original representation of Petermichl [2000]; it is taken from Hytönen [2008] and based on the idea of random dyadic systems from Nazarov, Treil, and Volberg [2003].

Our derivation of the preliminary version of the Littlewood–Paley inequality, as recorded in Proposition 5.1.10, follows the original approach of Bourgain [1984a, 1986b]; embedding this into the framework of dyadic shifts is a superficial detail, but this unified deduction of the two central analytic consequences of the UMD inequality is new compared to the existing literature.

The invariance properties mentioned in Remark 5.1.6 almost characterise the Hilbert transform: the only $L^2(\mathbb{R})$-bounded, translation and dilation invariant linear operators are linear combinations of the identity and the Hilbert transform. See Stein [1970a, Proposition III.1] for the (relatively straightforward) proof.

For many concrete Banach spaces X, the boundedness of the Hilbert transform can be checked directly, i.e., without passing through the UMD property. See for instance Clément and De Pagter [1991] for the case of $X = L^p(S; L^q(T))$, $1 < p, q < \infty$.

General singular integral operators and $T(1)$ theorems

The result of Figiel [1990] mentioned above was in fact an extension of the $T(1)$ theorem of David and Journé [1984] to the vector-valued setting. Namely, we have

Theorem 5.8.2 (David and Journé [1984], Figiel [1990]). *Let T be an integral operator as in (5.64) and (5.65), initially defined on a suitable test function class. Then the following conditions are equivalent:*

(1) *T extends to $T \in \mathscr{L}(L^2(\mathbb{R}^d))$.*
(2) *$|\langle T(\mathbf{1}_Q), \mathbf{1}_Q\rangle| \leqslant C|Q|$ for all cubes Q, and $T(1), T^*(1) \in \mathrm{BMO}(\mathbb{R}^d)$.*
(3) *T extends to $T \in \mathscr{L}(L^p(\mathbb{R}^d; X))$ for all UMD spaces X and $p \in (1, \infty)$.*

Here $T(1)$ refers to the action of T on the constant function 1, which may be made precise in a variety of ways, as discussed in the cited papers. The equivalence (1)⇔(2) is the celebrated $T(1)$ theorem of David and Journé [1984], the first general characterisation of the boundedness of singular integral operators beyond convolution kernels $K(x, y) = k(x - y)$, while (2)⇔(3) is the $T(1)$ theorem of Figiel [1990]; together they give the equivalence (1)⇔(3).

This first vector-valued $T(1)$ theorem has been the inspiration of several variants and extensions, including the following:

- A Fourier-analytic proof and an extension to operator-valued kernels $K(x, y) \in \mathscr{L}(X, Y)$ by Hytönen and Weis [2006].

- A UMD-extension of the "$T(b)$ theorem" of David, Journé, and Semmes [1985] by Hytönen [2006], replacing the constant function 1 by more general test functions in condition (2) above.
- An extension of Figiel's decomposition and the $T(1)$ theorem to abstract metric space domains (in place of $\mathbb{R}^d$) with a doubling measure by Müller and Passenbrunner [2012]. The underlying combinatorial arguments have been further analysed by Lechner and Passenbrunner [2014].
- A UMD-extension of the non-homogeneous $T(b)$ theorem of Nazarov, Treil, and Volberg [2003] (for a non-doubling measure on $\mathbb{R}^d$) by Hytönen [2014]. A version on abstract metric space domains was obtained by Martikainen [2012], and a "local" Tb theorem à la Nazarov, Treil, and Volberg [2002] (using a family of test functions b_Q indexed by all cubes) by Hytönen and Vähäkangas [2015]. (Contrary to the other listed results, the last mentioned paper assumes somewhat stronger conditions than mere UMD.)
- New proofs for $\mathbb{R}^d$ and the Lebesgue measure, but paying attention to the quantitative dependence on the UMD constant $\beta_{p,X}$, have been given by Pott and Stoica [2014] and Hänninen and Hytönen [2016].

Applications of the Hilbert transform in spectral theory

Despite the large arsenal of singular integral bounds available in UMD spaces, discussed in the main text and the Notes above, a number of deep consequences can already be deduced as more or less direct applications of the Hilbert transform alone. A highly influential early example is the theorem of Dore and Venni [1987] on sums of closed operators:

Theorem 5.8.3 (Dore and Venni [1987]). *Let X be a complex UMD space, and for $i = 0, 1$, let $A_i : \mathsf{D}(A_i) \to X$ be closed linear operators with dense domains $\mathsf{D}(A_i) \subseteq X$. Suppose also that, for both $j = 0, 1$, the following conditions are satisfied:*

(i) *the resolvent $\rho(A_j) := \{\lambda \in \mathbb{C} : (\lambda - A_j)^{-1} \in \mathscr{L}(X)\}$ contains $(-\infty, 0]$ and*
$$\|(t + A_j)^{-1}\|_{\mathscr{L}(X)} \leqslant \frac{M}{1+t} \qquad \forall t \in [0, \infty);$$

(ii) *the resolvents of A_0 and A_1 commute:*
$$(\lambda_0 - A_0)^{-1}(\lambda_1 - A_1)^{-1} = (\lambda_1 - A_1)^{-1}(\lambda_0 - A_0)^{-1} \qquad \forall \lambda_j \in \rho(A_j);$$

(iii) *the imaginary powers $s \in \mathbb{R} \mapsto A_j^{is}$, which can be defined in view of (i), are strongly continuous in $\mathscr{L}(X)$ with bounds*
$$\|A_j^{is}\|_{\mathscr{L}(X)} \leqslant K e^{\theta_j |s|},$$
where $\theta_0 \geqslant 0$ and $\theta_1 \geqslant 0$ satisfy
$$\theta_0 + \theta_1 < \pi.$$

Then the operator $A_0 + A_1$, defined on the domain $\mathsf{D}(A_0 + A_1) = \mathsf{D}(A_0) \cap \mathsf{D}(A_1)$, is a closed operator. Moreover, $A_0 + A_1$ has a bounded inverse $(A_0 + A_1)^{-1} \in \mathscr{L}(X)$.

The proof is based on an analysis of the operator-valued integral

$$S := \frac{1}{2i} \int_{c-i\infty}^{c+i\infty} \frac{A_0^{-z} A_1^{z-1}}{\sin(\pi z)} \, \mathrm{d}z, \qquad c \in (0,1),$$

which is eventually shown to provide a representation of $(A_0 + A_1)^{-1}$. After various reductions, the critical step is proving the existence of the limit

$$\lim_{\varepsilon \to 0} \int_{\varepsilon \leqslant |s| \leqslant 1} \frac{A_0^{-is} A_1^{is} x}{\pi s} \, \mathrm{d}s = Hf(0), \tag{5.67}$$

where

$$f(s) := \mathbf{1}_{(-1,1)}(s) A_0^{-is} A_1^{is} x$$

with $x \in X$, so that $f \in L^2(\mathbb{R}; X)$ (say) by (iii). For a UMD space X, the principal value of the Hilbert transform is guaranteed to exist at *almost every* point (by Theorem 5.1.1). Thus, there is a small $t > 0$ at which $Hf(t)$ and hence $A_0^{-it} A_1^{it} Hf(t)$ exists, where

$$A_0^{-it} A_1^{it} Hf(t) = \lim_{\varepsilon \to 0} \int_{\substack{|s| \geqslant \varepsilon \\ |s-t| \leqslant 1}} \frac{A_0^{-is} A_1^{is} x}{\pi s} \, \mathrm{d}s$$

differs from (5.67) by the non-singular integrals over $(-1, t-1) \cup (1, 1+t)$.

An extension of Theorem 5.8.3 to non-commuting operators, relaxing the assumption (ii), is due to Monniaux and Prüss [1997] and is based on similar reasoning. On the other hand, a non-symmetric version of Theorem 5.8.3 by Kalton and Weis [2001], allowing more general operators A_1 than (iii) when more is known about A_0, depends on different technology, and we return to this in a subsequent Volume. The applications of Theorem 5.8.3, and its extensions, are too numerous to recount here, and we refer the reader to the many papers citing these works. We refer to Prüss and Simonett [2016] for an excellent exposition and an extensive collection of references.

Applications of dyadic shifts to weighted norm inequalities

Petermichl [2000] originally devised her dyadic shift method for the estimation of the norms of commutators of the Hilbert transform and matrix-valued BMO functions, and its most celebrated application was in her proof of the sharp weighted norm inequality (Petermichl [2007])

$$\|Hf\|_{L^2(w)} \leqslant c[w]_{A_2} \|f\|_{L^2(w)}, \tag{5.68}$$

where

$$[w]_{A_2} := \sup_{\substack{I \subseteq \mathbb{R} \\ \text{interval}}} \Big(\fint_I w(x)\, \mathrm{d}x \Big) \Big(\fint_I \frac{1}{w(x)}\, \mathrm{d}x \Big)$$

is Muckenhoupt's A_2 constant. The bound (5.68) is optimal in its dependence on $[w]_{A_2}$, in that $[w]_{A_2}$ cannot be replaced by $\phi([w]_{A_2})$ for any ϕ with $\phi(t)/t \to 0$ as $t \to \infty$. Sharp weighted inequalities of this type have been the topic of considerable attention, and (5.68) was generalised to arbitrary Calderón–Zygmund operators T in place of H by Hytönen [2012], a result known as the "A_2 theorem". This was again based on a representation of these operators as averages of dyadic shifts, but of a more general form than treated here.

By now, several proofs of the A_2 theorem are available. In particular, the simple proof by Lerner [2013] turned out amenable for a vector-valued extension, which was obtained by Hänninen and Hytönen [2014]. This result does not involve the UMD condition directly, in that taking the (unweighted) $L^p(\mathbb{R}^d; X)$-boundedness of a Calderón–Zygmund operator T as in (5.64) and (5.65) as an assumption, the argument is valid in an arbitrary Banach space X. However, to verify the aforementioned assumption for concrete operators, UMD is typically needed. A combination of Theorem 5.1.1 and the main result of Hänninen and Hytönen [2014] shows in particular that (5.68) remains true with $L^2(w; X)$ in place of $L^2(w)$, replacing the universal c by a number c_X depending on the UMD space X.

Section 5.2

The converse direction of Theorem 5.1.1, the necessity of UMD for the Hilbert transform boundedness, was proved more or less by the original argument of Bourgain [1983]. A variant of this method using an "abstract conjugate function" (an analogue of the Hilbert transform on a compact abelian group) is given by Asmar, Kelly, and Montgomery-Smith [1996]. Bourgain's argument has two important features:

(a) It can be adapted to some other (Fourier-analytic) operators in place of the Hilbert transform, and it is essentially the only known method for deriving the defining property of UMD from the boundedness of such operators.
(b) In many situations, it is quantitatively efficient, in that it provides sharp lower bounds for the operator norms in terms of the UMD constant.

Some of the existing variants of this method yield the necessity of UMD for the L^p-boundedness of:

(1) the Fourier multipliers with symbols $|\xi|^{is}$ (Guerre-Delabrière [1991]);
(2) a large class of Fourier multipliers with even, homogeneous symbols

$$m(\xi) = m(\lambda\xi) \qquad \forall \xi \in \mathbb{R}^d \setminus \{0\}, \quad \forall \lambda \in \mathbb{R} \setminus \{0\},$$

including the second order Riesz transform $R_j R_k$ (Geiss, Montgomery-Smith, and Saksman [2010]);

and for the *two-sided* L^p-boundedness (i.e., boundedness, and boundedness from below) of:

(3) a version of the Littlewood–Paley square function for the Poisson semigroup (Hytönen [2007]).

The result of Geiss, Montgomery-Smith, and Saksman [2010], in particular, has important implications on the size of operator norms. Among other consequences, it shows that

$$\|2R_jR_k\|_{\mathscr{L}(L^p(\mathbb{R}^d;X))} = \beta_{p,X} \quad \text{for} \quad j \neq k,$$

where "$\leqslant$" was known before, whereas "$\geqslant$" was new even for $X = \mathbb{C}$. This is interesting, since $-2R_1R_2$ is the imaginary part of the Beurling transform $B = (R_1 - iR_2)^2$, whereas $\beta_{p,\mathbb{C}} = p^* - 1$ is the conjectured norm of the full operator B on $\mathscr{L}(L^p(\mathbb{R}^2))$.

Spectral theory and transference

After its birth within the context of martingales and singular integrals, the study of UMD spaces soon expanded in the direction of spectral theory, as it was realised that the boundedness of the vector-valued Hilbert transform and related objects on $L^p(\mathbb{R};X)$ could be *transferred*, via the classical method of Coifman and Weiss [1976], to the analysis of various operators on X itself. A pioneering contribution was the paper of Berkson, Gillespie, and Muhly [1986], who showed, among several other results, that every *power-bounded* operator $U \in \mathscr{L}(X)$ (which means that $\sup_{n\in\mathbb{Z}} \|U^n\|_{\mathscr{L}(X)} < \infty$) can be written as the Fourier–Stieltjes transform of an appropriate projection-valued measure (a *spectral family of projections*) and, as a corollary, has a logarithm in $\mathscr{L}(X)$. As a working engine for their transference, Berkson, Gillespie, and Muhly [1986] established the equivalence of the UMD property with the $\ell^p(\mathbb{Z};X)$-boundedness of the discrete Hilbert transform. Boundedness of vector-valued "Hilbert transforms" over more general groups, and their applications via transference to spectral decompositions of a UMD space X, were studied by Berkson, Gillespie, and Muhly [1989], and Asmar, Berkson, and Gillespie [1990].

The next decade witnessed new applications of the transference method involving more general harmonic analysis operators than just the Hilbert transform. In this direction, the work of Berkson and Gillespie [1994] was also instrumental in highlighting the role of R-boundedness, which will be discussed more thoroughly in Volume II. Asmar and Kelly [1996] explored applications of maximal convolution operators to ergodic theory, while Hieber and Prüss [1998] transferred the $L^p(\mathbb{R};X)$-boundedness of Mihlin-type Fourier multipliers to obtain an H^∞-calculus for generators of bounded C_0-groups on a UMD space X. Extensions to unbounded groups have been explored by Haase [2009].

Convergence of Fourier series

The connection of the Hilbert transform boundedness and the L^p-norm convergence of Fourier series, as in Proposition 5.2.7, has been well understood since the days of Riesz [1924]. Many variants of this argument appear in the literature; for instance, a similar result on the inversion of the Laplace transform in UMD spaces is contained in Haase [2008].

In contrast to the L^p-norm convergence, the pointwise convergence of Fourier series is a much more delicate issue. In the case of scalar-valued functions, this was a celebrated breakthrough of Carleson [1966], one of the highlights of 20th century Fourier analysis and well ahead of its time. It was only relatively recently that the pioneering ideas of Carleson were placed into a general context by Lacey and Thiele [2000], marking the birth of modern time–frequency analysis.

As for vector-valued functions, Rubio de Francia [1986] proved the pointwise convergence of Fourier series in $L^p(\mathbb{T};X)$, when $p \in (1,\infty)$ and X is a UMD *Banach function space*, leaving in particular open the case where $X = \mathscr{C}^q$ is a Schatten space with $q \in (1,\infty) \setminus \{2\}$. The method of Rubio de Francia [1986] ultimately reduced the problem to a situation where Carleson's classical theorem could be applied as a black box, pointwise on the Banach function space.

By extending the approach of Lacey and Thiele [2000] to the vector-valued situation, Hytönen and Lacey [2013] were able to extend these results to the more general class of *intermediate UMD spaces*, defined by the condition that $X = [Y,H]_\theta$ is a complex interpolation space between a Hilbert space H and a UMD space Y. That these spaces in fact include all UMD Banach function spaces is a deep theorem of Rubio de Francia [1986]. They also include the Schatten spaces $\mathscr{C}^q$ for $q \in (1,\infty)$, as is immediate by writing $\mathscr{C}^q = [\mathscr{C}^p, \mathscr{C}^2]_\theta$ for a suitable p such that $1 < p < q < 2$ or $2 < q < p < \infty$. The following related problems, already raised by Rubio de Francia [1986], remain open:

(a) Is every UMD space an intermediate UMD space (or a closed subspace thereof)?
(b) Is Carleson's theorem on pointwise convergence of Fourier series true in all UMD spaces?

By the results of Hytönen and Lacey [2013], a positive solution of (a) would provide a positive solution of (b), and conversely a counterexample to (b) would be a counterexample to (a).

In the direction of (b), Parcet, Soria, and Xu [2013] proved that the pointwise norms of the partial sums $\|P_n f(x)\|$ grow at most at the rate $o(\log\log n)$ as $n \to \infty$, for any UMD space X. It is also shown by Hytönen and Lacey [2013] that UMD is a necessary condition for (a version of) the pointwise convergence result.

Convergence of orthogonal polynomial expansions

While the connection of the (periodic) Hilbert transform with Fourier series is most intimate, its boundedness can, although less obviously, also be linked with the L^p norm convergence of some other classical orthogonal expansions. The following theorem is representative of such results:

Theorem 5.8.4 (König and Nielsen [1994]). *Let X be a Banach space, $1 \leqslant p \leqslant \infty$, $\alpha, \beta > -1$, and $w_{\alpha\beta}(t) := (1-t)^{\alpha}(1+t)^{\beta}$. Let Q_n denote the orthogonal projection of $L^2((-1,1), w_{\alpha,\beta})$ onto polynomials of degree at most n, and also the natural extension of this operator to $L^p((-1,1), w_{\alpha,\beta}; X)$. Then the following conditions are equivalent:*

(1) *$Q_n f \to f$ for all $f \in L^p((-1,1), w_{\alpha,\beta}; X)$, in the norm of this space;*
(2) *X is a UMD space and $m(\alpha,\beta) < p < m(\alpha,\beta)'$, where*

$$m(\alpha,\beta) := \widetilde{m}(\max\{-\tfrac{1}{2}, \alpha, \beta\}), \qquad \widetilde{m}(\gamma) := \frac{4(\gamma+1)}{2\gamma+3}.$$

As in the proof of Proposition 5.2.7, the convergence (1) is easily equivalent to the uniform boundedness of the operators Q_n, and the core of the argument of König and Nielsen [1994] consists of relating this to the boundedness of the Hilbert transform. A perhaps striking feature of (2) is that it deals with a range of exponents p different (if $\max(\alpha,\beta) > -\frac{1}{2}$) from the 'usual' $(1,\infty)$, which appears in both the UMD condition and the Hilbert transform boundedness. This is explained by the fact that König and Nielsen [1994] split the operator Q_n into two parts, one of which is essentially the Hilbert transform (and hence well-behaved for all $p \in (1,\infty)$, but only for UMD spaces X) and another one, which is dominated by a positive operator: hence its L^p-boundedness properties can be deduced from purely scalar-valued considerations, but they will only be valid in a certain range of p.

The orthogonal polynomials in $L^2((-1,1), w_{\alpha,\beta})$ are given by the *Jacobi polynomials*; results analogous to Theorem 5.8.4 for *Hermite polynomials* are due to König [1994]. Both papers also obtain related characterisations of interpolation inequalities for the respective quadrature methods.

On the definition of $H^1(\mathbb{T}; X)$

Classically, the periodic Hilbert transform, or conjugate function, is closely related to the theory of the Hardy space $H^1(\mathbb{T})$. In the Notes of Chapter 1 we have introduced the atomic Hardy space $H^1_{\mathrm{at}}(\mathbb{T}; X)$ as the space of all strongly measurable functions $f : \mathbb{T} \to X$ that can be represented an ℓ^1-sum of atoms, i.e.,

$$f = \sum_{n\geqslant 1} \lambda_n a_n$$

with each a_n an atom (in the sense of Definition 1.4.8) and $\sum_{n\geqslant 1} |\lambda_n| < \infty$. It is a celebrated result in harmonic analysis due to Burkholder, Gundy, and

Silverstein [1971] and Coifman [1974] (see Stein [1993] for the full story) that, in the scalar-valued case, the space $H^1_{\rm at}(\mathbb{T})$ coincides with the spaces $H^1_{\max}(\mathbb{T})$ and $H^1_{\rm conj}(\mathbb{T};X)$ up to equivalent norms. In the vector-valued context, the latter spaces are defined as follows. Firstly,

$$H^1_{\max}(\mathbb{T};X) := \{f \in L^1(\mathbb{T};X) : \sup_{0<r<1} \|P_r * f(\cdot)\| \in L^1(\mathbb{T})\},$$

with $(P_r)_{0<r<1}$ the Poisson kernel. This space is a Banach space with respect to the norm

$$\|f\|_{H^1_{\max}(\mathbb{T};X)} := \Big\| \sup_{0<r<1} \|P_r * f(\cdot)\| \Big\|_1.$$

Secondly,

$$\begin{aligned} &H^1_{\rm conj}(\mathbb{T};X) \\ &\qquad := \{f \in L^1(\mathbb{T};X) : \lim_{r\uparrow 1} Q_r * f \text{ exists a.e. and belongs to } L^1(T;X)\}, \end{aligned}$$

with $(Q_r)_{0<r<1}$ the conjugate Poisson kernel. The existence of the *conjugate function* as the almost everywhere limit

$$\widetilde{f} := \lim_{r\uparrow 1} Q_r * f$$

is automatic in the scalar-valued case and more generally when X is UMD, but has to be imposed as a separate condition in the case of a general Banach space X. The space $H^1_{\rm conj}(\mathbb{T};X)$ is a Banach space with respect to the norm

$$\|f\|_{H^1_{\rm conj}(\mathbb{T};X)} := \|f\|_1 + \|\widetilde{f}\|_1.$$

The equality $H^1_{\max}(\mathbb{T};X) = H^1_{\rm conj}(\mathbb{T};X)$ in the vector-valued case can be proved, for any Banach space X, by a routine extension of the proof for the scalar-valued case. Concerning the equivalence of these spaces with $H^1_{\rm at}(\mathbb{T};X)$ one has the following result of Blasco [1988]:

Theorem 5.8.5 (Blasco). *For any Banach X the following assertions are equivalent:*

(1) $H^1_{\max}(\mathbb{T};X) = H^1_{\rm conj}(\mathbb{T};X) = H^1_{\rm at}(\mathbb{T};X)$ *with equivalent norms;*
(2) X *is a UMD space.*

For a detailed treatment of vector-valued Hardy spaces we refer the reader to Pisier [2016].

Section 5.3

Most of the generalities on Fourier multipliers are straightforward adaptations of classical considerations in the scalar-valued case, as discussed e.g., in Stein

[1970a], Stein and Weiss [1971]. In the scalar-valued situation, a convenient initial domain of definition of the multiplier operators is given by $L^2(\mathbb{R}^d)$, which is invariant under the Fourier transform by Plancherel's theorem. This is no longer the case in the generality that we consider, which is the reason for using the space $\check{L}^1(\mathbb{R}^d; X)$ instead. This has the same key features as the classical choice $L^2(\mathbb{R}^d) = \check{L}^2(\mathbb{R}^d)$ of being preserved by the action of the operators T_m, while also being a space of proper functions rather than distributions.

The characterisation of the Fourier multipliers of $L^1(\mathbb{R}^d)$ as the Fourier transforms of complex Borel measures described in Remark 5.3.6 is discussed in Grafakos [2008, Section 2.5.4].

Theorem 5.3.15 on the necessity of R-bounded range is from Clément and Prüss [2001]; a weaker version of this phenomenon was also exhibited in Weis [2001], where the operator-valued Mihlin-type Theorem 5.3.18 was first obtained. The notion of R-boundedness has an interesting history of its own, but we postpone its discussion for Volume II.

The scalar-symbol case of Theorem 5.3.18 and the Littlewood–Paley inequality (Theorem 5.3.24) had already been obtained (essentially, with minor variations) by McConnell [1984] and, with a different method, by Bourgain [1984a, 1986b]. Indeed, the extension of the classical multiplier theorems to UMD spaces took place over two periods of activity: the case of scalar-valued multipliers was settled during the 1980's by McConnell [1984] and Bourgain [1984a, 1986b], and the operator-valued generalisation around and after the turn of the millennium, starting with Weis [2001]. We return to these development in the discussion of multipliers in several variables further below.

Our treatment of the multiplier theorems is derived from Bourgain [1984a, 1986b], which has been the more popular starting point for further developments in the literature. We offer a sketch of McConnell's alternative approach further below (see page 486).

Historically, and also in most other treatments, the simpler Littlewood–Paley inequality serves as an intermediate result in the proof of the more general multiplier theorem; we have partially abandoned this step-by-step approach to obtain the best available constants. Namely, instead of using the "final" form of the Littlewood–Paley inequality as in Theorem 5.3.24, we have only applied its "preliminary" form in Proposition 5.1.10 which, while being slightly more cumbersome to use, serves essentially the same needs with the added advantage of being valid with a better constant. The explicit constants in Theorems 5.3.18 and Theorem 5.3.24 seem to be new.

Prehistory

The theory of this section was first developed in its periodic version (and of course for scalar-valued functions) in the time between the two World Wars. The original Littlewood–Paley inequality (Theorem 5.3.24) for functions $f : \mathbb{T} \to \mathbb{C}$ was announced in Littlewood and Paley [1931] and proved in Littlewood and Paley [1936]. They formulated it as a square function estimate, but also stated an inequality exactly like that in Theorem 5.3.24 for fixed signs ϵ_I rather than the random ε_I. Their approach to the inequality was quite different from what we have seen here; it was based on lengthy computations dealing first with the case when $p = 2k$ is an even integer. This was most likely inspired by the approach of Riesz [1928] to the boundedness of the Hilbert transform, and thereby somewhat similar to, but rather more complicated than our computations in the proof of Proposition 5.4.2.

A periodic, scalar version of the Multiplier Theorem 5.3.18 was then obtained by Marcinkiewicz [1939b]. The scalar-valued theorem on the line $\mathbb{R}$, closely resembling our formulation of Theorem 5.3.18, was announced by Mihlin [1956] and proved in Mihlin [1957]. His method was an early instance of the transference method discussed in Section 5.7, taking the theorem of Marcinkiewicz [1939b] as a starting point.

It is hence clear that the credit for the primary invention of the multiplier theory belongs to Marcinkiewicz [1939b]. Our choice to of the name "Mihlin's multiplier theorem" is not meant in any way to diminish the pioneering role of Marcinkiewicz; rather, we were led to this choice to make a clear distinction between Mihlin's multiplier class and that of Marcinkiewicz, which is somewhat *larger*: while Mihlin's condition requires a uniform bound on $|\xi||m'(\xi)|$, the condition of Marcinkiewicz (or, to be precise, its direct analogue on the line) only assumes an integrated version:

$$\sup_{I\in\mathscr{I}} \int_I |m'(\xi)| \,\mathrm{d}\xi \leqslant C.$$

This more general condition could replace the Mihlin bound also in the vector-valued multiplier theorem with scalar-valued symbol; in fact, the result of Bourgain [1984a, 1986b] already took this form. For operator symbols, suitable modifications involving R-bounds can be made; see Haller, Heck, and Noll [2002], Štrkalj and Weis [2007] for two possible formulations.

A singular integral point of view

Using the duality of the product and convolution under the Fourier transform, the multiplier operators $T_m f = (m\widehat{f})^{\vee}$ admit, at least formally, an equivalent description

$$T_m f(x) = k * f(x) = \int_{\mathbb{R}} k(x-y) f(y) \,\mathrm{d}y, \qquad k := \check{m},$$

as convolution operators with in general singular kernel k. In the analysis of scalar-valued functions, the interplay of the multiplier and convolution points of view was systematically explored by Hörmander [1960], who gave new proofs and extensions of the multiplier theorem of Mihlin [1956] on the one hand, and results of Calderón and Zygmund [1952] on singular integrals on the other hand. Of course, the convolution kernels $k(x-y)$ are special cases of the kernels $K(x,y)$ discussed above in these Notes, but the convolution case was understood prior to the general case, both in the scalar-valued and the vector-valued situations. In the latter context of our primary interest, Bourgain [1984a, 1986b] obtained the following result, whose proof was a variant of his proof of the multiplier theorem:

Theorem 5.8.6 (Bourgain [1984a, 1986b]). *Let X be a UMD space and let $p \in (1,\infty)$. Let $k : \mathbb{R} \setminus \{0\} \to \mathbb{K}$ be an odd kernel (i.e., $k(-x) = -k(x)$) such that $K(x,y) := k(x-y)$ satisfies the standard estimates (5.65). Then*

$$\|k * f\|_{L^p(\mathbb{R};X)} \leqslant C_{p,X} \|f\|_{L^p(\mathbb{R};X)}.$$

(The oddness here is not an essential assumption, and has been relaxed in subsequent papers cited below, but some form of cancellation is needed.) In the special case when X is a UMD space with an unconditional basis $(e_j)_{j=1}^\infty$, Bourgain [1984b] gave a stronger version of this result with a completely different proof using techniques from weighted norm inequalities: If each k_j satisfies the assumptions of k in Theorem 5.8.6 with uniform constants, then

$$\Big\| \sum_{j=1}^\infty (k_j * f_j) e_j \Big\|_{L^p(\mathbb{R};X)} \leqslant C_{p,X} \Big\| \sum_{j=1}^\infty f_j e_j \Big\|_{L^p(\mathbb{R};X)}.$$

The operator on the left can be interpreted as a convolution of $f = \sum_{j=1}^\infty f_j e_j$ with an operator-valued kernel $\mathbf{k}$, where

$$\mathbf{k}(y)\Big(\sum_{j=1}^\infty x_j e_j \Big) := \sum_{j=1}^\infty k_j(y) x_j e_j \qquad \forall x = \sum_{j=1}^\infty x_j e_j \in X.$$

Convolution operators with operator-valued kernels in UMD spaces have been further studied by Hytönen and Weis [2007] and Hytönen and Portal [2008].

Littlewood–Paley inequalities for different systems of intervals

One may inquire about the validity of the Littlewood–Paley-type inequality

$$\Big\| \sum_{I \in \widetilde{\mathscr{I}}} \varepsilon_I \Delta_I f \Big\|_{L^p(\mathbb{R}\times\Omega;X)} \lesssim \|f\|_{L^p(\mathbb{R};X)} \tag{5.69}$$

for other collections $\widetilde{\mathscr{I}}$ of (disjoint) intervals in place of $\mathscr{I} = \{\pm(2^k, 2^{k+1}) : k \in \mathbb{Z}\}$ appearing in Theorem 5.3.24.

For $X = \mathbb{C}$, (5.69) holds for all collections $\widetilde{\mathscr{I}}$ of disjoint intervals of equal length if and only if $p \in [2,\infty)$; this was first proved by Carleson [1967], and then in different ways by both Córdoba [1981] and Rubio de Francia [1983]. Finally, Rubio de Francia [1985] proved that (5.69) remains valid in the same range of p for an *arbitrary* family $\widetilde{\mathscr{I}}$ of disjoint intervals. Another approach and a certain extension of this result is due to Bourgain [1985], who showed that the reverse inequality ($\lesssim$ replaced by $\gtrsim$) in (5.69) holds for all $p \in [1, 2]$ provided that $\widetilde{\mathscr{I}}$ is a partition of $\mathbb{R}$; the range $p \in (1, 2]$ is equivalent to the result of Rubio de Francia [1985] by straightforward duality considerations, but the case $p = 1$ is presents a genuine novelty.

A vector-valued analogue of the inequality (5.69) of Rubio de Francia [1985] for arbitrary systems of intervals $\widetilde{\mathscr{I}}$, under the assumption that X is a UMD space of type 2, was announced by Potapov and Sukochev [2005], together with applications to non-commutative L^p spaces, including an extension of old results of Birman and Solomyak [1967] on Schur multipliers. Unfortunately, the proof turned out to have a gap and was never published, leaving the validity of the announced result a major open problem in the area.

Already earlier, Berkson, Gillespie, and Torrea [2004] axiomatised the *Littlewood–Paley–Rubio de Francia property* LPR_p of a Banach space X as the validity of the bound (5.69) for all disjoint collections $\widetilde{\mathscr{I}}$ and all functions $f \in L^p(\mathbb{R}; X)$, and they explored some basic properties of this notion. Hytönen, Torrea, and Yakubovich [2009] showed that the case of equal-length intervals is valid if and only if X is a UMD space of type 2; in particular, these are necessary conditions for the general LPR_p property, but the converse to this statement remains open. The best available sufficient condition is due to Potapov, Sukochev, and Xu [2012]: they show that every Banach lattice whose 2-*concavification* is a UMD Banach lattice has LPR_p for all $p \in [2, \infty)$.

A corollary of the general Littlewood–Paley inequality of Rubio de Francia [1985] is a sharper form of the Marcinkiewicz multiplier theorem, obtained by Coifman, Rubio de Francia, and Semmes [1988]. Analogously, under the assumption of the LPR_p property of a Banach space X, improved multiplier theorems in $L^p(\mathbb{R}; X)$ were proved by Hytönen and Potapov [2006].

Littlewood–Paley inequalities for diffusion semigroups

Classical analysis abounds in estimates that deserve to be called Littlewood–Paley inequalities. In contrast to the discrete summations that we have treated, many of them involve a continuous quadratic integral. This is in particular the case when $(P_t)_{t>0}$ is a semigroup of operators (i.e., $P_{t+s} = P_t P_s$, together with some continuity assumptions) such as the classical Poisson semigroup, and one studies the square function (or "g function")

$$Gf(x) := \Big(\int_0^\infty |t\partial_t P_t f(x)|^2 \frac{\mathrm{d}t}{t} \Big)^{1/2}$$

or one of its many variants. (In the case of the Poisson semigroup on $\mathbb{T}$, identified with the complex unit circle, it is customary to study an analogous expression involving the re-reparametrisation $r = e^{-t}$ so that $f(rz) := P_t f(z)$ is the Poisson extension of f from the circle to the disk.) In Volume II, we shall present a systematic framework, developed by Kalton and Weis [2016], for randomised formulations of such continuous quadratic expressions, which leads to vector-valued versions with far-reaching applications. However, for the sake of Littlewood–Paley theory in its own right, one might study a more direct vector-valued generalisation by simply replacing absolute values with norms, leading to

$$G_q f(x) := \Big(\int_0^\infty \|t\partial_t P_t f(x)\|_X^q \frac{\mathrm{d}t}{t} \Big)^{1/q},$$

where f is an X-valued function, and we have also allowed a generic exponent q in place of the classical value of 2; the need for this will be explained shortly.

A version of the classical Littlewood–Paley inequality says that

$$A_p \|f - \langle f \rangle\|_{L^p(\mathbb{T})} \leqslant \|Gf\|_{L^p(\mathbb{T})} \leqslant B_p \|f\|_{L^p(\mathbb{T})}$$

for all $p \in (1, \infty)$. A vector-valued extension reads as follows:

Theorem 5.8.7 (Xu [1998]). *Let X be a Banach space and $p, q \in (1, \infty)$.*

(1) *There is a constant $A = A_{p,q,X} > 0$ such that*

$$A \|f - \langle f \rangle\|_{L^p(\mathbb{T};X)} \leqslant \|G_q f\|_{L^p(\mathbb{T})} \tag{5.70}$$

if and only if X has martingale type q.

(2) *There is a constant $B = B_{p,q,X} < \infty$ such that*

$$\|G_q f\|_{L^p(\mathbb{T})} \leqslant B \|f\|_{L^p(\mathbb{T};X)} \tag{5.71}$$

if and only if X has martingale cotype q.

Thus, (5.70) can only hold for $q \in (1, 2]$ and (5.71) for $q \in [2, \infty)$, and both of them simultaneously only if $q = 2$ and X is isomorphic to a Hilbert space.

Martínez, Torrea, and Xu [2006] extended this result to abstract subordinated diffusion semigroups (in the sense of Stein [1970b]) in place of the classical Poisson semigroup. In this case, the average $\langle f \rangle$ should be replaced by the projection of f onto the space of functions invariant under the action of the semigroup.

If one likes to have a two-sided inequality (both (5.70) and (5.71)) for a non-Hilbertian space X, then G_q has to be replaced by a different vector-valued extension of the classical square function, using a framework that we shall discuss in Volume II. For diffusion semigroups, this programme was carried out by Hytönen [2007], adapting the classical approach of Stein [1970b] to UMD spaces.

Section 5.4

Our proof of the Hilbert transform boundedness in the Schatten spaces (Proposition 5.4.2) is taken from Bourgain [1986b]. However, it is mentioned by Pisier and Xu [1997] (p. 668) that the possibility of adapting M. Riesz's proof for this purpose had been known "for many years", and that Pisier indeed learned it from P. Muhly already in 1976.

The quantitative norm bound obtained in Proposition 5.4.2 and Corollary 5.7.8 is exact in the case when $p = 2^n$, $n = 1, 2, 3, \ldots$. In fact, the norm of the scalar-valued Hilbert transform is known to be exactly

$$\|\widetilde{H}\|_{\mathscr{L}(L^p(\mathbb{T};\mathbb{K}))} = \cot \frac{\pi}{2p^*}, \qquad \forall p \in (1, \infty),$$

as was first proved by Pichorides [1972]. The case of powers of two was accomplished somewhat earlier by Gohberg and Krupnik [1968], who also proved the lower bound and conjectured the upper bound in the general case. This also shows that the bound of Corollary 5.7.8 is at most a factor 2 away from optimal for a general $p \in (1, \infty)$.

From Proposition 5.4.2 and Corollary 5.2.11 (the comparison of the UMD and Hilbert transform constants), we would deduce that $\beta_{p,\mathscr{C}^p}$ grows at most at the rate $O(p^*)^2$. This, however, can be improved to

$$\beta_{p,\mathscr{C}^p} \leqslant Cp^*,$$

for some constant C, as shown by Randrianantoanina [2002] as a consequence of a more general bound for *non-commutative martingale transforms*. This bound also has the correct order, observing that $\beta_{p,\mathscr{C}^p} \geqslant \beta_{p,\mathbb{K}} = p^* - 1$.

Theorem 5.4.3 ($\|M^e_m\|_{\mathscr{L}(\mathscr{C}^p)} \leqslant \|T_m\|_{\mathscr{L}(L^2(\mathbb{R};\mathscr{C}^p))}$) is implicit in the literature. A variant of its Corollary 5.4.6 (for m in the Mihlin class) is formulated by Potapov and Sukochev [2011, Theorem 4], but a transference argument akin to that of Theorem 5.4.3 is precisely what goes into the proof.

What we obtained as Corollary 5.4.5 is originally much older, and attributed to Macaev [1961]; see Gohberg and Kreĭn [1970, pp. 118–119] for a discussion of the original approach and some alternative strategies. While we proved Corollary 5.4.5 by transferring Proposition 5.4.2 on the boundedness of the $\mathscr{C}^p$-valued Hilbert transform, Gohberg and Kreĭn [1970] provide an argument more or less equivalent to the proof of Proposition 5.4.2, but directly on the level of $\mathscr{C}^p$ instead of $L^p(\mathbb{T};\mathscr{C}^p)$.

Theorem 5.4.7 ($\|f(u) - f(v)\|_{\mathscr{C}^p} \leqslant c_p \|f\|_{\mathrm{Lip}} \|u - v\|_{\mathscr{C}^p}$) is the special case for compact operators of the main result of Potapov and Sukochev [2011], which resolved an old problem related to a conjecture of Kreĭn [1964]. We refer to Potapov and Sukochev [2011] for an account of the intermediate progress on this question. Let us only mention that even the case that $f(t) = |t|$ is the absolute value was *a priori* not clear, and was first resolved by Davies [1988]. (In the above formalism, which is borrowed from Potapov and Sukochev [2011], this case is a rather quick consequence of Corollary 5.4.5, noting that $\phi_f(\lambda, \mu) = \mathrm{sgn}(\lambda - \mu)$, so that the Schur multiplier Φ_f is essentially just the triangular truncation in this case.)

As already mentioned, the compactness condition in Theorem 5.4.7 is actually redundant, and absent in the formulation of the main theorem of Potapov and Sukochev [2011]. In the general case, the proof of Potapov and Sukochev [2011] involves some additional approximation arguments, and relies on the theory of *double operator integrals* developed by De Pagter, Witvliet, and Sukochev [2002].

A quantitative refinement of Theorem 5.4.7 is obtained by Caspers, Montgomery Smith, Potapov, and Sukochev [2014], who show that the optimal constant c_p is of the order p^*. This depends on a more delicate construction of the auxiliary Fourier multipliers from which the relevant Schur multiplier is recovered. They also prove an extension of Proposition 5.4.8 to commutators of n self-adjoint operators.

Section 5.5

The d-dimensional Multiplier Theorem 5.5.10, in the operator-valued generality as stated, was first obtained by Štrkalj and Weis [2007], a paper that took over seven years to appear. ("Received by the editors October 1, 1999"; "electronically published on March 20, 2007".) In the mean time, another proof by Haller, Heck, and Noll [2002] was also published. They were both based on the analogous result and its proof for scalar-valued multipliers due to Zimmermann [1989], which in turn was an extension of the one-dimensional strategy of Bourgain [1984a, 1986b]. An earlier d-dimensional multiplier theorem by McConnell [1984] required the bounds $|\xi|^{|\alpha|}|\partial^\alpha m(\xi)| \leqslant C$ for a larger number of derivatives, namely, all those with total order $|\alpha| \leqslant d+1$.

Extensive applications of the multiplier theorem to elliptic and parabolic problems were investigated by Denk, Hieber, and Prüss [2003], Kunstmann and Weis [2004], Prüss and Simonett [2016]. The number of required derivatives $\partial^\alpha m$ in the Mihlin conditions, under additional conditions on the Banach spaces X and Y, have been investigated by Girardi and Weis [2003] and Hytönen [2004, 2010]. Analogues of multiplier theorems for more general *pseudo-differential operators*

$$T_a f(x) = \int_{\mathbb{R}^d} a(x,\xi)\widehat{f}(\xi)e^{i2\pi x\cdot\xi}\,\mathrm{d}\xi$$

were studied by Portal and Štrkalj [2006] and Hytönen and Portal [2008].

In the original scalar-valued case, Theorem 5.5.10 corresponds to the formulation of Mihlin [1957]. A closer $\mathbb{R}^d$-analogue of the periodic original of Marcinkiewicz [1939b] was given by Lizorkin [1963], using less restrictive derivative bounds of the form

$$|\xi^\alpha||\partial^\alpha m(\xi)| \leqslant C, \qquad \xi^\alpha := \prod_{i=1}^d \xi_i^{\alpha_i}$$

for all $\alpha \in \{0,1\}^d$; note that $|\xi^\alpha| \leqslant |\xi|^{|\alpha|}$, and it can be much smaller. It was discovered by Zimmermann [1989] that the vector-valued analogue of Lizorkin's theorem is *not* valid in all UMD spaces, but requires an additional condition on the Banach space X.

Fourier multiplier operators from $L^p(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$ with $p \neq q$ have not been considered in this chapter. It is well known that only the zero operator is bounded in the case $1 \leqslant q < p \leqslant \infty$. The theory of multiplier operators from $L^p(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$ with $1 \leqslant p < q \leqslant \infty$ is non-trivial. We refer to Hörmander [1960] and Lizorkin [1963] for detailed treatments. Fourier multiplier operators from $L^p(\mathbb{R}^d; X)$ to $L^q(\mathbb{R}^d; Y)$ have been studied in Rozendaal and Veraar [2016a,b] under various geometric conditions on the Banach space X and Y. Here R-boundedness plays a crucial role again, but the UMD property is not required. Results on the torus are derived as well by applying the transference technique of Lemma 2.4.19. The methods of Section 5.7 seem to be inapplicable if $p < q$.

McConnell's approach to Mihlin's multiplier theorem

Following McConnell [1984], we outline a different route to the Fourier multiplier theorems than the one followed above. It exploits the UMD condition in a more direct way but at the cost of requiring more sophisticated probabilistic arguments. The key steps of this method will be reproduced here without going to the full details.

The starting point is the classical-style Poisson integral representation (valid at least for f and m in appropriate test function classes)

$$\begin{aligned} T_m f(x) &= K * f(x) \\ &= \frac{(-4)^{d+2}}{(d+1)!} \int_0^\infty \int_{\mathbb{R}^d} s^{d+1} \partial_s^{d+1} P^s K(y) \partial_s P^{3s} f(x-y) \,\mathrm{d}y \,\mathrm{d}s, \end{aligned} \tag{5.72}$$

where K is the inverse Fourier transform of m and P^s denotes the Poisson semigroup,

$$P^s f(x) = \int_{\mathbb{R}^d} p_s(x-y) f(y) \,\mathrm{d}y, \qquad p_s(y) := \frac{c_d \cdot s}{(|y|^2 + s^2)^{(d+1)/2}},$$

where $c_d = \Gamma(\frac{1}{2}(d+1))\pi^{-(d+1)/2}$.

After truncation and discretisation of the s-integral in (5.72), one is led to estimate the function

$$g(x) := \sum_{i=0}^{N-1} \int_{\mathbb{R}^d} s_i^{d+1} \partial_{s_i}^{d+1} P^{s_i} K(y) [P^{2s_{i+1}+s_i} f(x-y) - P^{3s_i} f(x-y)] \,\mathrm{d}y, \tag{5.73}$$

with bounds independent of the discretising sequence $0 = s_0 < s_1 < \ldots < s_N$.

Probabilistic preparations. Now some probabilistic armoury is brought to the battlefield. Let W_t and B_t be independent d-dimensional and 1-dimensional

Brownian motions on a probability space Ω. We introduce the stopping times $\tau(s) := \inf\{t > 0 : B_t \leqslant s\}$, and denote by $\mathbb{P}_{(x,s)}$ and $\mathbb{E}_{(x,s)}$ the probability and expectation governing the $(d+1)$-dimensional Brownian motion (W_t, B_t) when the initial value is $(x,s) \in \mathbb{R}_+^{d+1} := \mathbb{R}^d \times (0,\infty)$. Furthermore, use will be made of the random variables $(W_t^{/x}, B_t^{/x})$ which correspond (in a certain rigorous way) to the intuitive notion of Brownian motion conditioned to exit $\mathbb{R}_+^{d+1}$ at the assigned point $(x,0)$. The probability and the expectation associated to the initial value (y,s) of these random variables are denoted by $\mathbb{P}_{(y,s)}^{/x}$ and $\mathbb{E}_{(y,s)}^{/x}$, respectively. Finally for $(x,r) \in \mathbb{R}_+^{d+1}$, one defines

$$\mathbb{P}^{x,r} := \int_{\mathbb{R}^d} p_r(x-y)\mathbb{P}_{(y,r)}^{/x}\,\mathrm{d}y,$$

which intuitively corresponds to the probability distribution of a Brownian motion known to exit $\mathbb{R}_+^{d+1}$ at $(x,0)$, but without specifying its point of origin: McConnell shows that the distribution of $W_{\tau(s)}$ under $\mathbb{P}^{x,r}$ is actually independent of the initial level r whenever $0 < s < r$, so that Kolmogorov's extension theorem provides a consistent limit measure $\mathbb{P}^x = \lim_{r\to\infty}\mathbb{P}^{x,r}$. When u is a bounded harmonic function on $\mathbb{R}_+^{d+1}$, McConnell proves the equality

$$u(x, s_1 + s_2) = \mathbb{E}^x u(W_{\tau(s_1)}, s_2); \tag{5.74}$$

furthermore, under $\mathbb{P}^x$, the process $(W_{\tau(s)})_{s>0}$ has independent increments, whose distribution is given by the density function $p_{s-s'}(y)$.

With the help of (5.74), the Poisson integrals in (5.73) admit a probabilistic formulation,

$$\begin{aligned} P^{2s_{i+1}+s_i} f(x-y) &= \mathbb{E}^x P^{s_{i+1}+s_i} f(W_{\tau(s_{i+1})} - y), \\ P^{3s_i} f(x-y) &= \mathbb{E}^x P^{2s_i} f(W_{\tau(s_i)} - y). \end{aligned}$$

Writing $s_i^{d+1}\partial_s^{d+1}P^{s_i}K(y) =: e(y,s_i)p_{s_i}(y)$, the y-integration in (5.73) may also be interpreted as an expectation $\widetilde{\mathbb{E}}^0$ of an independent Brownian motion (on another probability space $\widetilde{\Omega}$) conditioned to exit at $(0,0)$:

$$\begin{aligned} g(x) &= \widetilde{\mathbb{E}}^0\mathbb{E}^x \sum_{i=0}^{N-1} e(\widetilde{W}_{\widetilde{\tau}(s_i)}, s_i)[P^{s_{i+1}+s_i} f(W_{\tau(s_{i+1})} - \widetilde{W}_{\widetilde{\tau}(s_i)}) \\ &\qquad\qquad\qquad\qquad - P^{2s_i} f(W_{\tau(s_i)} - \widetilde{W}_{\widetilde{\tau}(s_i)})] \\ &=: \widetilde{\mathbb{E}}^0\mathbb{E}^x \sum_{i=0}^{N-1} e(\widetilde{W}_{\widetilde{\tau}(s_i)}, s_i)\Delta_i f =: \widetilde{\mathbb{E}}^0\mathbb{E}^x G. \end{aligned}$$

By Jensen's inequality and the definition of $\mathbb{E}^x$, we have

$$\begin{aligned}
\int_{\mathbb{R}^d} \|g(x)\|_X^p \,\mathrm{d}x &\leqslant \int_{\mathbb{R}^d} \widetilde{\mathbb{E}}^0 \mathbb{E}^x \|G\|_X^p \,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} \widetilde{\mathbb{E}}^0 \int_{\mathbb{R}^d} p_r(x-y) \mathbb{E}_{(y,r)}^{/x} \|G\|_X^p \,\mathrm{d}y \,\mathrm{d}x \qquad (5.75)\\
&= \int_{\mathbb{R}^d} \widetilde{\mathbb{E}}^0 \int_{\mathbb{R}^d} p_r(x-y) \mathbb{E}_{(y,r)}^{/x} \|G\|_X^p \,\mathrm{d}x \,\mathrm{d}y = \int_{\mathbb{R}^d} \widetilde{\mathbb{E}}^0 \mathbb{E}_{(y,r)} |G|_X^p \,\mathrm{d}y,
\end{aligned}$$

the last equality being the equivalence of the expectation with the average of conditional expectations, and for r we can take any number $r > s_N$.

The key step. Now comes the heart of the argument. For a fixed point $\widetilde{\omega} \in \widetilde{\Omega}$, $(\Delta_i f)_{i=0}^{N-1}$ is a martingale difference sequence (with respect to a decreasing filtration) on the probability space $(\Omega, \mathbb{P}_{(y,r)})$. Moreover, the transforming sequence $(e(\widetilde{W}_{\widetilde{\tau}(s_i)}, s_i))_{i=0}^{N-1}$ is a bounded one, with a uniform bound (depending only on d) for all multipliers m which satisfy the estimates $|\xi|^{|\alpha|}|\partial^\alpha m(\xi)| \leqslant 1$ for all multi-indices $|\alpha| \leqslant d+1$: this is a standard computation starting from the formula

$$e(y,s) = c_d^{-1}(|y|^2+s^2)^{(d+1)/2} s^n \frac{1}{2\pi}\int_{\mathbb{R}^d} e^{i\xi\cdot y} m(\xi)|\xi|^{d+1} e^{-|\xi|s}\,\mathrm{d}\xi.$$

Thus we have

$$\int_{\mathbb{R}^d} \|g(x)\|_X^p \,\mathrm{d}x \leqslant (C_d \beta_p^+)^p \int_{\mathbb{R}^d} \mathbb{E}_\varepsilon \widetilde{\mathbb{E}}^0 \mathbb{E}_{(y,r)} \Big\| \sum_{i=0}^{N-1} \varepsilon_i \Delta_i f \Big\|_X^p \,\mathrm{d}y,$$

and by (5.75), we may replace $\mathbb{E}_{(y,r)}$ by $\mathbb{E}^y$ inside the integral. The next point to observe is the equality of the distributions of $(W_{\tau(s_i)} - \widetilde{W}_{\widetilde{\tau}(s_i)}, W_{\tau(s_{i+1})} - \widetilde{W}_{\widetilde{\tau}(s_i)})_{i=0}^{N-1}$ under $\mathbb{P}^y \otimes \widetilde{\mathbb{P}}^0$ and the distribution of $(W_{\tau(2s_i)}, W_{\tau(s_{i+1}+s_i)})_{i=0}^{N-1}$ under $\mathbb{P}^y$. Replacing $\mathbb{E}^y$ by $\mathbb{E}_{(y,r)}$ again, the previous integral is equal to

$$\int_{\mathbb{R}^d} \mathbb{E}_\varepsilon \mathbb{E}_{(y,r)} \Big\| \sum_{j=0}^{2(N-1)} \varepsilon_j \mathbf{1}_{2\mathbb{N}}(j) [P^{r_{j+1}} f(W_{\tau(r_{j+1})}) - P^{r_j} f(W_{\tau(r_j)})] \Big\|_X^p \,\mathrm{d}y,$$

where $r_{2i} = 2s_i$, $r_{2i+1} = s_{i+1}+s_i$. The quantities in brackets constitute a martingale difference sequence with respect to $\mathbb{P}_{(y,r)}$, so that we may continue with

$$\begin{aligned}
&\leqslant (\beta_p^-)^p \int_{\mathbb{R}^d} \mathbb{E}_{(y,r)} \Big\| \sum_{j=0}^{2(N-1)} P^{r_{j+1}} f(W_{\tau(r_{j+1})}) - P^{r_j} f(W_{\tau(r_j)}) \Big\|_X^p \,\mathrm{d}x \\
&= (\beta_p^-)^p \int_{\mathbb{R}^d} \mathbb{E}^x \Big\| f(x) - P^r f(W_{\tau(r)}) \Big\|_X^p \,\mathrm{d}x \\
&= (\beta_p^-)^p \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \Big\| f(x) - P^r f(x+y) \Big\|_X^p p_r(y)\,\mathrm{d}y\,\mathrm{d}x, \qquad r := r_{2N-1}.
\end{aligned}$$

The double integral approaches $\|f\|_{L^p_X}^p$ as $r = s_N + s_{N-1} \to \infty$. On the other hand, as the grid $0 < s_1 < \ldots < s_N$ gets finer and $s_N \to \infty$, our g approaches $T_m f$. Thus we have shown that $\|T_m f\|_p \leqslant C_d \beta_p^+ \beta_p^- \|f\|_p$. As in the estimation of the Hilbert transform, a two-fold application of the UMD inequality (or more precisely, one application of both the decoupling and coupling inequalities) was required to derive the desired boundedness property.

Section 5.6

Our proofs of Theorems 5.6.1 and 5.6.9 and some other results in this section follow the ones for the scalar case in Seeley [1971]. The same method can be applied to prove similar results on domains of fractional powers for operators other than $1 - \Delta$. The vector-valued extensions presented here have been obtained by Amann [2009] in a more general context of anisotropic function spaces. To deal with this setting, Amann [2009] needs a different version of the multiplier theorem, for which he assumes an additional property (besides UMD) on the underlying Banach space X. Specialising his results to the isotropic function spaces considered here, this additional assumption becomes redundant.

A different proof tracking the dependence of the equivalence constants on $\beta_{p,X}$ is found in Hytönen, Li, and Naor [2016], where this dependence is needed as a step in deriving a quantitative form of a *uniform affine approximation property* arising from the work of Bates, Johnson, Lindenstrauss, Preiss, and Schechtman [1999]. Unlike the present approach built on the general multiplier theorems, their proof makes selective use of special multipliers for which one can obtain sharper estimates than those available via the general theory.

Littlewood–Paley inequalities for the Bessel potential spaces have been obtained in Meyries and Veraar [2015], where they were used to derive the boundedness of the mapping $f \mapsto \mathbf{1}_{\mathbb{R}^d_+} f$ on $H^{s,p}(\mathbb{R}^d; X)$ for UMD spaces X, $p \in (1,\infty)$, and $s \in (-1/p', 1/p)$. Lindemulder [2015] contains a different approach based on an equivalent difference norm for the Bessel potential spaces.

Section 5.7

The vector-valued transference results presented here are straightforward adaptations of well-known results in the scalar-valued case, as treated e.g., in Grafakos [2008, Section 3.6.2]. Corollary 5.7.6 on homogeneous multipliers is explicitly stated in the vector-valued setting in Geiss, Montgomery-Smith, and Saksman [2010].

As we already discussed in the Notes right above, the periodic versions of the multiplier theorems and the Littlewood–Paley inequality are actually those that came first, in the historical line of development, and the transference method was originally used, by Mihlin [1957], to obtain the result in $\mathbb{R}^d$ from those on $\mathbb{T}^d$ that were already established by Marcinkiewicz [1939b]. As in the case of $\mathbb{R}^d$, we have only presented a simpler Mihlin-type formulation of the Multiplier Theorem 5.7.9, which appears in this form in Arendt and Bu [2002]. More general forms of this theorem and its d-dimensional version (Theorem 5.7.11) can be found in Štrkalj and Weis [2007].

Multipliers of analytic Hardy spaces and the AUMD property

As we have seen in this Chapter, the UMD property is the correct generality for various results on the boundedness of Fourier multipliers on Bochner spaces. The situation changes, however, if we consider the action of the same operators on the analytic Hardy spaces

$$\begin{aligned} H^p_{\mathrm{an}}(\mathbb{T};X) &:= \{f \in L^p(\mathbb{T};X) : \widehat{f}(k) = 0 \ \forall k < 0\}, \\ H^p_{\mathrm{an}}(\mathbb{R};X) &:= \{f \in L^p(\mathbb{R};X) : \operatorname{supp} \widehat{f} \subseteq [0,\infty)\}, \end{aligned}$$

where $\operatorname{supp} \widehat{f}$ must, in general, be interpreted in a distributional sense. The analytic UMD property introduced by Garling [1988] (as discussed in the Notes of Chapter 4, see page 372) turns out to provide the correct set-up for such questions, as first shown by Blower [1990]:

Theorem 5.8.8 (Blower [1990]). *Let X be a complex Banach space. Then X is an analytic UMD space if and only if $\widetilde{T}_m \in \mathscr{L}(H^1_{\mathrm{an}}(\mathbb{T};X))$ for every $m : \mathbb{Z} \to \mathbb{C}$ which satisfies the strong Mihlin condition*

$$\begin{aligned} \sup_{k\geqslant 0} |m(k)| &+ \sup_{k\geqslant 1} k|m(k) - m(k-1)| \\ &+ \sup_{k\geqslant 2} k^2 |m(k) - 2m(k-1) + m(k-2)| < \infty. \end{aligned} \tag{5.76}$$

Bu and Le Merdy [2007] showed that (5.76) remains sufficient for $\widetilde{T}_m \in \mathscr{L}(H^p_{\mathrm{an}}(\mathbb{T};X))$ for all $p \in [1,\infty)$, while its continuous analogue

$$\sup_{\xi \in (0,\infty)} \Big(|m(\xi)| + \xi |m'(\xi)| + \xi^2 |m''(\xi)| \Big) < \infty$$

is sufficient for $\widetilde{T}_m \in \mathscr{L}(H^p_{\mathrm{an}}(\mathbb{R};X))$; moreover, replacing the supremum by an R-bound yields a sufficient condition for operator-valued multipliers. A multi-dimensional extension is due to Bu [2011].

The proof devised by Blower [1990] for Theorem 5.8.8 was a relative of the approach of McConnell [1984] to the L^p version, outlined in the Notes above. The extension by Bu and Le Merdy [2007] was based on the same strategy.

Transference to Hermite, Bessel, Laguerre and Schrödinger settings

Classical Fourier analysis is sometimes legitimately referred to as the "harmonic analysis of the Laplace operator Δ"; indeed, the variable ξ of the Fourier transform is itself (up to a constant factor depending on the chosen normalisation) the transform of the gradient ∇ appearing in the standard factorisation $-\Delta = -\operatorname{div}\nabla = \nabla^*\nabla$, while the important radial Fourier multipliers, like the imaginary powers $|\xi|^{is}$, are simply functions of the transformed Laplacian $|\xi|^2$ itself. Given another (differential) operator $\mathcal{L}$, especially one with a similar factorisation $\mathcal{L} = \mathcal{D}^*\mathcal{D}$, the consideration of analogous questions for the "Riesz transforms" $\mathcal{D}\mathcal{L}^{-1/2}$, "Poisson semigroups" $e^{-t\mathcal{L}^{1/2}}$, imaginary powers $\mathcal{L}^{is}$, etc., leads to an entire new "harmonic analysis of the operator $\mathcal{L}$".

The formal functional expressions $\phi(\mathcal{L})$ here may be defined either directly via an explicit orthogonal transformation that diagonalises $\mathcal{L}$, via the spectral theorem for self-adjoint operators, or via the general framework of *functional calculus*, as presented e.g., in Haase [2006] or a subsequent volume of the present work. This is available for all the operators $\mathcal{L}$ of interest in the subsequent discussion.

The following particular cases have been studied in some detail in the UMD-valued situation:

- on the measure space $(\mathbb{R}^d, \gamma)$, where $\mathrm{d}\gamma(x) := e^{-|x|^2/2}\,\mathrm{d}x/\sqrt{(2\pi)^d}$,
  - the Ornstein–Uhlenbeck operator $\mathcal{O} = -\Delta + x\cdot\nabla$;
- on the measure space $(\mathbb{R}^d, \mathrm{d}x)$,
  - the Hermite operator $H = -\Delta + |x|^2$, and more general
  - Schrödinger operators $H_V = -\Delta + V$, where $V \geqslant 0$ is a potential;
- on the measure space $(\mathbb{R}_+, \mathrm{d}x)$, the one-parameter families of
  - Bessel operators $S_\alpha = -\Delta + \alpha(\alpha-1)/x^2$, and
  - Laguerre operators $L_\alpha = -\Delta + x^2 + \alpha(\alpha-1)/x^2$,

  where $\alpha > 0$ in both cases and $\Delta = \mathrm{d}^2/\mathrm{d}x^2$.

The Hermite, Bessel and Laguerre operators are closely related to classical special functions. The latter two appear in various forms in the literature, as listed in Almeida, Betancor, and Castro [2014, Table on p. 23]. Results concerning one of the versions can often be translated into the other versions by a simple coordinate change, as in Almeida, Betancor, and Castro [2014, Lemma 1.3].

A sample of results achieved in these settings is as follows:

Theorem 5.8.9. *Let X be a Banach space and let $p \in (1,\infty)$. Let $\alpha > -1$, $d \geqslant 1$, $n \geqslant 3$, and let $V : \mathbb{R}^n \to [0,\infty)$ satisfy the* reverse Hölder inequality

$$\Big(\fint_B V^q\,\mathrm{d}x\Big)^{1/q} \leqslant C \fint_B V\,\mathrm{d}x$$

for some $q \geqslant n/2$, uniformly over all balls $B \subseteq \mathbb{R}^n$. Then the following conditions are equivalent:

(1) *X is a UMD space;*
(2) *the "Riesz transforms" $\partial_j \mathcal{O}^{-1/2}$ are bounded on $L^p(\mathbb{R}^d,\gamma)$ for all $j = 1,\dots,d$;*

(3) *the imaginary powers $\mathcal{L}^{is}$ are bounded on $L^p(\Omega;X)$ for all $s\in\mathbb{R}$, for any of the following choices:*

$$(\mathcal{L},\Omega)\in\big\{(H,\mathbb{R}),\ (L_\alpha,\mathbb{R}_+),\ (S_\alpha,\mathbb{R}_+),\ (H_V,\mathbb{R}^n)\ \textit{with}\ n\geqslant 3\big\}.$$

References are Harboure, Torrea, and Viviani [2003] (for (2)), and Betancor, Castro, Curbelo, and Rodríguez-Mesa [2013a], Betancor, Castro, and Rodriguez-Mesa [2016], Betancor, Crescimbeni, Fariña, and Rodríguez-Mesa [2013b] (for (3)). Several further characterisations of UMD in terms of the respective Poisson semigroups $e^{-t\mathcal{L}^{1/2}}$ depend on notions that we discuss in more detail in Volume II, namely, Littlewood–Paley inequalities based on γ-*radonifying norms* (Betancor, Castro, Curbelo, Fariña, and Rodríguez-Mesa [2012, 2014], Betancor, Castro, and Rodríguez-Mesa [2015]), and L^p-*maximal regularity* (Almeida, Betancor, and Castro [2014]).

All these results are proved by comparison (or "transference") of a "new" operator $T^{\mathcal{L}}$ with an "old" $T^{\mathcal{L}_0}$, whose boundedness is already known to characterise UMD; here $\mathcal{L}_0$ is either the negative Laplacian $-\Delta$, or another operator, for which the result is already established. Recall that the characterisation of UMD by the classical Riesz transforms $R_j=\partial_j(-\Delta)^{-1/2}$ is Theorem 5.5.1; for the imaginary powers $(-\Delta)^{is}$, it is a result of Guerre-Delabrière [1991], which in turn is a variant of the argument of Bourgain [1983] for Theorem 5.2.10.

The working engine of the comparison method is a partition of the respective operators into their "local" and "global" parts, a strategy that is nicely explained by Harboure, Torrea, and Viviani [2003, Section 2], and attributed by them to the work of Muckenhoupt [1969] in the scalar-valued setting. If T is one of the operators appearing in Theorem 5.8.9, or its analogue in the context of the Laplacian Δ, then it has a representation

$$Tf(x)=\int_\Omega K(x,y)f(y)\,\mathrm{d}y,$$

and the local and global parts are defined by

$$\begin{aligned}T_{\mathrm{loc}}f(x)&:=\int_\Omega K(x,y)\varphi(x,y)f(y)\,\mathrm{d}y,\\ T_{\mathrm{glob}}f(x)&:=\int_\Omega K(x,y)(1-\varphi(x,y))f(y)\,\mathrm{d}y,\end{aligned}$$

for a suitable smooth cut-off φ such that $\mathbf{1}_N(x,y)\leqslant\varphi(x,y)\leqslant\mathbf{1}_{\widetilde N}(x,y)$ for some neighbourhoods $\widetilde N\supseteq N$ of the diagonal $x=y$ in $\Omega\times\Omega$. The point is then that the local parts of the "new" and "old" operators are sufficiently close to each other, whereas the global parts can be estimated independently.

O

Open problems

The $p^* - 1$ problem

The most important problem in the UMD circle of ideas concerns the norm of a distinguished singular integral operator with respect to the area measure A in the complex plane $\mathbb{C}$, given by

$$Bf(z) = -\frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{\mathbb{C} \setminus D(z,\varepsilon)} \frac{f(y)}{(z-y)^2} \, \mathrm{d}A(y), \quad D(z,\varepsilon) := \{y \in \mathbb{C} : |y - z| < \varepsilon\}.$$

This is sometimes referred to as the two-dimensional Hilbert transform (due to the obvious resemblance), but more commonly as the Beurling, or Ahlfors–Beurling, transform. This operator is fundamental in complex analysis, since it intertwines the two partial differential operators

$$\partial_z := \frac{1}{2}(\partial_x - i\partial_y), \qquad \partial_{\bar{z}} := \frac{1}{2}(\partial_x + i\partial_y)$$

in the sense that

$$B \circ \partial_{\bar{z}} = \partial_z.$$

In the language of Fourier multipliers we have $B = T_m$, where the multiplier $m : \mathbb{C} \setminus \{0\} \to \mathbb{C}$ is given as $m(z) = \overline{z}/z$. From this we infer that $\|B\|_{\mathscr{L}(L^2(\mathbb{C}))} = 1$.

Relatively simple examples discovered by Lehto [1965] show the lower bound

$$\|B\|_{\mathscr{L}(L^p(\mathbb{C}))} \geqslant p^* - 1 := \max(p, p') - 1,$$

but the conjectured converse estimate remains open:

Problem O.1 (Iwaniec [1982]). Show that for all $p \in (1, \infty)$

$$\|B\|_{\mathscr{L}(L^p(\mathbb{C}))} = p^* - 1 := \max(p, p') - 1.$$

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 63, DOI 10.1007/978-3-319-48520-1

The reason for including this problem in the 'UMD circle of ideas' is two-fold. First of all, the conjectured number $p^* - 1$ is precisely the UMD constant of the scalar field, $p^* - 1 = \beta_{p,\mathbb{R}} = \beta_{p,\mathbb{C}}$, although this result of Burkholder [1982] only became available slightly after Iwaniec [1982] formulating his conjecture. (The manuscript of Iwaniec [1982] was received January 15, 1981; that of Burkholder [1982] on June 30, 1982.) Were it only for the equality of two numbers, one might well regard this as a mere coincidence; but, for another thing, the most significant partial progress on Problem O.1 has also been based on Burkholder's results and methods, and a more or less direct identification of the operator B with a martingale transform, in a certain analogy with our estimates for the one-dimensional Hilbert transform H in terms of the UMD constant. We refer the reader to Bañuelos and Janakiraman [2008] for the bound $\|B\|_{\mathscr{L}(L^p(\mathbb{C}))} \leqslant 1.575(p^* - 1)$, and the subsequent work of these authors and their collaborators.

In the hope that the martingale approach is the correct one to this problem, we also formulate a more general version:

Problem O.2. Determine whether

$$\|B\|_{\mathscr{L}(L^p(\mathbb{C};X))} = \beta_{p,X},$$

whenever X is a complex UMD space and $p \in (1, \infty)$.

The estimates

$$\beta_{p,X} \leqslant \|B\|_{\mathscr{L}(L^p(\mathbb{C};X))} \leqslant 2\beta_{p,X}$$

can be found in Geiss, Montgomery-Smith, and Saksman [2010, Theorem 1.1] and follow by considering the real and imaginary parts of B separately.

Interpolation of UMD spaces

It was shown in Proposition 4.2.17(5) that the UMD property is inherited under complex interpolation: if (X_0, X_1) is an interpolation couple of UMD spaces, then the complex interpolation space $[X_0, X_1]_\theta$ is again a UMD space. In particular, $[H, Y]_\theta$ is a UMD space whenever H is a Hilbert space and Y is a UMD space.

Problem O.3 (Rubio de Francia [1986]). Is every UMD space X a complex interpolation space $[H, Y]_\theta$ between a Hilbert space and another UMD space Y, or at least isomorphic to a closed subspace of such a space?

A number of results have been established for 'intermediate UMD spaces' $[H, Y]_\theta$ only, so that an affirmative answer to Problem O.3 would have the important consequence of promoting all such assertions to general theorems about all UMD spaces X. A list of such results includes the following:

- An improved discrete Fourier multiplier theorem on $\ell^p(\mathbb{Z}; X)$ by Berkson and Gillespie [2005] with applications to the functional calculus of power bounded operators.

- Littlewood–Paley-type inequalities by Hytönen [2007], and maximal inequalities and pointwise convergence by Taggart [2009], for the X-valued extensions of abstract diffusion semigroups in the sense of Stein [1970b].
- Pointwise convergence of Fourier series of $f \in L^p(\mathbb{T};X)$ with $p>1$, by Hytönen and Lacey [2013]; this is discussed in more detail further below.

More generally, there are probably several further situations where a known scalar-valued argument could be extended to the intermediate UMD spaces, but much less obviously to general UMD spaces. It is not uncommon that operators T arising from classical Fourier analysis admit an expansion

$$T = \sum_{k=0}^{\infty} T_k, \tag{O.1}$$

where $\|T_k \otimes I_Y\|_{\mathscr{L}(L^p(\mathbb{R}^d;Y))}$ is uniformly bounded (or perhaps moderately increasing) in k, for an arbitrary UMD space Y and $p \in (1,\infty)$, whereas a decay estimate $\|T_k \otimes I_H\|_{\mathscr{L}(L^2(\mathbb{R}^d;H))} \leqslant Ce^{-\varepsilon k}$ may be obtained, say with help of Plancherel's theorem or other orthogonality arguments in $L^2(\mathbb{R}^d;H)$. Interpolation between these estimates would give a favourable bound for $\|T_k \otimes I_X\|_{\mathscr{L}(L^p(\mathbb{R}^d;X))}$, with $X=[H,Y]_\theta$, to sum the series (O.1). For a concrete example of such a situation, see e.g., Watson [1990, Eqs. (13)–(15)].

In the support of an affirmative answer to Problem O.3, there is the following partial result obtained in the same paper where the problem was posed.

Theorem O.1 (Rubio de Francia [1986]). Let X be a UMD Banach function space over a σ-finite measure space $(S,\mathscr{A},\mu)$. Then there exists another UMD Banach function space Y over the same measure space such that $X=[L^2(\mu),Y]_\theta$.

This covers in particular the case of the usual L^p spaces, for which the conclusion that $L^p=[L^2,L^q]_\theta$ for $1/p=\theta/2+(1-\theta)/q$ is already immediate from Theorem 2.2.6. For the Schatten classes $\mathscr{C}^p$, which are not covered by Theorem O.1, we have a similar formula $\mathscr{C}^p=[\mathscr{C}^2,\mathscr{C}^q]_\theta$ from Proposition D.3.1, and $\mathscr{C}^2$ is again a Hilbert space. Actually, by results due to Dodds, Dodds, and De Pagter [1992], the interpolation properties of non-commutative spaces coincide with those of their commutative counterparts under fairly general conditions (see Dodds, Dodds, and De Pagter [1992] for details), and so the Theorem O.1 implies the interpolation property also for many further UMD spaces.

In the support of a negative answer, one may argue that the known examples of UMD spaces are altogether not so many, and a prospective counterexample would presumably have to extend this list in an interesting and non-trivial way.

Pointwise convergence of Fourier series

While the norm convergence $\sum_{k=-n}^{n}\widehat{f}(k)e_k \to f$ in $L^p(\mathbb{T};X)$ is a straightforward consequence of the Hilbert transform boundedness, and hence of the

UMD property, the pointwise convergence of this series (a vector-valued extension of the classical theorem of Carleson [1966]) is a much more delicate issue, and the following question remains open:

Problem O.4 (Rubio de Francia [1986]). Is it true that

$$\lim_{n\to\infty}\sum_{k=-n}^{n}\widehat{f}(k)e_k(x)=f(x)\qquad\text{for a.e. } x\in\mathbb{T},$$

whenever $f\in L^p(\mathbb{T};X)$ for some $p>1$ and X is a UMD space?

Rubio de Francia [1986] gave an affirmative answer for UMD Banach lattices, and Hytönen and Lacey [2013] extended this to all interpolation spaces $X=[H,Y]_\theta$ as considered in Problem O.3. This provides a link between the two problems: An affirmative answer to Problem O.3 would also resolve Problem O.4 in the positive, whereas a counterexample to Problem O.4 would also be a counterexample to Problem O.3.

The closest available result in a general UMD space is the following 'little Carleson theorem' due to Parcet, Soria, and Xu [2013]:

$$\sum_{k=-n}^{n}\widehat{f}(k)e_k(x)=o(\log\log n)\qquad\text{for a.e. } x\in\mathbb{T}.$$

Littlewood–Paley inequality for arbitrary intervals

Yet another problem arising from the work of Rubio de Francia is as follows:

Problem O.5. Prove (or disprove) that the inequality

$$\Big\|\sum_{I\in\tilde{\mathscr{I}}}\varepsilon_I\Delta_I f\Big\|_{L^p(\mathbb{R}\times\Omega;X)}\leqslant C\|f\|_{L^p(\mathbb{R};X)}$$

holds for a finite $C=C_{p,X}$, whenever X is a UMD space of type 2, $p\in[2,\infty)$, $f\in L^p(\mathbb{R};X)$, and $\tilde{\mathscr{I}}$ is an arbitrary disjoint collection of intervals $I\subset\mathbb{R}$. In the negative alternative, give a characterisation of Banach spaces X in which the mentioned estimate is valid.

The positive result would be a vector-valued extension of a well-known theorem of Rubio de Francia [1985] for $X=\mathbb{C}$. See page 481 for a discussion of partial progress on, and motivation for this problem on so-called LPR_p-spaces.

Comparison of the UMD constant and the Hilbert transform norm

A Banach space X is a UMD space if and only if the Hilbert transform H acts boundedly on $L^p(\mathbb{R};X)$ for some/all $p\in(1,\infty)$. The UMD constants of X then relate to the norm $\hbar_{p,X}:=\|H\|_{\mathscr{L}(L^p(\mathbb{R};X))}$ by

$$\beta_{p,X}^{1/2} \leqslant \hbar_{p,X} \leqslant c\beta_{p,X}^2. \tag{O.2}$$

The proof we presented in this Volume gave $c \approx 1.7651\ldots$ (cf. Remark 5.1.14). It is known that the right-hand side inequality actually holds with $c = 1$ and in fact it is even true (as shown in Garling [1986]) that

$$\hbar_{p,X} \leqslant \beta_{p,X}^{-}\beta_{p,X}^{+}.$$

However, a more substantial qualitative improvement would be of interest (see Burkholder [2001, Section 3], Geiss, Montgomery-Smith, and Saksman [2010] and Pietsch and Wenzel [1998, 8.8.2]):

Problem O.6. Is there a linear estimate $\beta_{2,X} \leqslant C\hbar_{2,X}$ or $\hbar_{2,X} \leqslant C\beta_{2,X}$?

A proof of either estimate should presumably reveal some new and more direct connections between the Hilbert transform and martingale transforms, whereas the currently available arguments always create one of the operators from a composition of two copies of the other one.

Aside from intrinsic interest, such a direct connection could shed some light on the following problem of the UMD and Hilbert transform properties of an *operator* $T \in \mathscr{L}(X,Y)$: Consider the estimates

$$\Big\| \sum_{k=1}^n \epsilon_k T(df_k) \Big\|_{L^p(S;Y)} \leqslant \beta_{p,T} \Big\| \sum_{k=1}^n df_k \Big\|_{L^p(S;X)}$$

for all martingale difference sequences $(df_k)_{k=1}^n$ in $L^p(S;X)$ over a σ-finite measure space S, and

$$\|H(Tf)\|_{L^p(\mathbb{R};Y)} \leqslant \hbar_{p,T}\|f\|_{L^p(\mathbb{R};X)},$$

for all functions $f \in L^p(\mathbb{R};X)$. In this language, the UMD property of the space X (as defined before) becomes the UMD property of the operator $T = I_X$, and likewise for the Hilbert transform. Now, an attempt to imitate the proof of (O.2) in the operator case would only give (assuming in addition that $Y = X$)

$$\beta_{p,T^2} \leqslant (\hbar_{p,T})^2, \qquad h_{p,T^2} \leqslant c(\beta_{p,T})^2,$$

i.e., one estimate for T implies the other one for $T^2 = T \circ T$. Accordingly, the following related problem is also open (see Geiss, Montgomery-Smith, and Saksman [2010], Pietsch and Wenzel [1998, 8.8.2]):

Problem O.7. For any $T \in \mathscr{L}(X,Y)$, is it true that $\beta_{2,T} < \infty$ if and only if $\hbar_{2,T} < \infty$, or even that $c\beta_{2,T} \leqslant \hbar_{p,T} \leqslant C\beta_{2,T}$?

As explained in Pietsch and Wenzel [1998, 8.5.23 and 8.8.2] and Wenzel [2004] a possible candidate for a counterexample is the summation operator $\Sigma_n : \ell_n^1 \to \ell_n^\infty$ defined by

$$\Sigma_n(x) = \Big(\sum_{j=1}^k x_j\Big)_{k=1}^n.$$

It is known that $\hbar_{2,\Sigma_{2^n}} \eqsim n$ and $c\sqrt{n} \leqslant \beta_{2,\Sigma_{2^n}} \leqslant Cn$.

Problem O.8 (Wenzel [2004]). Find the asymptotic behaviour of $\beta_{2,\Sigma_{2^n}}$ as $n \to \infty$.

Showing that $\lim_{n\to\infty} \beta_{2,\Sigma_{2^n}}/n = 0$ would give a counterexample to Problem O.7. On the other hand the estimate $\beta_{2,\Sigma_{2^n}} \gtrsim n$ seems to be of interest for the theory of geometry of Banach spaces as well (see Pietsch and Wenzel [1998, 8.5.25]). Some numerical evidence for $\beta_{2,\Sigma_{2^n}} \eqsim \sqrt{n}$ is presented in Wenzel [2004].

Randomised UMD

Proposition 4.2.3 shows that the UMD property can be equivalently phrased as randomised two-sided estimate. It was explained in the Notes of Chapter 5 how this observation leads to the introduction of two new Banach space properties, called UMD$^+$ and UMD$^-$, which express the validity of the lower, respectively upper, randomised bound. An example of a UMD$^-$ space that is not UMD is ℓ^1. No such examples for UMD$^+$ are known.

Problem O.9 (Geiss [1999]). Is it true that UMD$^+$ implies UMD?

For Banach function spaces the answer is affirmative. This can be deduced from unpublished work by Kalton and Weis. Some evidence towards a possible negative solution of the general problem, due to Geiss [1999], has been discussed in the Notes of Chapter 4. In particular there exist UMD$^+$ operators that are not UMD.

The reader may check that our proof of the equivalence of the UMD property with its dyadic counterpart breaks down for the UMD$^\pm$ properties. Therefore we ask:

Problem O.10. Are the UMD$^-$ and UMD$^+$ properties equivalent to their dyadic counterparts?

This problem is less innocent than it may look at first sight, as it is known that, at least for the the UMD$^-$ property, the corresponding constants cannot the same in general; this has been explained in the Notes of Chapter 4.

Decoupling inequalities

Let $(r_k)_{k\geqslant 1}$ and $(\tilde{r_k})_{k\geqslant 1}$ be real Rademacher sequences on probability spaces $(\Omega, \mathbb{P})$ and $(\tilde{\Omega}, \tilde{\mathbb{P}})$ respectively. Specialising the UMD$^-$-property to Paley–Walsh martingales, we obtain the following property of a Banach space X:

A Banach space X is said to have the *decoupling property for Paley–Walsh martingales* if for every $p \in (1,\infty)$ there is a constant C_p such that for all $n \geqslant 1$ and all $\phi_k : \{-1,1\}^{k-1} \to X$ the following estimate holds:

$$\Big\| \sum_{k=1}^n r_k \phi_k(r_1,\dots,r_{k-1}) \Big\|_{L^p(\Omega;X)} \leqslant C_p \Big\| \sum_{k=1}^n \tilde{r}_k \phi_k(r_1,\dots,r_{k-1}) \Big\|_{L^p(\Omega\times\tilde{\Omega};X)}. \tag{O.3}$$

It is known that if this estimate holds for some $p \in (1,\infty)$, then it holds for all $p \in (1,\infty)$ (see Garling [1990]). One can even consider $p \in (0,1]$ for this property. Also L^1-spaces satisfy (O.3). The motivation for studying (O.3) comes from the theory of vector-valued stochastic integration. It is not difficult to see that the theory of Itô stochastic integration extends to the case of processes with values in a martingale type 2 Banach space with one-sided estimates instead of the classical Itô isometry. In the equivalent formulation for 2-smooth Banach spaces, this was pioneered by Neidhardt [1978]. The martingale type 2 theory was developed by Dettweiler [1989, 1991] and Brzeźniak [1995]. An extension of the Itô stochastic integration theory for processes with values in UMD Banach spaces which allows two-sided estimates, was developed in Garling [1986], McConnell [1989] and Van Neerven, Veraar, and Weis [2007]. For a quick introduction to this material we refer the interested reader to the survey paper Van Neerven, Veraar, and Weis [2015].

Based on (O.3) it is possible to develop a more general theory which allows one-sided estimates as well. This leads to the following natural question:

Problem O.11. Does martingale type 2 imply the decoupling property for Paley–Walsh martingales? Or more generally does finite (martingale) cotype of a Banach space imply the decoupling property for Paley–Walsh martingales?

Positive or negative answers to these questions on Banach lattices are already of interest. More generally one might ask whether martingale type 2 implies UMD^-. It follows from Garling [1990] that for every $p \in [1,2)$ there is an example of a Banach lattice which has martingale type p but does not satisfy UMD^-.

As an alternative generalisation of Problem O.11 one might ask whether martingale type 2 implies the decoupling property as introduced in the Notes of Chapter 4.

Motivated by a result in the scalar case by Hitczenko [1993] we also have the following question on the behaviour of the constant in (O.3) if one varies p. Let $\beta^{-,\Delta}_{p,X}$ denote the optimal constant C in (O.3).

Problem O.12. Let X be a Banach space for which the decoupling property holds. Does one have $\sup_{p\in(1,\infty)} \beta^{-,\Delta}_{p,X} < \infty$?

More general decoupling inequalities with a uniform estimate in p have been proved in Hitczenko [1994] in the scalar case and extended to the Hilbert space case in Cox and Veraar [2011]. A positive answer to Problem O.12 yields an L^p-estimate for stochastic integrals with the correct dependence on the exponent as $p \to \infty$.

Rademacher maximal inequality

The Rademacher maximal function M_{Rad} was introduced in Definition 3.6.8 mainly to illustrate some of the developed techniques in Chapter 3. It has proved to be quite a useful tool in analysis in Banach spaces (see Di Plinio and Ou [2015] and Hytönen, McIntosh, and Portal [2008]). Recall that a Banach space is said to have the RMF property if M_{Rad} is bounded from $L^p(S;X)$ into $L^p(S)$. It is known that every space with type 2 has the RMF property. Also every UMD lattice has the RMF property. Conversely, every RMF space has non-trivial type. It would be interesting to know whether the converse is also true. A natural question is the following:

Problem O.13. Does every UMD space have the RMF property? Does every space with non-trivial type have the RMF property?

The RMF property does not imply UMD, because by James [1978] there exists a non-reflexive (thus non-UMD) space with type 2.

A

Measure theory

It is assumed that the reader is familiar with the elements of measure theory and Lebesgue integration. The purpose of this appendix is to bring together a rather diverse collection of results that are useful in the main text.

A.1 Measure spaces

In order to fix the terminology, we begin by recalling some standard definitions.

A.1.a Basic definitions

A *measurable space* is a pair $(S, \mathscr{A})$, where S is a set and $\mathscr{A}$ is a *σ-algebra* in S, i.e., a non-empty collection of subsets of S which is closed under taking complements and countable unions.

When $(S, \mathscr{A})$ and $(T, \mathscr{B})$ are measurable spaces, a mapping $f : S \to T$ is said to be *$\mathscr{A}/\mathscr{B}$-measurable* (or just *measurable* if it is clear to which σ-algebras is being referred) if

$$f^{-1}(B) \in \mathscr{A} \quad \text{for all } B \in \mathscr{B}.$$

In the vector-valued context, the range space of the functions under consideration is often a topological space. The *Borel σ-algebra* of a topological space T is the smallest σ-algebra containing all open subsets of T. This σ-algebra is denoted by $\mathscr{B}(T)$. The sets in $\mathscr{B}(T)$ are called the *Borel sets* of T. When considering functions f from a measurable space $(S, \mathscr{A})$ into a topological space T, we always think of T as being endowed with its Borel σ-algebra and call such a function *$\mathscr{A}$-measurable* if $f^{-1}(B) \in \mathscr{A}$ for all $B \in \mathscr{B}(T)$, i.e., if it is $\mathscr{A}/\mathscr{B}(T)$-measurable. The collection of all $B \in \mathscr{B}(T)$ satisfying $f^{-1}(B) \in \mathscr{A}$ is easily seen to be a σ-algebra. As a consequence, f is $\mathscr{A}$-measurable if and only if $f^{-1}(U) \in \mathscr{A}$ for all open sets U in T. A function f from a topological

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik
und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys
in Mathematics 63, DOI 10.1007/978-3-319-48520-1

space S into a measurable space $(T, \mathscr{B})$ is called *Borel measurable* (or simply, a *Borel function* if g is $\mathscr{B}(S)/\mathscr{B}$-measurable. When both S and T are topological spaces, every continuous function $g : S \to T$ is Borel measurable.

A *measure* on a measurable space $(S, \mathscr{A})$ is a mapping $\mu : \mathscr{A} \to [0, \infty]$ with the property that $\mu(\varnothing) = 0$ and

$$\mu\Big(\bigcup_{n\geqslant 1} A_n\Big) = \sum_{n\geqslant 1} \mu(A_n), \tag{A.1}$$

whenever the sets $A_1, A_2, \dots$ in $\mathscr{A}$ are disjoint; we refer to the condition (A.1) as the *countable additivity* of μ. A *measure space* is a triple $(S, \mathscr{A}, \mu)$, where $(S, \mathscr{A})$ is a measurable space and μ is a measure on $(S, \mathscr{A})$. A measure μ is called a *probability measure* if $\mu(S) = 1$, a *finite measure* if $\mu(S) < \infty$, and a *σ-finite measure* if there exist sets $S^{(1)} \subseteq S^{(2)} \subseteq \dots$ in $\mathscr{A}$ such that $\mu(S^{(n)}) < \infty$ for all $n \geqslant 1$ and $S = \bigcup_{n\geqslant 1} S^{(n)}$. Such a sequence is called an *exhausting sequence* for μ. A measure space $(S, \mathscr{A}, \mu)$ is said to be a *probability space*, respectively a *finite measure space* or a *σ-finite measure space*, when μ has the corresponding property.

A set $N \in \mathscr{A}$ is said to be a *μ-null set* if $\mu(N) = 0$. If f is a function defined on S, a statement about the values f is said to hold *μ-almost everywhere* if there is a null set $N \in \mathscr{A}$ such that the statement holds for all values $f(s)$ with $s \in \complement N$, the complement of N.

Remark A.1.1. We emphasise that it is part of the definition of a μ-null set that it belongs to $\mathscr{A}$. This remark is relevant in the light of the fact many textbooks define a μ-null set to be any set that can be covered, for any given $\varepsilon > 0$, with countably many sets in $\mathscr{A}$ whose joint μ-measure does not exceed ε. Such a set need not be itself in $\mathscr{A}$, and therefore this definition is more general than ours. However, such a set is always contained in a μ-null set belonging to $\mathscr{A}$, and this is usually sufficient for applications.

It is sometimes desirable to replace the closedness under countable unions in the definition of a σ-algebra by closedness under finite unions; the resulting structure is called an *algebra*. An algebra $\mathscr{B}$ is said to *generate* the σ-algebra $\mathscr{A}$ if $\mathscr{A}$ is the smallest σ-algebra containing $\mathscr{B}$. The following approximation lemma is frequently useful.

Lemma A.1.2. *Let $(S, \mathscr{A}, \mu)$ be a finite measure space and let $\mathscr{B}$ be an algebra in S which generates $\mathscr{A}$. Then for all $A \in \mathscr{A}$ and $\varepsilon > 0$ there exists a set $B \in \mathscr{B}$ such that $\mu(A \Delta B) < \varepsilon$.*

Here,

$$A \Delta B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A)$$

is the *symmetric difference* of A and B.

Proof. Let $\mathscr{C}$ be the collection of all sets $A \in \mathscr{A}$ with the property that for all $\varepsilon > 0$ there exists a set $B \in \mathscr{B}$ such that $\mu(A \Delta B) < \varepsilon$. It is easily checked that $\mathscr{C}$ is a sub-σ-algebra of $\mathscr{A}$. Clearly, this σ-algebra contains $\mathscr{B}$, and therefore it contains $\mathscr{A}$. As a consequence we see that $\mathscr{C} = \mathscr{A}$. $\square$

This lemma fails for σ-finite measures, as can be seen from the example where $\mathscr{B}$ is the algebra generated by the intervals $[k, k+1)$, $k \in \mathbb{Z}$, and $\mathscr{A}$ is the σ-algebra generated by $\mathscr{B}$.

A *complex measure* on a measurable space $(S, \mathscr{A})$ is a mapping $\mu : \mathscr{A} \to \mathbb{C}$ satisfying $\mu(\varnothing) = 0$ and $\mu(\bigcup_{n\geqslant 1} A_n) = \sum_{n\geqslant 1} \mu(A_n)$ whenever the sets $A_1, A_2, \dots$ in $\mathscr{A}$ are disjoint. A *real measure* is a complex measure which takes real values only. Note that in the definition of a measure we allow sets to have infinite measure, whereas in the definition of a real measure this is not permitted. Thus a measure is a real measure if and only if it is finite.

The next uniqueness result, known as *Dynkin's lemma*, gives a convenient criterion for deciding when two complex measures agree.

Lemma A.1.3 (Dynkin). *Let μ_1 and μ_2 be complex measures defined on a measurable space $(S, \mathscr{A})$. Let $\mathscr{C} \subseteq \mathscr{A}$ be a collection of sets with the following properties:*

(i) *$S \in \mathscr{C}$;*
(ii) *$\mathscr{C}$ is closed under finite intersections;*
(iii) *$\sigma(\mathscr{C})$, the σ-algebra generated by $\mathscr{C}$, equals $\mathscr{A}$.*

If $\mu_1(C) = \mu_2(C)$ for all $C \in \mathscr{C}$, then $\mu_1 = \mu_2$.

Proof. Let $\mathscr{D}$ denote the collection of all sets $D \in \mathscr{A}$ with $\mu_1(D) = \mu_2(D)$. Then $\mathscr{C} \subseteq \mathscr{D}$ and $\mathscr{D}$ is a *Dynkin system*, that is,

- $S \in \mathscr{D}$;
- if $D_1 \subseteq D_2$ with $D_1, D_2 \in \mathscr{D}$, then also $D_2 \setminus D_1 \in \mathscr{D}$;
- if $D_1 \subseteq D_2 \subseteq \dots$ with all $D_n \in \mathscr{D}$, then also $\bigcup_{n\geqslant 1} D_n \in \mathscr{D}$.

By assumption we have $\mathscr{D} \subseteq \mathscr{A} = \sigma(\mathscr{C})$; we will show that $\sigma(\mathscr{C}) \subseteq \mathscr{D}$. To this end let $\mathscr{D}_0$ denote the smallest Dynkin system in $\mathscr{A}$ containing $\mathscr{C}$. We will show that $\sigma(\mathscr{C}) \subseteq \mathscr{D}_0$. In view of $\mathscr{D}_0 \subseteq \mathscr{D}$, this will prove the lemma.

Let $\mathscr{E} = \{D_0 \in \mathscr{D}_0 : \ D_0 \cap C \in \mathscr{D}_0 \text{ for all } C \in \mathscr{C}\}$. Then $\mathscr{E}$ is a Dynkin system and $\mathscr{C} \subseteq \mathscr{E}$ since $\mathscr{C}$ is closed under taking finite intersections. It follows that $\mathscr{D}_0 \subseteq \mathscr{E}$, since $\mathscr{D}_0$ is the smallest Dynkin system containing $\mathscr{C}$. But obviously, $\mathscr{E} \subseteq \mathscr{D}_0$, and therefore $\mathscr{E} = \mathscr{D}_0$.

Now let $\mathscr{E}' = \{D_0 \in \mathscr{D}_0 : \ D_0 \cap D \in \mathscr{D}_0 \text{ for all } D \in \mathscr{D}_0\}$. Then $\mathscr{E}'$ is a Dynkin system and the fact that $\mathscr{E} = \mathscr{D}_0$ implies that $\mathscr{C} \subseteq \mathscr{E}'$. Hence $\mathscr{D}_0 \subseteq \mathscr{E}'$, since $\mathscr{D}_0$ is the smallest Dynkin system containing $\mathscr{C}$. But obviously, $\mathscr{E}' \subseteq \mathscr{D}_0$, and therefore $\mathscr{E}' = \mathscr{D}_0$.

It follows that $\mathscr{D}_0$ is closed under taking finite intersections. But a Dynkin system with this property is a σ-algebra. Thus, $\mathscr{D}_0$ is a σ-algebra, and now $\mathscr{C} \subseteq \mathscr{D}_0$ implies that also $\sigma(\mathscr{C}) \subseteq \mathscr{D}_0$. $\square$

A.1.b The structure of sub-σ-algebras

The following structure result will be important in Chapter 3. Let us call a measure *purely infinite* if it takes values in $\{0, \infty\}$.

If $\mathscr{F}$ is a sub-σ-algebra of $\mathscr{A}$, then for a set $F \in \mathscr{F}$ we denote by $\mathscr{F}|_F$ the σ-algebra in F defined by

$$\mathscr{F}|_F = \{F \cap F' : \ F' \in \mathscr{F}\} = \{F' \in \mathscr{F} : \ F' \subseteq F\}.$$

Proposition A.1.4. *Let $(S, \mathscr{A}, \mu)$ be a σ-finite measure space and $\mathscr{F} \subseteq \mathscr{A}$ be a sub-σ-algebra. Then there exist disjoint sets $S_0, S_1 \in \mathscr{F}$ such that $S_0 \cup S_1 = S$ such that the restriction of μ to $\mathscr{F}|_{S_0}$ is σ-finite and the restriction of μ to $\mathscr{F}|_{S_1}$ is purely infinite.*

The proof is based on two lemmas. Denote $\mathscr{A}^+ := \{A \in \mathscr{A} : 0 < \mu(A) < \infty\}$.

Lemma A.1.5. *There exists $\phi \in L^1(S, \mathscr{A}, \mu)$ such that $\phi > 0$ everywhere and $\int_S \phi \,\mathrm{d}\mu = 1$.*

Proof. Let $A_i \in \mathscr{A}^+$, $i = 1, 2, \ldots$, be a disjoint cover of S (which exists by σ-finiteness), and set $\phi := \sum_{i \geqslant 1} 2^{-i} \mathbf{1}_{A_i} / \mu(A_i)$. $\square$

Lemma A.1.6. *Any disjoint sub-collection of $\mathscr{A}^+$ is at most countable.*

Proof. Let $\mathscr{D}$ be such a sub-collection, and let $\mathscr{F} \subseteq \mathscr{D}$ be finite. Then

$$1 = \int_S \phi \,\mathrm{d}\mu \geqslant \sum_{F \in \mathscr{F}} \int_F \phi \,\mathrm{d}\mu \geqslant \frac{1}{k} \cdot \#\Big\{F \in \mathscr{F} : \int_F \phi \,\mathrm{d}\mu > \frac{1}{k}\Big\}.$$

Thus

$$\#\Big\{F \in \mathscr{F} : \int_F \phi \,\mathrm{d}\mu > \frac{1}{k}\Big\} \leqslant k$$

for any finite $\mathscr{F} \subseteq \mathscr{D}$, and hence

$$\#\Big\{F \in \mathscr{D} : \int_F \phi \,\mathrm{d}\mu > \frac{1}{k}\Big\} \leqslant k.$$

Since $\mathscr{D} \subseteq \mathscr{A}^+$ (in particular, $\mu(A) > 0$ for every $A \in \mathscr{D}$), we have

$$\begin{aligned} \mathscr{D} &= \{A \in \mathscr{D} : \mu(A) > 0\} \\ &= \Big\{A \in \mathscr{D} : \int_A \phi \,\mathrm{d}\mu > 0\Big\} = \bigcup_{k=1}^{\infty} \Big\{A \in \mathscr{D} : \int_A \phi \,\mathrm{d}\mu > \frac{1}{k}\Big\}, \end{aligned}$$

which is at most countable. $\square$

Proof of Proposition A.1.4. Consider the family $\mathfrak{D}$ of all disjoint sub-collections of $\mathscr{F}^+$. Then $\mathfrak{D}$ is partially ordered by set inclusion, and we can argue by Zorn's lemma that there exists a maximal member $\mathscr{D} \in \mathfrak{D}$. By Lemma A.1.6, $\mathscr{D} \subseteq \mathscr{A}^+$ is at most countable. Let us fix an enumeration $\mathscr{D} = \{D_j\}_{j\geqslant 1}$ and put

$$S_0 := \bigcup_{j\geqslant 1} D_j.$$

This set is in $\mathscr{F}$. We will check that it has the required properties. First, the above decomposition of S_0 into $D_j \in \mathscr{F}^+$ immediately shows the σ-finiteness of μ on S_0.

Second, let $S_1 := \complement S_0$ and let $F \subseteq S_1$ be a set in $\mathscr{F}$. We cannot have $0 < \mu(F) < \infty$, as this would contradict the maximality of $\{D_j\}_{j\geqslant 1}$, since $\{D_j\}_{j\geqslant 1} \cup \{F\}$ would then be larger disjoint sub-collection of $\mathscr{F}^+$. This only leaves the possibility that $\mu(F) \in \{0, \infty\}$, and we are done. □

The σ-finiteness assumption on $(S, \mathscr{A}, \mu)$ cannot be dropped from Proposition A.1.4, as is shown by the example of the unit interval with the counting measure on its power set.

A.1.c Divisibility

Let $(S, \mathscr{A}, \mu)$ be a measure space. A set $A \in \mathscr{A}$ is called an *atom* if $\mu(A) > 0$ and $A = A_0 \cup A_1$ with disjoint $A_0, A_1 \in \mathscr{A}$ implies that $\mu(A_0) = 0$ or $\mu(A_1) = 0$. The measure space $(S, \mathscr{A}, \mu)$ is called *non-atomic* if $\mathscr{A}$ has no atoms, and *atomic* if $\mathscr{A}$ is generated by its atoms.

Proposition A.1.7. *Let $(S, \mathscr{A}, \mu)$ be a σ-finite atomic measure space. Then there is a μ-null set $N \in \mathscr{A}$ such that its complement $\complement N$ is the union of at most countably many disjoint atoms of finite μ-measure.*

Proof. We first claim that $\mathscr{A}$ contains at most countably many atoms. Indeed, suppose we had atoms B_i, $i \in I$, with I an uncountable index set and $B_i \neq B_j$ when $i \neq j$, and let $(S^{(n)})_{n\geqslant 1}$ be an exhaustion of S by sets of finite μ-measure. For all pairs (n, i) we have either $\mu(S^{(n)} \cap B_i) = 0$ or $\mu(S^{(n)} \cap B_i) = \mu(B_i)$. Hence then there must be an integer $k \geqslant 1$ such that at least one of the $S^{(n)}$ intersects infinitely many (even uncountably many) atoms B_i in a set of measure $\geqslant \frac{1}{k}$. But since any two atoms are either equal, disjoint, or intersect in a set of measure 0, this is impossible since $\mu(S^{(n)}) < \infty$.

Let $(B_i)_{i\in I}$, with I countable, the set of all atoms in $\mathscr{A}$, and consider the set $N_0 := S \setminus \bigcup_{i\in I} B_i$. We wish to prove that $\mu(N_0) = 0$.

Suppose the contrary. Then $\mu(N_0) > 0$. The set N_0 contains no atoms, for if it did these atoms would be disjoint with all the B_i, contradicting the assumption that I label all atoms. In particular N_0 itself is not an atom, and therefore it admits a decomposition $N_0 = N_0' \cup N_0''$ with $N_0', N_0'' \in \mathscr{A}$ and $\mu(N_0') > 0$ and $\mu(N_0'') > 0$.

For $j,k \in I$ with $j \neq k$ let $N_{jk} := B_j \cap B_k$, and let $N' := \bigcup_{\substack{j,k\in I \\ j\neq k}} N_{jk}$. Then $\mu(N') = 0$ and the sets $A_i := B_i \setminus N'$ are disjoint atoms. Let $\mathscr{B}$ be the collection of all sets that can be represented as a countable union of the atoms A_i, sets in $\mathscr{A}|_{N'}$, and N_0. This is a sub-σ-algebra of $\mathscr{A}$ containing every B_i. Since $\mathscr{A}$ is generated by its atoms, it follows that $\mathscr{A} = \mathscr{B}$. In particular $N_0', N_0'' \in \mathscr{B}$. But both these sets contain no atoms and are strictly contained in N_0, and therefore both are countable unions of sets in $\mathscr{A}|_{N'}$. But then $\mu(N_0') = \mu(N_0'') = 0$.

This contradiction concludes the proof that $\mu(N_0) = 0$. Since also $\mu(N') = 0$, the lemma follows (with $N = N_0 \cup N'$). $\square$

A measure space $(S, \mathscr{A}, \mu)$ is called *divisible* if for all real numbers $0 < t < 1$ and all sets $A \in \mathscr{A}$ there exist two disjoint sets $A_0, A_1 \in \mathscr{A}$ satisfying $A_0 \subseteq A$, $A_1 \subseteq A$, and

$$\mu(A_0) = (1-t)\mu(A), \quad \mu(A_1) = t\mu(A).$$

Divisible measure spaces play a role in Sections 1.3 and 3.6. In the former, the fact is mentioned that a σ-finite measure space $(S, \mathscr{A}, \mu)$ is divisible if and only if it is non-atomic. For reasons of completeness we include a proof of this fact.

Proposition A.1.8. *A σ-finite measure space $(S, \mathscr{A}, \mu)$ is divisible if and only if it is non-atomic.*

Proof. It is obvious that divisibility implies non-atomicity. Hence only the 'if' part needs proof, and we assume from now on that $(S, \mathscr{A}, \mu)$ is non-atomic. We also understand that all sets are taken from $\mathscr{A}$ without specifying this explicitly every time.

By exhaustion, it suffices to prove the divisibility property for sets of *finite* measure.

Step 1 – We claim that for every $A \in \mathscr{A}$ of finite measure, there exists $B \subseteq A$ and $t \in (\frac{1}{4}, \frac{1}{2}]$ such that $\mu(B) = t\mu(A)$.

Let us denote

$$t_A := \sup \big\{ t \in [0, \frac{1}{2}] : \ \exists\, B \subseteq A, \ \mu(B) = t\mu(A) \big\}.$$

As there are no atoms, we must have $t_A > 0$ and the claim is that $t_A > \frac{1}{4}$. We assume the opposite, i.e., that $t_A \leqslant \frac{1}{4}$.

We first show that there cannot be a set $B \subseteq A$ for which the equality $\mu(B) = t_A\mu(A)$ is reached. Indeed, assuming that B was such a set, we construct another set $C \subseteq A$ with $\mu(C)/\mu(A) \in (t_A, \frac{1}{2}]$, contradicting the definition of t_A. To this end, let $D \subseteq A \setminus B$ be a set with $0 < \mu(D) < \mu(A \setminus B)$. Write $\mu(D) = t\mu(A)$ so that $t \in (0, 1 - t_A)$. Depending on whether $t \in (0, t_A]$, $t \in (t_A, \frac{1}{2}]$ or $t \in (\frac{1}{2}, 1 - t_A)$, the contradiction is reached with the set $C = B \cup D$, $C = D$ or $C = A \setminus D$.

It remains to construct a set verifying the forbidden equality to get a contradiction. Pick a sequence of subsets $A_i \subseteq A$ so that $\mu(A_i) = (t_A - \varepsilon_i)\mu(A)$, where the numbers ε_i decrease rapidly to zero. Then $\mu(A_i \cup A_j) \leqslant \mu(A_i) + \mu(A_j) < 2t_A\mu(A) \leqslant \frac{1}{2}\mu(A)$ so, by the definition of t_A,

$$t_A\mu(A) \geqslant \mu(A_i \cup A_j) = \mu(A_i) + \mu(A_j \setminus A_i) = (t_A - \varepsilon_i)\mu(A) + \mu(A_j \setminus A_i).$$

Thus $\mu(A_j \setminus A_i) \leqslant \varepsilon_i\mu(A)$, and hence

$$\mu\Big(\bigcup_{i=1}^{n} A_i\Big) \leqslant \mu(A_1) + \sum_{i=1}^{n-1} \mu(A_{i+1} \setminus A_i) \leqslant (t_A - \varepsilon_1)\mu(A) + \sum_{i=1}^{n-1} \varepsilon_i\mu(A) < \frac{1}{2}\mu(A),$$

when the ε_i decrease sufficiently rapidly. Again by the definition of t_A,

$$(t_A - \varepsilon_n)\mu(A) = \mu(A_n) \leqslant \mu\Big(\bigcup_{i=1}^{n} A_i\Big) \leqslant t_A\mu(A),$$

and this implies that $\mu\big(\bigcup_{i=1}^{\infty} A_i\big) = t_A\mu(A)$. But this is impossible, and the proof of the claim is complete.

Step 2 – Next we claim that if $A_1 \subseteq A$ with $0 < \mu(A_1) \leqslant \frac{1}{2}\mu(A) < \infty$, there exists another set A_2 with $A_1 \subseteq A_2 \subseteq A$ such that

$$0 \leqslant \frac{1}{2} - t_2 \leqslant \frac{3}{4}(\frac{1}{2} - t_1), \qquad t_i := \frac{\mu(A_i)}{\mu(A)}.$$

Let $t_1 := \mu(A_1)/\mu(A) \in (0, \frac{1}{2}]$. If $t_1 = \frac{1}{2}$, we can simply take $A_2 := A_1$. Assume therefore that $t_1 \in (0, \frac{1}{2})$.

Let $B_0 := A \setminus A_1$. Using Step 1 we then choose sets B_i, $i = 1, 2, \ldots$, by the following procedure. Given B_{i-1}, choose $B_i \subseteq B_{i-1}$ with $\mu(B_i) = s_i\mu(B_{i-1})$ for some $s_i \in (\frac{1}{4}, \frac{1}{2}]$. Stop at the first B_n which satisfies $\mu(B_n) < (\frac{1}{2} - t_1)\mu(A)$. Note that $\mu(B_0) = (1 - t_1)\mu(A) > (\frac{1}{2} - t_1)\mu(A)$, so the process cannot stop at B_0, and $\mu(B_i) \leqslant \frac{1}{2}\mu(B_{i-1}) \leqslant \ldots \leqslant 2^{-i}\mu(B_0)$, so the process will eventually stop. If it stops at B_n, then

$$(\frac{1}{2} - t_1)\mu(A) > \mu(B_n) > \frac{1}{4}\mu(B_{n-1}) \geqslant \frac{1}{4}(\frac{1}{2} - t_1)\mu(A).$$

We define $A_2 := A_1 \cup B_n$. Then $\mu(A_2) = \mu(A_1) + \mu(B_n)$, so that

$$\frac{1}{2}\mu(A) > \mu(A_2) > t_1\mu(A) + \frac{1}{4}(\frac{1}{2} - t_1)\mu(A)$$

and, writing $\mu(A_2) = t_2\mu(A)$, we have

$$0 < \frac{1}{2} - t_2 < \frac{1}{2} - t_1 - \frac{1}{4}(\frac{1}{2} - t_1) = \frac{3}{4}(\frac{1}{2} - t_1),$$

which proves the claim.

Step 3 – For every A of finite measure, there exists $B \subseteq A$ such that $\mu(B) = \frac{1}{2}\mu(A)$.

We construct a sequence of sets $A_n \subseteq A$ by induction. For A_1, we simply pick any $A_1 \subseteq A$ such that $\mu(A_1) = t_1\mu(A)$ and $t_1 \in (0, \frac{1}{2}]$.

We iterate the procedure of Step 2: If A_{n-1} is already chosen, we find A_n so that $A_{n-1} \subseteq A_n \subseteq A$, and $t_n := \mu(A_n)/\mu(A)$ satisfies

$$0 \leqslant \frac{1}{2} - t_n \leqslant \frac{3}{4}(\frac{1}{2} - t_{n-1}) \leqslant \ldots \leqslant (\frac{3}{4})^{n-1}(\frac{1}{2} - t_1).$$

Thus $t_n \uparrow \frac{1}{2}$, and hence $B = \bigcup_{n=1}^{\infty} A_n$ verifies the assertion.

Step 4 – Assume for induction that we have found disjoint sets $A_j \subseteq A$, $j = 1, \ldots, n$, so that $\mu(A_j) = 2^{-j}\mu(A)$. (This is trivially true for $n = 0$.) Then, using Step 3, we pick $A_{n+1} \subseteq A \setminus \bigcup_{j=1}^{n} A_j$ with

$$\mu(A_{n+1}) = \frac{1}{2}\mu\Big(A \setminus \bigcup_{j=1}^{n} A_j\Big) = \frac{1}{2}\Big(1 - \sum_{j=1}^{n} 2^{-j}\Big)\mu(A) = 2^{-(n+1)}\mu(A).$$

Let $t \in (0,1)$ have the binary expansion $t = \sum_{j=1}^{\infty} \delta_j 2^{-j}$ with $\delta_j \in \{0,1\}$. For any set C, let us write $C^0 := \varnothing$ and $C^1 := C$. Then

$$B := \bigcup_{j=1}^{\infty} A_j^{\delta_j} \subseteq A$$

satisfies

$$\mu(B) = \sum_{j=1}^{\infty} \mu(A_j^{\delta_j}) = \sum_{j=1}^{\infty} \delta_j 2^{-j}\mu(A) = t\mu(A).$$

□

Corollary A.1.9. *Let $(S, \mathscr{A}, \mu)$ be a σ-finite measure space and $\mathscr{F}$ be a sub-σ-algebra of $\mathscr{A}$, such that $(S, \mathscr{F}, \mu|_{\mathscr{F}})$ is σ-finite and divisible. Then $(S, \mathscr{A}, \mu)$ is divisible as well.*

Proof. Let $A \in \mathscr{A}$ have strictly positive measure. By the σ-finiteness and divisibility of $\mathscr{F}$, we can partition S into countably many subsets $F_i \in \mathscr{F}$, each of positive measure $\mu(F_i) < \mu(A)$. Since $0 < \mu(A) = \sum_{i=1}^{\infty} \mu(A \cap F_i)$, there is at least one i with

$$0 < \mu(A \cap F_i) \leqslant \mu(F_i) < \mu(A),$$

so that $A \cap F_i$ is a non-trivial subset of A. Since this is true for any $A \in \mathscr{A}$ of positive measure, we find that $\mathscr{A}$ is non-atomic, and hence divisible. □

We conclude our discussion of divisibility with a characterisation of divisible probability spaces in terms of independence. In probability theory, measurable scalar-valued functions defined on a probability space are called *random variables*, and we shall adopt this terminology here.

Proposition A.1.10. *A probability space $(\Omega, \mathscr{A}, \mathbb{P})$ is divisible if and only if, for any finite σ-algebra $\mathscr{E}$ contained in $\mathscr{A}$, there is a uniformly distributed random variable $u : \Omega \to [0,1)$ independent of $\mathscr{E}$.*

Proof. 'If': Given a set $E \in \mathscr{E}$, let $u : \Omega \to [0,1)$ be a uniformly distributed random variable (i.e., for all Borel sets B in $[0,1)$ we have $\mathbb{P}(\{u \in B\}) = |B|$, the Lebesgue measure of B) which is independent of $\sigma(E)$. Then $\mu(E \cap \{u \leqslant t\}) = t\mu(E)$ for each $t \in [0,1)$.

'Only if': Suppose that Ω is divisible, let $\mathscr{E}$ be a finite σ-algebra of its measurable subsets, and $\mathscr{E}^*$ the collection of its atoms, which is a finite partition of Ω. For each $A \in \mathscr{E}^*$, we use divisibility to divide $A_{[0,1)} := A$ into two halves $A_{[0,1/2)}$ and $A_{[1/2,1)}$ of equal measure. By induction, we obtain subsets A_I of A indexed by the dyadic sub-intervals

$$I \in \mathscr{D}[0,1) := \bigcup_{k=0}^{\infty} \mathscr{D}_k[0,1), \quad \mathscr{D}_k[0,1) := \{2^{-k}[j-1,j) : j = 1,2,\dots,2^k\},$$

such that $A_I \subseteq A_J$ if and only if $I \subseteq J$, and $\mu(A_I) = |I|\mu(A)$.

We then define

$$u := \sum_{A \in \mathscr{E}^*} u_A, \qquad u_A := \sum_{k=1}^{\infty} 2^{-k} \sum_{I \in \mathscr{D}_{k-1}[0,1)} \mathbf{1}_{A_{I_{\text{right}}}},$$

where $I_{\text{right}} := [(a_I + b_I)/2, b_I)$ is the right half of the interval $I = [a_I, b_I) \in \mathscr{D}[0,1)$. To complete the proof, we need to check that

(a) u is uniformly distributed on $[0,1)$,
(b) u is independent of $\mathscr{E}$.

Let us first consider the functions

$$i(t) := \sum_{k=1}^{\infty} 2^{-k} i_k(t), \quad i_k(t) := \sum_{I \in \mathscr{D}_{k-1}[0,1)} \mathbf{1}_{I_{\text{right}}}(t),$$

defined for $t \in [0,1)$. One can check that $i_k(t)$ is precisely the kth digit $t_k \in \{0,1\}$ in the binary expansion $t = \sum_{k=1}^{\infty} 2^{-k} t_k$ (chosen to terminate with a tail of zeros instead of a tail of ones in case of ambiguity), and therefore $i(t) = t$. It is also clear that

$$\{i > t\} = (t,1) = \bigcup_{I \in \mathscr{D}[t]} I = \bigcup_{I \in \mathscr{D}[t]^\star} I,$$

where $\mathscr{D}[t] := \{I = [a_I, b_I) \in \mathscr{D}[0,1) : a_I > t\}$, and $\mathscr{D}[t]^\star$ is the collection of maximal (with respect to set inclusion) members of $\mathscr{D}[t]$.

Due to the bijective and isomorphic (relative to set inclusions) correspondence of the intervals $I \in \mathscr{D}[0,1)$ and the sets A_I, and the pairwise disjointness of the sets $A \in \mathscr{E}^*$, we infer that

$$\{u_A > t\} = \bigcup_{I \in \mathscr{D}[t]^\star} A_I, \qquad \{u > t\} = \bigcup_{A \in \mathscr{E}^*} \bigcup_{I \in \mathscr{D}[t]^\star} A_I.$$

Thus, if $E \in \mathscr{E}$ is arbitrary, say $E = \bigcup_{A \in \mathscr{E}'} A$ for some subset $\mathscr{E}' \subseteq \mathscr{E}^*$, we have

$$\mu(\{u > t\} \cap E) = \sum_{A \in \mathscr{E}'} \sum_{I \in \mathscr{D}[t]^\star} \mu(A_I) = \sum_{A \in \mathscr{E}'} \mu(A) \sum_{I \in \mathscr{D}[t]^\star} |I| = \mu(E)|(t,1)|.$$

With $E = \Omega$ this proves (a), and then with arbitrary $E \in \mathscr{E}$ it proves (b). $\square$

A.2 Convergence in measure

In this section we introduce some convergence notions which are usually stated for random variables defined on probability spaces only, but which can be extended to the σ-finite setting without much difficulty.

Definition A.2.1. *Let $(f_n)_{n\geqslant 1}$ be a sequence of scalar-valued measurable functions defined on a measure space S.*

(1) *The sequence $(f_n)_{n\geqslant 1}$* converges μ-almost surely *to f if there is a μ-null set $N \in \mathscr{A}$ such that $f_n \to f$ pointwise on $S \setminus N$.*
(2) *The sequence $(f_n)_{n\geqslant 1}$* converges in μ-measure *to f if for all $r > 0$ and all $A \in \mathscr{A}$ satisfying $\mu(A) < \infty$ we have*

$$\lim_{n\to\infty} \mu(A \cap \{|f_n - f| > r\}) = 0.$$

(3) *The sequence $(f_n)_{n\geqslant 1}$ is* Cauchy in μ-measure *if for all $r > 0$ and all $A \in \mathscr{A}$ satisfying $\mu(A) < \infty$ we have*

$$\lim_{n,m\to\infty} \mu(A \cap \{|f_n - f_m| > r\}) = 0.$$

When the measure μ is understood we will omit the prefix 'μ-'.

Lemma A.2.2. *Suppose that $(S, \mathscr{A}, \mu)$ is σ-finite and let $(f_n)_{n\geqslant 1}$ be a sequence of scalar-valued measurable functions on S. If $f_n \to f$ almost surely, then $f_n \to f$ in measure.*

Proof. Let $r > 0$ and let $A \in \mathscr{A}$ be a set of finite μ-measure. Then, by the dominated convergence theorem, $\mu(A \cap |f_n - f| > r) \to 0$ as $n \to \infty$. $\square$

In the conversely direction we have the following result.

Lemma A.2.3. *Suppose that $(S, \mathscr{A}, \mu)$ is σ-finite and let $(f_n)_{n\geqslant 1}$ be a sequence of scalar-valued measurable functions on S which is Cauchy in measure. Then there exists a measurable function f on S and a subsequence $(f_{n_k})_{k\geqslant 1}$ such that $f_{n_k} \to f$ almost everywhere.*

Proof. Let $A \in \mathscr{A}$ be a set of finite μ-measure. We claim that

$$\lim_{n,m\to\infty} \int_S |f_n - f_m| \wedge \mathbf{1}_A \,\mathrm{d}\mu = 0.$$

To prove this let $\varepsilon > 0$ be arbitrary. Fix $r \in (0,1)$ so small that $r\mu(A) < \varepsilon$. Then

$$\begin{aligned}\int_S |f_n - f_m| \wedge \mathbf{1}_A \,\mathrm{d}\mu &\leqslant \mu(A \cap \{|f_n - f_m| > r\}) + \int_{A\cap\{|f_n-f_m|\leqslant r\}} r\,\mathrm{d}\mu \\ &< \mu(A \cap \{|f_n - f_m| > r\}) + \varepsilon.\end{aligned}$$

Since the first term on the right-hand side tends to 0 as $n, m \to \infty$ and $\varepsilon > 0$ is arbitrary, this proves the claim.

Let $S^{(1)} \subseteq S^{(2)} \subseteq \dots$ be an exhausting sequence for μ. By the claim there is a sequence $n_1 < n_2 < \dots$ such that $\int_S |f_{n_k} - f_{n_j}| \wedge \mathbf{1}_{S^{(j)}} \,\mathrm{d}\mu < 2^{-j}$ whenever $k \geqslant j$. In particular, $\lim_{j,k\to\infty} |f_{n_k} - f_{n_j}| = 0$ in $L^1(S^{(m)})$ for all $m \geqslant 1$. Hence we may pass to a subsequence that converges almost everywhere on $S^{(1)}$, and then to a further subsequence that converges almost everywhere on $S^{(2)}$, and so on. A diagonal argument produces a subsequence converges almost everywhere on all of S. □

Next we prove that convergence in measure is a metrisable:

Proposition A.2.4. *Suppose* $(S, \mathscr{A}, \mu)$ *is* σ*-finite, and let* $S^{(1)} \subseteq S^{(2)} \subseteq \dots$ *be an exhausting sequence for* μ*. With respect to the metric*

$$d(f,g) := \sum_{k\geqslant 1} \frac{1}{2^k} d^{(k)}(f,g),$$

with

$$d^{(k)}(f,g) := \inf\Big\{r > 0 :\ \mu\big(S^{(k)} \cap \{|f-g| \geqslant r\}\big) \leqslant r\Big\} \wedge 1,$$

$L^0(S)$ *is a complete metric space, and we have* $\lim_{n\to\infty} f_n = f$ *in measure if and only* $\lim_{n\to\infty} d(f_n, f) = 0$.

Proof. Let us first check that d is indeed a metric. If $d(f,g) = 0$, then for all $k \geqslant 1$ we have $d^{(k)}(f,g) = 0$ and there exist $r_n \downarrow 0$ such that $\mu(S^{(k)} \cap \{|f-g| > r_n\}) \leqslant r_n$, which forces that $f = g$ μ-almost everywhere on $S^{(k)}$. In the converse direction it is clear that $d(f,f) = 0$. It is also clear that $d(f,g) = d(g,f)$. To prove the triangle inequality, fix $k \geqslant 1$. suppose $r, s > 0$ are such that $\mu(S^{(k)} \cap \{|f-g| \geqslant r\}) \leqslant r$ and $\mu(S^{(k)} \cap \{|g-h| \geqslant s\}) \leqslant s$. Then $\{|f-g| < r\} \cup \{|g-h| < s\} \subseteq \{|f-h| < r+s\}$ implies

$$\begin{aligned}&\mu(S^{(k)} \cap \{|f-h| > r+s\}) \\ &\qquad \leqslant \mu(S^{(k)} \cap \{|f-g| \geqslant r\}) + \mu(S^{(k)} \cap \{|f-g| \geqslant s\}) \leqslant r+s.\end{aligned}$$

It follows that

$$d^{(k)}(f,h) \leqslant \inf\{r>0:\ \mu(S^{(k)} \cap \{|f-g| \geqslant r\}) \leqslant r\}$$
$$+\inf\{s>0:\ \mu(S^{(k)} \cap \{|g-h| \geqslant s\}) \leqslant s\},$$

and since $d^{(k)}(f,h) \leqslant 1$ this in turn implies that $d^{(k)}(f,h) \leqslant d^{(k)}(f,g) + d^{(k)}(g,h)$.

Suppose next that $d(f_n,f) \to 0$. Then $\mu(S^{(k)} \cap \{|f_n - f| \geqslant r\}) \to 0$ for all $k \geqslant 1$ and $r>0$, and from this an easy exhaustion argument implies that if $\mu(A)<\infty$, then $\mu(A \cap \{|f_n - f| \geqslant r\}) \to 0$ for all $r>0$. It follows that $f_n \to f$ in measure. Conversely, if $f_n \to f$ in measure, then $d^{(k)}(f_n,f) \to 0$ for all $k \geqslant 1$, and since also $0 \leqslant d(f_n,f) \leqslant 1$, this implies $d(f_n,f) \to 0$.

Finally, to prove completeness, suppose that $d(f_n,f_m) \to 0$ as $n,m\to\infty$. Then $(f_n)_{n\geqslant 1}$ is Cauchy in measure. By Lemma A.2.3, every subsequence contains a subsequence converging almost everywhere, hence in measure by Lemma A.2.2. It follows that every subsequence of $(f_n)_{n\geqslant 1}$ contains a subsequence converging to f in the metric d. Therefore, $d(f_n,f) \to 0$. □

A.3 Uniform integrability

The results of this section will play a role in Chapters 2 and 3. Again, these results are usually stated for probability spaces only.

Definition A.3.1 (Uniform integrability). *Let $(S,\mathscr{A},\mu)$ be a measure space. A family $T \subseteq L^1(S)$ is called* uniformly integrable *if the following two conditions are satisfied:*

(i) *For all $\varepsilon>0$ there exists a number $r>0$ such that*

$$\sup_{f\in T}\int_S \mathbf{1}_{\{|f|>r\}}|f|\,\mathrm{d}\mu \leqslant \varepsilon.$$

(ii) *For all $\varepsilon>0$ there exists a set $B\in\mathscr{A}$ with $\mu(B)<\infty$ such that*

$$\sup_{f\in T}\int_{\complement B} |f|\,\mathrm{d}\mu \leqslant \varepsilon.$$

Condition (ii) is trivially fulfilled when μ is a finite measure. The standard unit vectors of ℓ^1 satisfy (i) but not (ii).

Example A.3.2. Let $(S,\mathscr{A},\mu)$ be a finite measure space, let $1<p<\infty$. We claim that every bounded subset T of $L^p(S)$ is uniformly integrable. To see this, put $C := \sup_{f\in T}\int_S |f|^p\,\mathrm{d}\mu$. On the sets $\{|f|>r\}$ we have $|f| \leqslant |f|^p/r^{p-1}$. Therefore,

$$\sup_{f\in T}\int_S \mathbf{1}_{\{|f|>r\}}|f|\,\mathrm{d}\mu \leqslant \sup_{f\in T}\int_S |f|^p/r^{p-1}\,\mathrm{d}\mu = C/r^{p-1},$$

which tends to 0 as $r\to\infty$.

Proposition A.3.3. *A subset $T \subseteq L^1(S)$ is uniformly integrable if and only if it is bounded in $L^1(S)$, satisfies condition (ii) of Definition A.3.1, and for every $\varepsilon > 0$ there exists a number $\delta > 0$ such that for all $A \in \mathscr{A}$ we have*

$$\mu(A) < \delta \implies \sup_{f \in T} \int_A |f| \, \mathrm{d}\mu \leqslant \varepsilon.$$

The boundedness assumption cannot be omitted here, as can be seen by considering the set $\{ne_1 : \ n \geqslant 1\} \subseteq \ell^1$ with e_n denoting the nth standard unit vector.

Proof. First we prove the 'if' part. Choose $\varepsilon > 0$ arbitrary and choose $\delta > 0$ according to the assumption. Put $C := \sup_{f \in T} \|f\|_1$ and choose $r > C/\delta$. Then $\mu(|f| > r) \leqslant C/r < \delta$ and therefore condition (i) of Definition A.3.1 follows if we apply the assumption with $A = \{|f| > r\}$.

For the proof of the 'only if' part let again $\varepsilon > 0$ be arbitrary and choose $r > 0$ such that

$$\sup_{f \in T} \int_S \mathbf{1}_{\{|f| > r\}} |f| \, \mathrm{d}\mu \leqslant \varepsilon.$$

Taking $\delta = \varepsilon/r$, for all $f \in T$ and $A \in \mathscr{A}$ with $\mu(A) < \delta$ we find

$$\int_A |f| \, \mathrm{d}\mu \leqslant \int_{A \cap \{|f| > r\}} |f| \, \mathrm{d}\mu + r\mu(A) \leqslant 2\varepsilon.$$

To see that T is bounded in $L^1(S)$ choose $B \in \mathscr{A}$ of finite μ-measure such that $\sup_{f \in T} \int_{\complement B} |f| \, \mathrm{d}\mu \leqslant \varepsilon$. Then

$$\|f\|_1 \leqslant \varepsilon + \int_S \mathbf{1}_{\{|f| > r\}} |f| \, \mathrm{d}\mu + \int_B \mathbf{1}_{\{|f| \leqslant r\}} |f| \, \mathrm{d}\mu \leqslant 2\varepsilon + r\mu(B).$$

□

The following result is an immediate consequence of the definition of uniform integrability and Proposition A.3.3.

Proposition A.3.4. *The following assertions hold:*

(1) *Every singleton is uniformly integrable.*
(2) *If T_1 and T_2 are uniformly integrable, then the algebraic sum*

$$T_1 + T_2 := \{f_1 + f_2 : f_1 \in T_1, f_2 \in T_2\}$$

is uniformly integrable.

(3) *If T is uniformly integrable, then the solid hull*

$$\{g : \ \exists f \in T \text{ such that } |g| \leqslant |f| \text{ almost everywhere}\}$$

is uniformly integrable.

Note that (1) is equivalent to the statement that $\nu(A) := \int_A |f| \, d\mu$ defines an absolute continuous measure.

Next we formulate the convergence result which can be proved under an uniform integrability assumption.

Proposition A.3.5. *Let $(f_n)_{n\geqslant 1}$ be a sequence of measurable functions. Then $f_n \to f$ in $L^1(S)$ if and only if $(f_n)_{n\geqslant 1}$ is uniformly integrable and $f_n \to f$ in measure.*

Proof. To prove the 'if' part, suppose that $(f_n)_{n\geqslant 1}$ is uniformly integrable and $f_n \to f$ in measure. By Proposition A.3.4 the family $\{|f_n - f| : n \geqslant 1\}$ is bounded in $L^1(S)$. Since also $|f_n - f| \to 0$ in measure, it suffices to consider the case $f = 0$. Moreover, by considering subsequences of subsequences it is enough to prove that there is a subsequence $(f_{n_k})_{k\geqslant 1}$ such that $\|f_{n_k}\|_1 \to 0$.

Let $\varepsilon > 0$ be arbitrary. Choose $r > 0$ and $A \in \mathscr{A}$ of finite measure such that

$$\|\mathbf{1}_{\complement A} f_n\|_1 < \varepsilon \quad \text{and} \quad \int_S \mathbf{1}_{\{|f_n|>r\}} |f_n| \, d\mu \leqslant \varepsilon \quad \text{for all } n \geqslant 1.$$

From Lemma A.2.3 we obtain a subsequence $(f_{n_k})_{k\geqslant 1}$ of $(f_n)_{n\geqslant 1}$ such that $\mathbf{1}_A f_{n_k} \to 0$ almost everywhere. Then,

$$\begin{aligned}
\|f_{n_k}\|_1 &\leqslant \varepsilon + \|\mathbf{1}_A f_{n_k}\|_1 \\
&\leqslant \varepsilon + \int_S \mathbf{1}_{\{|f_{n_k}|>r\}} |f_{n_k}| \, d\mu + \int_S \mathbf{1}_{\{|f_{n_k}|\leqslant r\}} \mathbf{1}_A |f_{n_k}| \, d\mu \\
&\leqslant 2\varepsilon + \int_S \mathbf{1}_{\{|f_{n_k}|\leqslant r\}} \mathbf{1}_A |f_{n_k}| \, d\mu.
\end{aligned}$$

By dominated convergence theorem it follows that $\limsup_{k\to\infty} \|f_{n_k}\|_1 \leqslant 2\varepsilon$. Since $\varepsilon > 0$ was arbitrary the result follows.

To prove the 'only if' part, suppose that $f_n \to f$ in $L^1(S)$. For $r > 0$ and one has

$$\mu(\{|f_n - f| > r\}) \leqslant \frac{1}{r} \|f_n - f\|_1 \to 0$$

as $n \to \infty$ and the convergence in μ-measure follows. To prove uniform integrability, by Proposition A.3.4 it is enough to prove that $\{f_n - f : n \geqslant 1\}$ is uniformly integrable. Thus it suffices to consider the case $f = 0$.

Fix an arbitrary $\varepsilon > 0$. Let $N \geqslant 1$ be such that $\|f_n\|_1 < \varepsilon$ for all $n \geqslant N$. For all $r > 0$ it is trivial that $\int_{\{|f_n|>r\}} |f_n| \, d\mu < \varepsilon$ for all $n \geqslant N$. For each $1 \leqslant n \leqslant N$ we can apply Proposition A.3.4 to find an $r_n > 0$ such that $\int_{\{|f_n|>r_n\}} |f_n| \, d\mu < \varepsilon$. Thus we have verified condition (i) of Definition A.3.1 for $r := \max\{r_n : 1 \leqslant n \leqslant N\}$.

To verify condition (ii) we argue similarly. With N as before, for each $1 \leqslant n \leqslant N$ we choose a set $B_n \in \mathscr{A}$ of finite measure such that $\|\mathbf{1}_{\complement B_n} f_n\|_1 < \varepsilon$. Then for $B := \bigcup_{n=1}^N B_n$ we have $\|\mathbf{1}_{\complement B} f_n\|_1 < \varepsilon$ for all $n \geqslant 1$. $\square$

A.4 Notes

Accessible references to measure theory and integration include Adams and Guillemin [1996] and Bauer [2001]. For advanced comprehensive treatments we refer the reader to Bogachev [2007a,b] and Fremlin [2000, 2001, 2002, 2003].

Section A.1

Proposition A.1.8 appears in Sierpinski [1922]; see also Fremlin [2001, Proposition 215D].

Section A.2

Although they are usually formulated for probability spaces only, the results of this section are standard. We refer the reader to the classical references on the subject, Billingsley [1999] and Parthasarathy [1967].

Section A.3

Our definition of uniform integrability is equivalent to the one in Fremlin [2001]. Good treatments of uniform integrability on finite measure spaces include Hewitt and Stromberg [1975] and Kallenberg [2002].

B

Banach spaces

Unless stated otherwise, all Banach spaces are taken over the scalar field $\mathbb{K}$ which may be either $\mathbb{R}$ or $\mathbb{C}$. When X is a Banach space, the norm of an element $x \in X$ is denoted by $\|x\|_X$, or, if no confusion can arise, by $\|x\|$. We denote by

$$B_X := \{x \in X : \ \|x\| < 1\}$$

the open unit ball of X.

B.1 Duality

The *dual* of a Banach space X is the Banach space X^* consisting of all continuous linear mappings $x^* : X \to \mathbb{K}$ endowed with the norm $\|x^*\|_{X^*} := \sup_{\|x\| \leqslant 1} |x^*(x)|$. We shall use the notation

$$\langle x, x^* \rangle := x^*(x)$$

to denote the duality pairing of the elements $x \in X$ and $x^* \in X^*$. Traditionally, the elements of X^* are usually referred to as the *functionals* on X

B.1.a Hahn–Banach theorems

The main result on duality is the so-called Hahn–Banach theorem, which provides an abundant supply of functionals. The Hahn–Banach theorem can be stated in various degrees of generality. At various occasions in these volumes we will need the theorem in its most general form.

Theorem B.1.1 (Hahn–Banach theorem for real vector spaces). *Let V be a vector space and let $W \subseteq V$ be a subspace.*

(1) *Let $\mathbb{K} = \mathbb{R}$ and suppose that $p : V \to \mathbb{R}$ is* sub-linear*, i.e., for all $x, y \in V$ and $t \geqslant 0$ we have*

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 63, DOI 10.1007/978-3-319-48520-1

$$p(x+y) \leqslant p(x) + p(y), \quad p(tx) = tp(x).$$

If $\phi : W \to \mathbb{R}$ is a linear mapping satisfying

$$\phi(x) \leqslant p(x) \quad \forall x \in W,$$

then there exists a linear mapping $\Phi : V \to \mathbb{R}$ that extends ϕ and satisfies

$$-p(-x) \leqslant \Phi(x) \leqslant p(x) \quad \forall x \in V.$$

(2) *Let $K = \mathbb{R}$ or $\mathbb{C}$ and suppose that $p : V \to [0, \infty)$ is a* seminorm*, i.e., for all $x, y \in V$ and $t \in \mathbb{K}$ we have*

$$p(x+y) \leqslant p(x) + p(y), \quad |p(tx)| = |t| p(x).$$

If $\phi : W \to \mathbb{R}$ is a linear mapping satisfying

$$|\phi(x)| \leqslant p(x) \quad \forall x \in W,$$

then there exists a linear mapping $\Phi : V \to \mathbb{R}$ that extends ϕ and satisfies

$$|\Phi(x)| \leqslant p(x) \quad \forall x \in V.$$

When applied to Banach spaces, this theorem takes the following form.

Theorem B.1.2 (Hahn–Banach extension theorem). *Let X be a Banach space and let $Y \subseteq X$ be a closed subspace. Then for every $y^* \in Y^*$ there exists a functional $x^* \in X^*$ that extends y^* and satisfies $\|x^*\|_{X^*} = \|y^*\|_{Y^*}$.*

As a consequence (take Y to be the linear span of x and consider $y^* : cx \mapsto c$), for all $x \in X$ we have

$$\|x\| = \sup_{\|x^*\| \leqslant 1} |\langle x, x^* \rangle|.$$

A vector space endowed with a topology is called a *topological vector space* if the vector space operations $(c, x) \mapsto cx$ and $(x, y) \to x + y$ are continuous (with respect to the product topologies of $\mathbb{K} \times X$ and $X \times X$). Such a space is called *Hausdorff* if every two distinct points are contained in disjoint open sets. A topological vector space is called *locally convex* if it is Hausdorff and every neighbourhood of 0 contains a convex neighbourhood of 0.

Theorem B.1.3 (Hahn–Banach separation theorem). *Let C and D be non-empty disjoint convex subsets of a locally convex topological vector space X.*

(1) *If C is open, then there exists a continuous linear mapping $x^* : X \to \mathbb{K}$ and a real number $t \in \mathbb{R}$ such that*

$$\Re\langle x, x^* \rangle \leqslant t < \Re\langle y, x^* \rangle \quad \forall x \in C, \, y \in D.$$

(2) *If C is compact, D is closed, and X is locally convex, then there exists a continuous linear mapping $x^* : X \to \mathbb{K}$ and real numbers $s, t \in \mathbb{R}$ such that*

$$\Re\langle x, x^*\rangle \leqslant s < t \leqslant \Re\langle y, x^*\rangle \quad \forall x \in C,\, y \in D.$$

As an application of the Hahn–Banach theorems, we identify the duals of subspaces and quotients of Banach spaces. The *annihilator* of a subspace $Y \subseteq X$ is the subspace $Y^\perp$ of X^* defined by

$$Y^\perp := \{x^* \in X^* :\ \langle y, x^*\rangle = 0 \text{ for all } y \in Y\}.$$

Likewise, the *pre-annihilator* of a subspace $Y \subseteq X^*$ is the subspace ${}^\perp Y$ of X defined by

$${}^\perp Y := \{x \in X :\ \langle x, y\rangle = 0 \text{ for all } y \in Y\}.$$

Annihilators and pre-annihilators are useful in describing the duals of subspaces and quotient spaces. We will use the notation $[x]$ to denote the equivalence class of an element x.

Proposition B.1.4. *Let X be a Banach space.*

(1) *Let $Y \subseteq X$ be a closed subspace. Then*

$$(X/Y)^* = Y^\perp$$

isometrically. The isomorphism is given by identifying $\phi \in Y^\perp$ with $\Phi \in (X/Y)^*$ if $\phi(x) = \Phi([x])$ for all $x \in X$.

(2) *Let $Y \subseteq X$ be a closed subspace. Then*

$$Y^* = X^*/Y^\perp$$

isometrically. The isomorphism is given by identifying $\phi \in Y^*$ with $[\Phi] \in X^*/Y^\perp$, where $\Phi \in X^*$ is any bounded linear extension of ϕ to all X.

Proof. (1): Observe first that if either $\phi \in Y^\perp \subseteq X^*$ or $\Phi \in (X/Y)^*$ is given, the formula $\phi(x) = \Phi([x])$ well-defines the other one. Moreover,

$$|\phi(x)| = |\Phi([x])| \leqslant \|\Phi\|_{(X/Y)^*}\|[x]\|_{X/Y} \leqslant \|\Phi\|_{(X/Y)^*}\|x\|,$$

so that $\|\phi\|_{X^*} \leqslant \|\Phi\|_{(X/Y)^*}$. Similarly, for all $y \in Y$,

$$|\Phi([x])| = |\phi(x)| = |\phi(x+y)| \leqslant \|\phi\|_{X^*}\|x+y\|_X;$$

thus $|\Phi([x])| \leqslant \|\phi\|_{X^*}\|[x]\|_{X/Y}$ and hence $|\Phi|(X/Y)^* \leqslant \|\phi\|_{X^*}$.

(2): Given $\phi \in Y^*$, the Hahn–Banach theorem guarantees an extension $\Phi_0 \in X^*$ with $\|\Phi_0\|_{X^*} = \|\phi\|_{Y^*}$. Moreover, by definition, every other extension $\Phi \in X^*$ satisfies $\Phi - \Phi_0 \in Y^\perp$. Thus $\phi \in Y^*$ is indeed identified with a unique equivalence class $[\Phi] = [\Phi_0] \in X^*/Y^\perp$, and its norm satisfies

$$\|[\Phi]\|_{X^*/Y^\perp} \leqslant \|\Phi_0\|_{X^*} = \|\phi\|_{Y^*}.$$

Conversely, given $[\Phi] \in X^*/Y^\perp$, we obtain an element $\phi \in Y^*$ by restricting Φ to Y; this restriction only depends on the equivalence class of Φ. Moreover,

$$\|\phi\|_{Y^*} \leqslant \inf_{\Psi \in Y^\perp} \|\Phi + \Psi\|_{X^*} = \|[\Phi]\|_{X^*/Y^\perp}.$$

□

Proposition B.1.5. *Let X be a Banach space.*

(1) *Let $Y \subseteq X$ be a linear subspace. Then ${}^\perp(Y^\perp) = \overline{Y}$.*
(2) *Let $Y \subseteq X^*$ be a linear subspace. Then $({}^\perp Y)^\perp = \overline{Y}^{\text{weak}^*}$.*

Proof. We first prove (2). It is clear that $Y \subseteq ({}^\perp Y)^\perp$ and that annihilators are weak* closed. This gives the inclusion $\overline{Y}^{\text{weak}^*} \subseteq ({}^\perp Y)^\perp$.

To prove the converse inclusion suppose $x^* \notin \overline{Y}^{\text{weak}^*}$. By the Hahn–Banach theorem (applied to the locally convex Hausdorff space (X^*, weak^*)) there is an $x \in {}^\perp Y$ such that $\langle x, x^*\rangle \neq 0$. The latter means that x^* does not vanish on ${}^\perp Y$, that is, $x^* \notin ({}^\perp Y)^\perp$.

The proof of (1) is entirely similar and produces the equality ${}^\perp(Y^\perp) = \overline{Y}^{\text{weak}}$. But $\overline{Y}^{\text{weak}} = \overline{Y}$ by the Hahn–Banach theorem. □

B.1.b Weak topologies

The *weak topology* of a Banach space X is the topology $\tau(X, X^*)$ on X generated by X^*. It is the is the coarsest topology on X with the property that for all $x^* \in X^*$ the functions $x \mapsto \langle x, x^*\rangle$ are continuous and is generated by all sets of the form

$$V(x_0, x_0^*, \varepsilon) := \{x \in X : \ |\langle x_0 - x, x_0^*\rangle| < \varepsilon\}$$

with given $x_0 \in X$, $x_0^* \in X^*$, and $\varepsilon > 0$. A set $O \subseteq X$ is open in the weak topology if and only if for all $x_0 \in O$ there exist $x_1^*, \dots, x_k^* \in X^*$ and a number $\varepsilon > 0$ such that

$$\bigcap_{j=1}^{k} \{x \in X : \ |\langle x_0 - x, x_j^*\rangle| < \varepsilon\} \subseteq O.$$

It is easy to check that $\lim_{n\to\infty} x_n = x$ in the weak topology if and only if $\lim_{n\to\infty}\langle x_n, x^*\rangle = \langle x, x^*\rangle$ for all $x^* \in X^*$.

As an immediate consequence of the Hahn–Banach separation theorem we record the following fact.

Theorem B.1.6. *A convex set in a Banach space X is closed if and only if it is weakly closed.*

The *weak*-topology* of X^* is the topology $\tau(X^*, X)$ on X^* generated by X. It is the is the coarsest topology on X^* with the property that for all $x \in X$ the functions $x^* \mapsto \langle x, x^*\rangle$ are continuous. It is generated by all sets of the form

$$V(x_0, y_0^*, \varepsilon) := \{x^* \in X^* : \ |\langle x_0, x_0^* - x^*\rangle| < \varepsilon\}$$

with given $x_0 \in X$, $x_0^* \in X^*$, and $\varepsilon > 0$.

A set $O \subseteq X^*$ is open in the weak*-topology if and only if for all $x_0^* \in O$ there exist $x_1, \dots, x_k \in X$ and a number $\varepsilon > 0$ such that

$$\bigcap_{j=1}^{k} \{x^* \in X^* : \ |\langle x_j, x_0^* - x^*\rangle| < \varepsilon\} \subseteq O.$$

It is easy to check that $\lim_{n\to\infty} x_n^* = x^*$ in the weak topology if and only if $\lim_{n\to\infty}\langle x, x_n^*\rangle = \langle x, x_n^*\rangle$ for all $x \in X$.

Theorem B.1.7 (Banach–Alaoglu theorem). *The closed unit ball of the dual of any Banach space is weak*-compact. If X is separable, this set is weak*-sequentially compact.*

The first assertion is a straightforward consequence of the Tychonov theorem on compactness of products of compact sets, noting that the mapping from $\overline{B}_{X^*}$ to $\prod_{x\in\overline{B}_X} \overline{B}_{\mathbb{K}}$ given by $x^* \mapsto (\langle x, x^*\rangle)_{x\in\overline{B}_X}$ is a homeomorphism into its image. The second statement is a consequence of the following proposition.

Proposition B.1.8. *Let X be a Banach space. The following assertions are equivalent:*

(1) *X is separable;*
(2) *The weak*-topology of $\overline{B}_{X^*}$ is metrisable.*

Proof. (1)⇒(2): Let $(x_n)_{n\geqslant 1}$ be a dense sequence in the closed unit ball $\overline{B}_X$ of X; such a sequence exists since X is separable. It is easily checked that the formula

$$d(x^*, y^*) := \sum_{n\geqslant 1} \frac{1}{2^n} \, \frac{|\langle x_n, x^* - y^*\rangle|}{1 + |\langle x_n, x^* - y^*\rangle|}$$

defines a metric on $\overline{B}_{X^*}$ and that the identity mapping $I_{X^*} : (\overline{B}_{X^*}, \text{weak}^*) \to (\overline{B}_{X^*}, d)$ is continuous. By the Banach–Alaoglu theorem, $(\overline{B}_{X^*}, \text{weak}^*)$ is compact, and therefore I_{X^*} is a homeomorphism.

(2)⇒(1): Let d be a metric which induces the weak*-topology of $\overline{B}_{X^*}$. By the Banach-Alaoglu theorem, this metric turns $\overline{B}_{X^*}$ into a compact metric space. By a standard result from topology, this implies that the Banach space $C(\overline{B}_{X^*})$ is separable. The separability of X follows from this by noting that the mapping $x \mapsto f_x$, where $f_x(x^*) := \langle x, x^*\rangle$, defines an isometric embedding of X into $C(\overline{B}_{X^*})$. $\square$

We continue with some further useful duality results.

Proposition B.1.9. *If X^* is separable, then X is separable.*

Proof. Choose a dense sequence $(x_n^*)_{n=1}^\infty$ in the closed unit sphere $S_{X^*} = \{x^* \in X^* : \ \|x^*\| = 1\}$, and let $(x_n)_{n=1}^\infty$ be a sequence of norm one vectors in X such that $|\langle x_n, x_n^*\rangle| \geqslant \frac{1}{2}$. We claim that the set of all $\mathbb{L}$-linear combinations of the elements x_n is dense in X, where $\mathbb{L}$ is any countable dense subset of $\mathbb{K}$. If this were false, by the Hahn–Banach theorem we find a norm one element $x^* \in X^*$ such that $\langle x_n, x^*\rangle = 0$ for all $n \geqslant 1$. Choosing $m \geqslant 1$ such that $\|x_m^* - x^*\| < \frac{1}{2}$ we obtain the contradiction $|\langle x_m, x_m^*\rangle| = |\langle x_m, x_m^* - x^*\rangle| < \frac{1}{2}$. □

A subset $Y \subseteq X^*$ is called *norming* for a subset $F \subseteq X$ if for all $x \in F$ we have

$$\sup_{x^* \in Y \setminus \{0\}} \frac{|\langle x, x^*\rangle|}{\|x^*\|} = \|x\|.$$

A subset of X^* which is norming for X is simply said to be *norming.*

Proposition B.1.10. *If $F \subseteq X$ is a separable subset and $Y \subseteq X^*$ is a norming subset for F, then Y contains a sequence which is norming for F.*

Proof. Choose a dense sequence $(x_n)_{n\geqslant 1}$ in F and choose a sequence of non-zero vectors $(x_n^*)_{n\geqslant 1}$ in Y such that $|\langle x_n, x_n^*\rangle| \geqslant (1-\varepsilon_n)\|x_n\|\|x_n^*\|$ for all $n \geqslant 1$, where the numbers $0 < \varepsilon_n \leqslant 1$ satisfy $\lim_{n\to\infty} \varepsilon_n = 0$. The sequence $(x_n^*)_{n\geqslant 1}$ is norming for F. To see this, fix an arbitrary $x \in F$ and let $\delta > 0$. Pick $n_0 \geqslant 1$ such that $0 < \varepsilon_{n_0} \leqslant \delta$ and $\|x - x_{n_0}\| \leqslant \delta$. Then,

$$\begin{aligned}
(1-\delta)\|x\|\|x_{n_0}^*\| &\leqslant (1-\varepsilon_{n_0})\|x\|\|x_{n_0}^*\| \\
&\leqslant (1-\varepsilon_{n_0})\|x_{n_0}\|\|x_{n_0}^*\| + (1-\varepsilon_{n_0})\delta\|x_{n_0}^*\| \\
&\leqslant |\langle x_{n_0}, x_{n_0}^*\rangle| + \delta\|x_{n_0}^*\| \\
&\leqslant |\langle x, x_{n_0}^*\rangle| + 2\delta\|x_{n_0}^*\|.
\end{aligned}$$

Since $\delta > 0$ was arbitrary it follows that $\|x\| \leqslant \sup_{n\geqslant 1} |\langle x, x_n^*\rangle|/\|x_n^*\|$. The opposite inequality is trivial. □

A subset $Y \subseteq X^*$ is said to *separate the points* of a subset F of X if for every pair $x, y \in F$ with $x \neq y$ there exists an $x^* \in Y$ with $\langle x, x^*\rangle \neq \langle y, x^*\rangle$. By the Hahn–Banach theorem, a linear subspace of X^* separates the points of X if and only if it is weak*-dense in X^*.

Proposition B.1.11. *If $F \subseteq X$ is a separable subset and $Y \subseteq X^*$ is a weak*-dense subset, then Y contains a sequence which separates the points of F.*

Proof. Let $G := \{x - y : x \in F, y \in F, x \neq y\}$. For each $g \in G$ there exists a vector $x_g^* \in Y \cap \{x^* \in X^* : \langle g, x^* \rangle \neq 0\}$. Defining

$$V_g := \{x \in G : \langle x, x_g^* \rangle \neq 0\}$$

we obtain an open cover $(V_g)_{g \in G}$ of the separable metric space G. By the Lindelöf theorem, open covers of separable metric spaces admit a countable sub-cover, and therefore $(V_g)_{g \in G}$ admits a countable sub-cover. It follows that there exists a sequence $(g_n)_{n=1}^\infty$ in G such that $(V_{g_n})_{n=1}^\infty$ covers G. Then every $x - y \in G$ with $x, y \in F$ and $x \neq y$ belongs to some V_{g_n}, which means that $\langle x - y, x_{g_n}^* \rangle \neq 0$. □

Theorem B.1.12 (Krein–Šmulian). *A linear subspace of X^* is weak*-closed if and only if its intersection with $\overline{B}_{X^*}$ is weak*-closed in $\overline{B}_{X^*}$.*

Corollary B.1.13. *An element $x^{**} \in X^{**}$ belongs to x if and only if it weak*-continuous on B_{X^*}.*

Corollary B.1.14. *If X is separable and Y is a linear subspace of X^* that is weak*-dense and weak*-sequentially closed, then $Y = X^*$.*

Proof. The closed unit ball $\overline{B}_{X^*}$ is weak*-closed, hence weak*-sequentially closed, in X^*. It follows that $\overline{B}_{X^*} \cap Y$ is a weak*-sequentially closed subset of $\overline{B}_{X^*}$. By Proposition B.1.10, the separability of X implies that the weak*-topology of $\overline{B}_{X^*}$ is metrisable. Since sequentially closed subsets of metric spaces are closed, $\overline{B}_{X^*} \cap Y$ is weak*-closed as a subset of $\overline{B}_{X^*}$. Hence, by the Krein–Šmulian theorem, Y is weak*-closed. By assumption Y is also weak*-dense, and therefore $Y = X^*$. □

The proof of the Krein–Šmulian is based on the fact that several natural topologies coincide on X^*.

Lemma B.1.15. *The following topologies coincide on the dual X^* of a Banach space X:*

(1) *The topology $\tau_{\rm n}$ of uniform convergence on null sequences of X;*
(2) *The topology $\tau_{\rm c}$ of uniform convergence on compact subsets of X;*
(3) *The finest topology $\tau_{\rm bw^*}$ that coincides with the weak*-topology on every bounded subset of X^*.*

To be explicit, the topologies $\tau_{\rm n}$ and $\tau_{\rm c}$ are the locally convex topologies generated by the seminorms $p_\xi(x^*) := \sup_j |\langle \xi_j, x^* \rangle|$ and $p_K(x^*) := \sup_{x \in K} |\langle x, x^* \rangle|$, with $\xi_j \to 0$ and K ranging over all null sequences and compact subsets of X, respectively. These topologies are generated by the sets of the form $\{x^* \in X^* : p_\xi(x^* - x_0^*) < \varepsilon\}$ and $\{x^* \in X^* : p_K(x^* - x_0^*) < \varepsilon\}$ respectively, with $x_0^* \in X^*$ and $\varepsilon > 0$. The topology $\tau_{\rm bw^*}$ is defined as consisting of those sets $U \subseteq X^*$ with the property that for every bounded set $B \subseteq X^*$ there exists a weak*-open set $U' \subseteq X^*$ such that $U \cap B = U' \cap B$. All three topologies are translation invariant and contained as sub-topologies in the norm topology of X^*.

Proof of Lemma B.1.15. Since null sequences are relatively compact, it is clear that $\tau_{\mathrm{n}} \subseteq \tau_{\mathrm{c}}$.

To prove that $\tau_{\mathrm{c}} \subseteq \tau_{\mathrm{bw}^*}$, by translation invariance it suffices to prove that every set of the form

$$U_\varepsilon := \Big\{x^* \in X^* : \sup_{x\in K} |\langle x, x^*\rangle| < \varepsilon\Big\},$$

with K compact in E, belongs to τ_{bw^*}.

Pick an arbitrary $\delta > 0$. Let $B \subseteq X^*$ be bounded, say $\|x^*\| \leqslant M$ for all $x^* \in B$, let $x_1, \dots, x_N$ be a δ/M-net in K, and for $\eta > 0$ consider the weak*-open set

$$U'_\eta := \{x^* \in X^* : |\langle x_n, x^*\rangle| < \eta, \ n = 1, \dots N\}.$$

It is clear that $U_\varepsilon \subseteq U'_\varepsilon$. Conversely, let $0 < \eta < \varepsilon$ and fix an $x^* \in U'_\eta \cap B$. Fix an $x \in K$ and choose an index $1 \leqslant n \leqslant N$ such that $\|x - x_n\| \leqslant \delta/M$. Then

$$|\langle x, x^*\rangle| \leqslant \|x - x_n\|\|x^*\| + |\langle x_n, x^*\rangle| < \delta + \eta.$$

Since $\delta > 0$ was arbitrary, it follows that $|\langle x, x^*\rangle| \leqslant \eta$. This being true for all $x \in K$, it follows that $\sup_{x\in K} |\langle x, x^*\rangle| \leqslant \eta < \varepsilon$ and consequently $x^* \in U_\varepsilon$. This shows that $U'_\eta \cap B \subseteq U_\varepsilon$. Taking the union over all $0 < \eta < \varepsilon$ we get $U'_\varepsilon \cap B = \bigcup_{0<\eta<\varepsilon} U'_\eta \cap B \subseteq U_\varepsilon$. Putting together what we proved, we see that $U_\varepsilon \cap B = U'_\varepsilon \cap B$.

It remains prove that $\tau_{\mathrm{bw}^*} \subseteq \tau_{\mathrm{n}}$. Let a non-empty $U \in \tau_{\mathrm{bw}^*}$ be given. By translation invariance, it suffices to prove that if $0 \in U$, then $0 \in V \subseteq U$ for some τ_{n}-open set V.

Since U is norm-open, it contains a closed ball $\overline{B}$ centred at 0 of positive radius r. Let U'_1 be a weak*-open set such that $U \cap \overline{B} = U'_1 \cap \overline{B}$. The set U' contains a subset of the form $V_1 := \{x^* \in X^* : |\langle x_n, x^*\rangle| \leqslant 1, \ n = 1, \dots, N_1\}$ with $x_1, \dots, x_{N_1} \in X$. Set $K_1 := \{x_1, \dots, x_{N_1}\}$. Thus

$$V_1 \cap \overline{B} \subseteq U.$$

We claim that there exists a finite set K_2 in X, all of whose elements have norm $\leqslant 1/r$, say $K_2 = \{x_{N_1+1}, \dots, x_{N_2}\}$ with $N_2 \geqslant N_1$, such that if we put $V_2 := \{x^* \in X^* : |\langle x_n, x^*\rangle| \leqslant 1, \ n = 1, \dots, N_2\}$, then

$$V_2 \cap 2\overline{B} \subseteq U.$$

Suppose this were false. Then for all finite sets K in X consisting of elements of norm $\leqslant 1/r$ the set C_K of all $x^* \in 2\overline{B} \cap \complement U$ such that $|\langle x, x^*\rangle| \leqslant 1$ for all $x \in K_1 \cup K$ is non-empty. If $K^{(1)}, \dots, K^{(M)}$ are such finite sets, so is their union, and therefore $C_{K^{(1)}} \cap \dots \cap C_{K^{(M)}} = C_{K^{(1)} \cup \dots \cup K^{(M)}}$ is non-empty. Stated differently, the sets C_K have the finite intersection property. Also, by the definition of τ_{bw^*}, the set $2\overline{B} \cap \complement U$ is relatively weak*-closed in the weak*-compact set $2\overline{B}$ and therefore weak*-compact. The finite intersection property of the sets C_K therefore implies that their intersection is non-empty. This means that there exists an $x^* \in 2\overline{B} \cap \complement U$ such that for all $x \in X$ of norm $\leqslant 1/r$ and also for all $x \in K_1$ we have $|\langle x, x^* \rangle| \leqslant 1$. But then $x^* \in V_1$ and $\|x^*\| \leqslant r$, so $x^* \in V_1 \cap \overline{B} \subseteq U$. This contradicts $x^* \in \complement U$ and the claim is proved.

Proceeding inductively, we obtain finite sets $K_1, K_2, K_3, \dots$ in X, with $K_{i+1} = \{x_{N_i+1}, \dots, x_{N_{i+1}}\}$ consisting of elements of norm $\leqslant 2^{-i}r$ for all $i \geqslant 1$, such that with $V_i := \{x^* \in X^* : \ |\langle x_n, x^* \rangle| \leqslant 1, \ n = 1, \dots, N_i\}$ we have

$$V_i \cap 2^i \overline{B} \subseteq U.$$

The sequence $K_1 \cup K_2 \cup \cdots = (x_n)_{n \geqslant 1}$ is null. The set $V := \{x^* \in X^* : \sup_{n \geqslant 1} |\langle x_n, x^* \rangle| < 1\}$ is τ_{n}-open and contained in each of the sets V_i. For any $x^* \in V$ we may select i so large that $x^* \in 2^i\overline{B}$ and infer that $x^* \in V_i \cap 2^i \overline{B} \subseteq U$. This shows that $V \subseteq U$. □

Lemma B.1.16. *For a linear mapping $\phi : X^* \to \mathbb{K}$ the following assertions are equivalent:*

(1) *ϕ is weak*-continuous;*
(2) *ϕ is τ_{n}-continuous;*
(3) *ϕ is τ_{c}-continuous.*

Proof. From $\tau_{\mathrm{w}^*} \subseteq \tau_{\mathrm{n}} = \tau_{\mathrm{c}}$ it is clear that the implications (1)⇒(2)⇔(3) hold. To prove the implication (2)⇒(1) suppose that ϕ is τ_{c}-continuous. Then there is a null sequence $\xi = (\xi_j)_{j \geqslant 1}$ in X such that

$$p_\xi(x^*) < 1 \text{ implies } |\phi(x^*)| < 1. \tag{B.1}$$

Consider the mapping $T : X^* \to c_0$, $Tx^* := (\langle \xi_j, x^* \rangle)_{j \geqslant 1}$. If $Tx^* = 0$, then $p_\xi(\lambda x^*) = \lambda p_\xi(x^*) = 0$ for all $\lambda > 0$, forcing $\phi(x^*) = 0$. Hence on the range of T we may define $h(Tx^*) := \phi(x^*)$; this is well defined by the preceding consideration. By (B.1), h is bounded on the range of T. By the Hahn–Banach theorem, h has an extension to a bounded functional $h : c_0 \to \mathbb{K}$. Representing h by an element $(h_j)_{j \geqslant 1}$ in ℓ^1, we have

$$\phi(x^*) = h(Tx^*) = \sum_{j \geqslant 1} h_j \langle \xi_j, x^* \rangle = \Big\langle \sum_{j \geqslant 1} h_j \xi_j, x^* \Big\rangle.$$

Thus ϕ is represented by $\sum_{j \geqslant 1} h_j \xi_j \in X$, which shows that ϕ is weak*-continuous. □

Proof of Theorem B.1.12. Only the 'if' part needs proof. Suppose Y is a linear subspace of X^* whose intersection with $\overline{B}_{X^*}$ is weak*-closed in $\overline{B}_{X^*}$. By the weak*-continuity of scalar multiplication, $Y \cap (r\overline{B}_{X^*})$ is weak*-closed in $r\overline{B}_{X^*}$ for every $r > 0$. This in turn implies that $Y \cap B$ is closed in the relative weak*-topology of B for every bounded set $B \subseteq X^*$. This is the same as saying that Y is closed in the topology τ_{bw^*}. By Lemma B.1.15, Y is closed in τ_{c}. By the Hahn–Banach theorem for the locally convex topology τ_{c}, Y is the intersection of null spaces of τ_{c}-continuous functionals. These functionals are weak*-continuous by Lemma B.1.16, and therefore Y is weak*-closed. □

B.1.c Reflexivity

The mapping $j : X \to X^{**}$ defined by

$$\langle x^*, jx\rangle := \langle x, x^*\rangle, \quad x^* \in X^*,$$

defines an isometric embedding of X into the bi-dual X^{**}; this is an immediate consequence of the Hahn–Banach theorem. We shall always identify X with its image in X^{**}. The following result is known as *Goldstine's theorem*

Proposition B.1.17. *$\overline{B}_X$ is weak*-dense in $\overline{B}_{X^{**}}$.*

Proof. First note that $\overline{B}_X \subseteq \overline{B_X}^{\text{weak}^*}$ in X^{**}. Suppose, for a contradiction that a non-zero $x^{**} \in \overline{B}_{X^{**}} \setminus \overline{B_X}^{\text{weak}^*}$ exists. By the Hahn–Banach theorem for the locally convex space (X^{**}, weak^*), there exists $x^* \in X^*$, which we may take to be of norm one, and real numbers $a < b$ such that $\Re\langle x, x^*\rangle \leqslant a < b \leqslant \Re\langle x^*, x^{**}\rangle$ for all $x \in B_X$. Using first that $|\Re\langle x^*, x^{**}\rangle| \leqslant \|x^*\|\|x^{**}\| \leqslant 1$, we obtain $\|x^*\| = \sup_{x\in B_X} |\Re\langle x^*, x^*\rangle| \leqslant a < b \leqslant 1$. But this contradicts our assumption $\|x^*\| = 1$. □

A Banach space X is called *reflexive* if this embedding is surjective. As a consequence of the Banach–Alaoglu theorem and the weak*-density of $\overline{B}_X$ in $\overline{B}_{X^{**}}$ we have the following characterisation of reflexivity:

Proposition B.1.18. *A Banach space X is reflexive if and only if $\overline{B}_X$ is weakly compact.*

Reflexivity passes on to (pre-)duals:

Proposition B.1.19. *A Banach space X is reflexive if and only if its dual X^* is reflexive.*

For if X is reflexive, then $(X^*)^{**} = (X^{**})^* = X^*$ under the usual natural identifications, and therefore X^* is reflexive. If X^* is reflexive, then so is X^{**}, and therefore so is its closed subspace X.

B.2 Bounded linear operators

Let X and Y be Banach spaces. A linear operator $T : X \to Y$ is bounded (i.e., $\|Tx\| \leqslant C\|x\|$ for some constant $C \geqslant 0$ and all $x \in X$) if and only if it is continuous, if and only if it is continuous at some point $x_0 \in X$. The space $\mathscr{L}(X, Y)$ of all bounded linear operators from X to Y is a Banach space with respect to the operator norm $\|T\| := \sup_{\|x\|\leqslant 1} \|Tx\|$.

Theorem B.2.1 (Uniform boundedness principle). *Let I be an index set. If $(T_i)_{i\in I}$ is a family of operators in $\mathscr{L}(X,Y)$ which has the property that*

$$\sup_{i\in I} \|T_i x\| < \infty,$$

then the family is uniformly bounded,

$$\sup_{i\in I} \|T_i\| < \infty.$$

This is a consequence of the *Baire category theorem*, which asserts that if a complete metric space is the union of countably many closed subsets, then at least one of these subsets has non-empty interior. This uniform boundedness theorem is obtained by applying this result to the closed sets $\{x \in X : \ \sup_{i\in I} \|T_i x\| \leqslant n\}$ whose union equals X.

Theorem B.2.2 (Closed graph theorem). *A linear operator $T : X \to Y$ whose graph $\{(x, Tx) : \ x \in X\}$ is closed in $X \times Y$ is bounded.*

Here $X \times Y$ may be endowed with any norm which satisfies

$$\max\{\|x\|, \|y\|\} \leqslant \|(x,y)\| \leqslant \|x\| + \|y\|.$$

All such norms are equivalent and turn $X \times Y$ into a Banach space.

Theorem B.2.3 (Open mapping theorem). *If a bounded linear operator $T : X \to Y$ is surjective, then it is open. In particular, a bijective bounded linear operator has a bounded inverse.*

The *strong operator topology* of $\mathscr{L}(X,Y)$ is the topology generated by the mappings $T \mapsto Tx$, where x ranges over X. A set $O \subseteq \mathscr{L}(X,Y)$ is open in this topology if and only if for all $S \in O$ there exist $x_1, \dots, x_k \in X$ and a number $\varepsilon > 0$ such that

$$\bigcap_{j=1}^{k} \{T \in \mathscr{L}(X,Y) : \ \|Sx_j - Tx_j\| < \varepsilon\} \subseteq O.$$

It is easy to check that $\lim_{n\to\infty} T_n = T$ in the strong operator topology if and only if $\lim_{n\to\infty} T_n x = Tx$ for all $x \in X$.

The *weak operator topology* of $\mathscr{L}(X,Y)$ is the topology generated by the mappings $T \mapsto \langle Tx, y^*\rangle$, where x and y^* range over X and Y^*. A set $O \subseteq \mathscr{L}(X,Y)$ is open in this topology if and only if for all $S \in O$ there exist $x_1, \dots, x_k \in X$, $y_1^*, \dots, y_k^* \in Y^*$, and a number $\varepsilon > 0$ such that

$$\bigcap_{j=1}^{k} \{T \in \mathscr{L}(X,Y) : \ |\langle Sx_j - Tx_j, y_j^*\rangle| < \varepsilon\} \subseteq O.$$

One has $\lim_{n\to\infty} T_n = T$ in the weak operator topology if and only if $\lim_{n\to\infty} \langle T_n x, y^*\rangle = \langle Tx, y^*\rangle$ for all $x \in X$ and $y^* \in Y^*$.

Proposition B.2.4. *A convex set $\mathscr{T} \subseteq \mathscr{L}(X,Y)$ is closed in the strong operator topology if and only if it is closed in the weak operator topology.*

Proof. Suppose $\mathscr{T} \subseteq \mathscr{L}(X,Y)$ is convex and closed in the strong operator topology. Consider an operator $T_0 \notin \mathscr{T}$. Choose $x_1, \dots, x_N \in X$ and $\varepsilon > 0$ such that

$$V := \big\{T \in \mathscr{L}(X,Y) : \ \|(T - T_0)x_n\| < \varepsilon, \ n = 1, \dots, N\big\}$$

does not intersect $\mathscr{T}$. This means that for all $T \in \mathscr{T}$ we have $\sup_{1 \leqslant n \leqslant N} \|(T - T_0)x_n\| \geqslant \varepsilon$. In X^N consider the closed convex set $C := \{(Tx_n)_{n=1}^N : \ T \in \mathscr{T}\}$ and the point $c_0 := (T_0 x_n)_{n \geqslant 1}$. Then $\|c - c_0\| \geqslant \varepsilon$ for all $c \in C$, and therefore the Hahn–Banach theorem provides us with a functional $x^* \in (X^N)^* = (X^*)^N$, say $x^* = (x_n^*)_{n=1}^N$, that separates C and c_0:

$$\Re\langle c_0, x^*\rangle \leqslant a < b \leqslant \Re\langle c, x^*\rangle, \quad \forall c \in C.$$

It follows that for all $T \in \mathscr{T}$ we can find $1 \leqslant n \leqslant N$ such that

$$\Re\langle (T - T_0)x_n, x_n^*\rangle \geqslant \frac{b-a}{N}.$$

Therefore the weakly open set

$$W = \Big\{T \in \mathscr{L}(X,Y) : \ |\langle (T - T_0)x_n, x_n^*\rangle| < \frac{b-a}{N}, \ n = 1, \dots, N\Big\},$$

which contains T_0, does not intersect $\mathscr{T}$. □

The *adjoint* of an operator $T \in \mathscr{L}(X,Y)$ is the operator $T^* \in \mathscr{L}(Y^*, X^*)$ defined by

$$\langle x, T^* y^*\rangle := \langle Tx, y^*\rangle.$$

Note that $\|T^*\| \leqslant \|T\|$ and, repeating the construction starting from T^*, $\|T^{**}\| \leqslant \|T^*\|$. But T^{**} extends T, and therefore $\|T\| \leqslant \|T^{**}\|$. It follows that

$$\|T\| = \|T^*\|.$$

From the Hahn–Banach theorem one deduces:

Proposition B.2.5. *Let X and Y be Banach spaces.*

(1) *A linear operator $T : X \to Y$ is bounded if and only if it is continuous with respect to the weak topologies of X and Y;*
(2) *A linear operator $T : Y^* \to X^*$ is the adjoint of a bounded operator from X to Y if and only if it is continuous with respect to the weak*-topologies of Y^* and X^*.*

B.3 Holomorphic mappings

A *domain* is a connected open subset of $\mathbb{C}$. If $D \subseteq \mathbb{C}$ is a domain and X is a Banach space, a function $f : D \to X$ is called *holomorphic* if for all $z_0 \in D$ the limit

$$f'(z_0) := \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists in X.

Proposition B.3.1. *Let Z be a weak*-dense linear subspace of X^*. For a locally bounded and strongly measurable function $f : D \to X$ the following assertions are equivalent:*

(1) *f is holomorphic;*
(2) *$\langle f, x^* \rangle$ is holomorphic for all $x^* \in Z$.*

Remark B.3.2. It will be clear from the proof that the local boundedness assumption may be dropped if Z induces an equivalent norm on X. Under this stronger assumption, the first part of the proof becomes redundant, and the use of the uniform boundedness principle towards the end of the proof may be combined with the observation that X embeds into $\overline{Z}^*$ isomorphically.

Proof. Evidently we only need to prove that (2) implies (1).

By the strong measurability assumption and Pettis's measurability theorem (Theorem 1.1.20) we may assume that X is separable. Since the pointwise limit of a locally bounded sequence of holomorphic functions is holomorphic, we can apply Corollary B.1.14 to conclude that $\langle f, x^* \rangle$ is holomorphic for all $x^* \in X^*$.

Fix $z_0 \in D$ and let $r > 0$ be so small that the closed disc $\{z \in \mathbb{C} : |z - z_0| \leqslant r\}$ is contained in D. Set

$$V = \Big\{ \frac{1}{h-g} \Big(\frac{f(z_0+h) - f(z_0)}{h} - \frac{f(z_0+g) - f(z_0)}{g} \Big) : \ |g|, |h| < \frac{r}{2} \Big\}.$$

Fix $x^* \in Z$. For all $|g|, |h| < r/2$ we have

$$\begin{aligned}
&\Big| \frac{1}{h-g} \Big\langle \frac{f(z_0+h) - f(z_0)}{h} - \frac{f(z_0+g) - f(z_0)}{g}, x^* \Big\rangle \Big| \\
&\quad = \Big| \frac{1}{2\pi i} \frac{1}{h-g} \int_{|z-z_0|=r} \langle f(z), x^* \rangle \\
&\qquad\quad \times \Big(\frac{1}{h} \Big(\frac{1}{z-(z_0+h)} - \frac{1}{z-z_0} \Big) - \frac{1}{g} \Big(\frac{1}{z-(z_0+g)} - \frac{1}{z-z_0} \Big) \Big) \,\mathrm{d}z \Big| \\
&\quad = \Big| \frac{1}{2\pi i} \int_{|z-z_0|=r} \langle f(z), x^* \rangle \frac{1}{(z-(z_0+h))(z-(z_0+g))(z-z_0)} \Big) \,\mathrm{d}z \Big| \\
&\quad \leqslant \sup_{0 \leqslant \theta \leqslant 2\pi} |\langle f(re^{i\theta}, x^* \rangle| \cdot \frac{2}{r} \cdot \frac{2}{r} \cdot \frac{1}{r},
\end{aligned}$$

which is finite. This shows that for each $x^* \in X^*$ the set $\{\langle v, x^*\rangle : v \in V\}$ are bounded. By the uniform boundedness principle this implies that V is a bounded set. It follows that there exists a finite constant M such that

$$\Big\|\frac{f(z_0+h)-f(z_0)}{h} - \frac{f(z_0+g)-f(z_0)}{g}\Big\| \leqslant M|h-g|$$

whenever $|g|, |h| < r/2$. As a consequence, $\lim_{h\to 0}(f(z_0+h)-f(z_0))/h$ exists in X. □

This theorem allows one to generalise many classical results on holomorphic functions to the vector-valued case. Such extensions can then often be proved either by applying the scalar counterpart after composing the functions with a functional, or by repeating the proof of their scalar-valued counterparts. For instance, the derivative $f' : D \to X$ is again a holomorphic function (since each of the functions $\langle f', x^*\rangle$ is holomorphic), and vector-valued analogues of Cauchy's theorem hold (since this theorem holds for each of the functions $\langle f, x^*\rangle$). Vector-valued holomorphic functions can be locally expressed as power series with 'coefficients' in X (repeat the scalar proof) and the maximum principle holds: if D is open and $f : D \to X$ is holomorphic, then $z \mapsto \|f(z)\|$ cannot take a local maximum at any point in D (if $\|f(z)\|$ takes a local maximum in $z_0 \in D$ and if $x^* \in X^*$ has norm one and satisfies $|\langle f(z_0), x^*\rangle| = \|f(z_0)\|$, then we can apply the scalar maximum principle to $\langle f, x^*\rangle$ as we have $|\langle f(z_0), x^*\rangle| = \|f(z_0)\| \geqslant \|f(z)\| \geqslant |\langle f(z), x^*\rangle|$ for all $z \in D$).

Corollary B.3.3. *Let D be a domain in $\mathbb{C}$ and let X and Y be Banach spaces and let Z be a norming subspace of Y^*. For an operator-valued function $f : D \to \mathscr{L}(X, Y)$ the following assertions are equivalent:*

(1) *f is holomorphic;*
(2) *fx is holomorphic for all $x \in X$;*
(3) *$\langle fx, y^*\rangle$ is holomorphic for all $x \in X$ and $y^* \in Z$.*

Proof. It is easy to see that the linear span of the functionals $x \otimes y^* \in (\mathscr{L}(X, Y))^*$, where x and y^* range over X and Z, is norming for $\mathscr{L}(X, Y)$. Now the equivalence of (1) and (3) follows from Proposition B.3.1. The equivalence with (2) is then evident. □

B.4 Complexification

The Cartesian product $X \times X$ of a real vector space X is a complex vector space with respect to the scalar multiplication $(a + bi)(x, y) := (ax - by, ay + bc)$. This complex vector space will be denoted by $X_{\mathbb{C}}$ and is called the *complexification* of X. Elements of $X_{\mathbb{C}}$ will be denoted by $x + iy$ rather than by (x, y).

If X is a real Banach space, the problem poses itself of finding norms on $X_{\mathbb{C}}$ which extend the given norm of X.

Proposition B.4.1. *If X is a real Banach space, then*

$$\|x+iy\|_{X_{\mathbb{C}}} := \sup_{\|x^*\|\leqslant 1} (\langle x,x^*\rangle^2 + \langle y,x^*\rangle^2)^{1/2}$$

defines a norm on $X_{\mathbb{C}}$ which extends the norm of X. With respect to this norm $X_{\mathbb{C}}$ is a complex Banach space.

Proof. $\langle x,x^*\rangle = \langle y,x^*\rangle = 0$ for all $x^*\in X^*$ and therefore $x=y=0$. Next,

$$\begin{aligned}\|(a+bi)(x+iy)\|_{X_{\mathbb{C}}}^2 &= \sup_{\|x^*\|\leqslant 1} \big(\langle ax-by,x^*\rangle^2 + \langle ay+bx,x^*\rangle^2\big)\\ &= \sup_{\|x^*\|\leqslant 1} \big(\langle ax,x^*\rangle^2 + \langle by,x^*\rangle^2 + \langle ay,x^*\rangle^2 + \langle bx,x^*\rangle^2\big)\\ &= (a^2+b^2) \sup_{\|x^*\|\leqslant 1} \big(\langle x,x^*\rangle^2 + \langle y,x^*\rangle^2\big)\\ &= (a^2+b^2)\|x+iy\|_{X_{\mathbb{C}}}^2.\end{aligned}$$

This proves the complex homogeneity of the norm. Finally, the triangle inequality with respect to this norm follows from the triangle inequality in $L^2(\Omega;X)$. ☐

We have a natural identification of vector spaces

$$(X^*)_{\mathbb{C}} = (X_{\mathbb{C}})^*$$

obtained associating to the vector $x^*+iy^*\in (X^*)_{\mathbb{C}}$ the mapping $\phi_{x^*+iy^*}\in (X_{\mathbb{C}})^*$ defined by

$$\phi_{x^*+iy^*}(x+iy) := (\langle x,x^*\rangle - \langle x,y^*\rangle) + i(\langle x,y^*\rangle + \langle y,x^*\rangle).$$

On the one hand, $(X^*)_{\mathbb{C}}$ can be turned into a complex Banach space by means of the procedure of Proposition B.4.1, i.e.,

$$\|x^*+iy^*\|_{(X^*)_{\mathbb{C}}} := \sup_{\|x^{**}\|\leqslant 1} (\langle x^*,x^{**}\rangle^2 + \langle y^*,x^{**}\rangle^2)^{1/2}.$$

On the other hand, $(X_{\mathbb{C}})^*$ is a complex Banach space with respect to the dual norm

$$\|x^*\|_{\mathbb{C}} := \sup_{\|x\|_{\mathbb{C}}\leqslant 1} |\langle x,x^*\rangle|.$$

Proposition B.4.2. *Under the above identification, we have an isomorphism of Banach spaces $(X^*)_{\mathbb{C}} = (X_{\mathbb{C}})^*$ and*

$$\|x^*+iy^*\|_{(X^*)_{\mathbb{C}}} \leqslant \|x^*+iy^*\|_{(X_{\mathbb{C}})^*} \leqslant 2\|x^*+iy^*\|_{(X^*)_{\mathbb{C}}}.$$

Proof. Since B_X is weak*-dense in $B_{X^{**}}$, we have

$$\begin{aligned}
\|x^* + iy^*\|_{(X^*)_\mathbb{C}} &= \sup_{\|x^{**}\|\leqslant 1} (\langle x^*, x^{**}\rangle^2 + \langle x^*, y^{**}\rangle^2)^{1/2} \\
&= \sup_{\|x\|\leqslant 1} (\langle x, x^*\rangle^2 + \langle x, y^*\rangle^2)^{1/2}. \\
&\leqslant \sup_{\|x'+iy'\|_{X_\mathbb{C}}\leqslant 1} ((\langle x', x^*\rangle - \langle y', y^*\rangle)^2 + (\langle x', y^*\rangle + \langle y', x^*\rangle)^2)^{1/2} \\
&= \sup_{\|x'+iy'\|_\mathbb{C}\leqslant 1} |\langle x' + iy', x^* + iy^*\rangle| = \|x^* + iy^*\|_{(X_\mathbb{C})^*}
\end{aligned}$$

since $\|x\| \leqslant 1$ implies $\|x\|_{X_\mathbb{C}} \leqslant 1$. This shows that

$$(X_\mathbb{C})^* \hookrightarrow (X^*)_\mathbb{C}$$

contractively. On the other hand, if $\|z^*\| = 1$, then

$$\begin{aligned}
\|z^*\|_{(X_\mathbb{C})^*} &= \sup_{\|x'+iy'\|_\mathbb{C}\leqslant 1} |\langle x, z^*\rangle + i\langle y, z^*\rangle| \\
&= \sup_{\|x'+iy'\|_\mathbb{C}\leqslant 1} (\langle x, z^*\rangle^2 + \langle y, z^*\rangle^2)^{1/2} \leqslant 1 = \|z^*\|,
\end{aligned}$$

where the last inequality follows from the fact that $\|x' + iy'\|_\mathbb{C} \leqslant 1$ is synonymous to the requirement that $\langle x, z^*\rangle^2 + \langle y, x^*\rangle^2 \leqslant 1$ whenever $\|z^*\| \leqslant 1$. Hence,

$$\|x^* + iy^*\|_{(X_\mathbb{C})^*} \leqslant \|x^*\| + \|y^*\| \leqslant 2\|x^* + iy^*\|_{(X^*)_\mathbb{C}},$$

where the last inequality follows from

$$\|x^* + iy^*\|^2_{(X^*)_\mathbb{C}} = \sup_{\|x^{**}\|\leqslant 1} (\langle x^*, x^{**}\rangle^2 + \langle y^*, x^{**}\rangle^2) \geqslant \sup_{\|x^{**}\|\leqslant 1} \langle x^*, x^{**}\rangle^2 = \|x^*\|^2$$

and similarly $\|x^* + iy^*\|^2_{(X^*)_\mathbb{C}} \geqslant \|y^*\|^2$. This shows that the identity mapping from $(X^*)_\mathbb{C}$ to $(X_\mathbb{C})^*$ has norm at most 2. $\square$

The fact that we do not get an isometry here can be understood in terms of tensor norms. We refer the reader to the Notes for an elaboration of this point.

Next we turn to the problem of complexifying bounded operators.

Proposition B.4.3. *Let T be a bounded linear operator on a real Banach space X. Then the operator*

$$T_\mathbb{C}(x + iy) := Tx + iTy$$

is complex linear and bounded on $X_\mathbb{C}$, of norm $\|T_\mathbb{C}\| = \|T\|$.

Proof. Additivity is evident, and regarding the complex scalar multiplication we have

$$T_{\mathbb{C}}((a+bi)(x+iy)) = aTx - bTy + i(aTy - bTx) = (a+bi)(T_{\mathbb{C}}(x+iy)).$$

It follows that T is complex linear. Moreover, for any $x+iy \in X_{\mathbb{C}}$,

$$\begin{aligned}\|T_C(x+iy)\|_{X_{\mathbb{C}}} &= \sup_{\|x^*\|\leqslant 1} (\langle x, T^*x^*\rangle^2 + \langle y, T^*x^*\rangle^2)^{1/2} \\ &\leqslant \|T\| \sup_{\|y^*\|\leqslant 1} (\langle x, y^*\rangle^2 + \langle y, x^*\rangle^2)^{1/2} = \|T\|\|x+iy\|_{X_{\mathbb{C}}}.\end{aligned}$$

This shows that $T_{\mathbb{C}}$ is bounded on $X_{\mathbb{C}}$ of norm $\|T_{\mathbb{C}}\| \leqslant \|T\|$. The opposite inequality $\|T_{\mathbb{C}}\| \geqslant \|T\|$ is clear by considering vectors $x \in X$. ◻

B.5 Notes

The theory of Banach spaces was initiated by Banach [1932], who obtained most of the classical results at least in the separable case. Several beautiful monographs have been devoted to this subject, among which Albiac and Kalton [2006], Li and Queffélec [2004], Wojtaszczyk [1991] are closest in spirit to the present volume.

A comprehensive history on Banach spaces and their linear operators has been written by Pietsch [2007]; for histories of functional analysis in a broader perspective see Dieudonné [1981] and Monna [1973].

Section B.1

The Hahn–Banach theorem goes back to Helly [1912]. For real Banach spaces it was first proved by Hahn [1927] and Banach [1932]; the complex version was found only several years later, independently by Murray [1936], Bohnenblust and Sobczyk [1938] and Soukhomlinoff [1938]. Modern proofs of the Hahn–Banach theorems can be found in every textbook on functional analysis. Standard texts include Brezis [2011], Conway [1990], Rudin [1991], Yosida [1980] and the German text Werner [2000].

The Banach–Alaoglu theorem goes back to Helly [1912], who essentially treated the case $X = C[a,b]$. For separable Banach spaces, the theorem appears in Banach [1932]. The extension to arbitrary Banach spaces is due to Alaoglu [1940]. For a discussion of the many contributions by other authors we refer to Pietsch [2007].

Corollary B.1.14 is taken from Brzeźniak and Van Neerven [2000] but might well have been known before. The Krein–Šmulian theorem was proved in Kreĭn and Šmulian [1940]. Our proof is taken from Dunford and Schwartz [1958]. The basic lemma B.1.15 holds true is a much more general locally convex setting, where it is known as the Banach–Dieudonné theorem; see for instance Kelley and Namioka [1976]. For this reason some authors refer to the Krein–Šmulian theorem under this name.

Section B.2

The material of this section is standard and goes back to Banach [1932]. Proofs can be found in most textbooks on functional analysis or operator theory.

Section B.3

In its form stated here, Proposition B.3.1 seems to be due to Grosse-Erdmann [1993]. A shorter proof was subsequently given in Arendt and Nikolski [2000, 2006], where it is shown that the local boundedness assumption cannot be removed if Z does not induce an equivalent norm in X. The special case of Proposition B.3.1 for $Z = X^*$ is stated in Kato [1995].

Section B.4

Complexifications of Banach spaces are discussed extensively in Ruston [1986] and Muñoz, Sarantopoulos, and Tonge [1999].

The complex norm in the main text has the equivalent representation

$$\|x + iy\|_{X_{\mathbb{C}}} = \sup_{0 \leqslant \theta \leqslant 2\pi} \|x \cos\theta + y \sin\theta\|.$$

In this form the norm seems to have first been used by Taylor [1938]. To prove the equivalence, note that the right-hand side is equal to

$$\begin{aligned}
&\sup_{\|x^*\| \leqslant 1} \sup_{0 \leqslant \theta \leqslant 2\pi} |\langle x\cos\theta + y\sin\theta, x^*\rangle| \\
&\qquad = \sup_{\|x^*\| \leqslant 1} \sup_{a^2+b^2=1} |a\langle x, x^*\rangle + b\langle y, x^*\rangle| \\
&\qquad = \sup_{\|x^*\| \leqslant 1} (\langle x, x^*\rangle^2 + \langle y, x^*\rangle^2)^{1/2} = \|x + iy\|_{\mathbb{C}}.
\end{aligned}$$

In Banach lattices, the formula

$$\|x + iy\| := \| (|x|^2 + |y|^2)^{1/2} \|$$

gives the 'correct' complexification of $L^p(\mu)$ and $C(K)$ in the sense that this formula reproduces the natural complex norms on $L^p(\mu; \mathbb{C})$ and $C(K; \mathbb{C})$. For $C(K)$ this norm is the same as the one introduced in the text. This can be understood once we realise that $X_{\mathbb{C}} = X \otimes_\varepsilon \ell_2^2$ isometrically, where $\otimes_\varepsilon$ stands for the injective tensor norm (see Diestel and Uhl [1977]). From this we see that $X_{*\mathbb{C}} = X \otimes_\pi \ell_2^2$ isometrically, where $\otimes_\pi$ stands for the projective tensor norm. In particular, $(L^1(\mu))_{*\mathbb{C}} = L^1(\mu;\mathbb{C})$ isometrically. This result was extended to abstract injective and projective tensor products by van Van Zyl [2008]. Further result on complexifications of Banach lattices can be found in Lotz [1968], Luxemburg and Zaanen [1971], Mittelmeyer and Wolff [1974], Van Neerven [1997], Schaefer [1974], and De Schipper [1973].

The question whether every complex Banach space arises (possibly up to an equivalent norm) as the complexification of a real Banach spaces is an interesting one. Let us start the discussion of this problem with the trivial observation that the answer is affirmative for Hilbert spaces. To see this, let H be a complex Hilbert space, and fix a maximal orthonormal system in H. Then the real-linear subspace consisting of those $h \in H$ whose expansion with respect to this system has all coordinates real has the structure of a real Hilbert space and its complexification is isometric to H.

Quite surprisingly, not every complex Banach space is the complexification of an underlying real Banach space. Before turning to this problem, let us first look at the related question in the opposite direction whether every real Banach space can be endowed with a complex structure. A necessary condition for this to be possible is easy to state: if X is a complex Banach space and $X_\mathbb{R}$ denotes the real Banach space obtained from X by restricting the complex scalar multiplication to the reals, then the mapping $x \mapsto ix$ induces an isometry R on $X_\mathbb{R}$ which satisfies $R^2 = -I$.

Conversely, if a real Banach space X admits an isometry R satisfying $R^2 = -I$, then one may extend the scalar multiplication on X to the complex scalars by putting $(a + bi)x := ax + bRy$. It is a simple matter to verify that this extension turns X into a complex vector space, and with respect to the equivalent norm

$$|||x||| := \sup_{\theta\in[0,2\pi]} \|e^{i\theta}x\|$$

X becomes a complex Banach space.

These simple observations are due to Dieudonné [1952], who proceeded to observe that there exist real Banach spaces that cannot be endowed with complex structure. In fact, the famous James space, introduced in James [1950], is a real Banach space J with the curious property that $\dim(J^{**}/J) = 1$. If this space had complex structure, it would turn J^{**}/J into a complex space of real dimension one, which is absurd.

Let us now turn to the problem whether any complex Banach space X is the complexification of a real Banach space. This amounts to the question whether $X_\mathbb{R}$ is of the form $Y \oplus Y$ for some real Banach space Y. Indeed, once we know this, $R(y_1, y_2) := (-y_2, y_1)$ is an isomorphism of $X_\mathbb{R}$ which satisfies $R^2 = -1$ and we find that $X = Y_\mathbb{C}$ up to an equivalent norm on X. Thus we have proved:

Proposition B.5.1. *A complex Banach space X is isomorphic to the complexification of the real Banach space Y if and only if $X_{\mathbb{R}}$ and $Y \oplus Y$ are isomorphic as real Banach spaces.*

Representing $x \in X = Y_{\mathbb{C}}$ as $x = y_1 + iy_2$ with $y_1, y_2 \in Y$ and noting that the 'conjugation' $(y_1, y_2) \mapsto (y_1, -y_2)$ is an isomorphism on $Y \oplus Y$, we see that a necessary condition in order that X be isomorphic to the complexification of a real Banach space Y is that X is isomorphic to its own conjugate $\bar{X}$, which, by definition, is the Banach space obtained by endowing X with the new complex multiplication $(c, x) \mapsto \bar{c}x$ (where the given scalar multiplication of X is used in the right-hand side). The first proof that there exist complex Banach spaces which are not isomorphic to their own conjugate is due to Bourgain [1986a], who employed a non-constructive probabilistic argument. A more direct approach using twisted sums of Hilbert spaces was subsequently invented by Kalton [1995].

C

Interpolation theory

This appendix presents a concise introduction to the theory of interpolation spaces.

C.1 Interpolation couples

Definition C.1.1. *An* interpolation couple *is an ordered pair* (X_0, X_1) *of Banach spaces, both of which are continuously embedded in a Hausdorff topological vector space* $\mathscr{X}$.

Example C.1.2. Let $(S, \mathscr{A}, \mu)$ be a measure space. For all $1 \leqslant p_0, p_1 \leqslant \infty$ the pair $(L^{p_0}(S), L^{p_1}(S))$ is an interpolation couple with respect to the embeddings of these spaces into $\mathscr{X} = L^0(S)$. Here on $L^0(S)$ we define a translation invariant metric by

$$d_\infty(f,g) := \inf\{a \geqslant 0 : \mu(|f-g| \geqslant a) \leqslant a\} \wedge 1.$$

If the measure space is σ-finite we can also use the topology of convergence in measure.

Proposition C.1.3. *If* (X_0, X_1) *is an interpolation couple, then the spaces*

$$\begin{aligned} X_0 \cap X_1 &= \{x \in \mathscr{X} : \ x \in X_0 \ \textit{and} \ x \in X_1\}, \\ X_0 + X_1 &= \{x \in \mathscr{X} : \ x = x_0 + x_1 \ \textit{with} \ x_0 \in X_0, \ x_1 \in X_1\} \end{aligned}$$

are Banach spaces with respect to the norms

$$\begin{aligned} \|x\|_{X_0 \cap X_1} &:= \max\{\|x\|_{X_0}, \|x\|_{X_1}\}, \\ \|x\|_{X_0 + X_1} &:= \inf\{\|x_0\|_{X_0} + \|x_1\|_{X_1} : \ x = x_0 + x_1, \ x_0 \in X_0, \ x_1 \in X_1\}. \end{aligned}$$

Proof. Clearly, $\|\cdot\|_{X_0 \cap X_1}$ is a norm. To check completeness let $(x_n)_{n \geqslant 1}$ be a Cauchy sequence in $X_0 \cap X_1$. Then $(x_n)_{n \geqslant 1}$ is a Cauchy sequence in both

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 63, DOI 10.1007/978-3-319-48520-1

X_0 and X_1. By the continuity of the embeddings of these spaces and the fact that $\mathscr{X}$ is Hausdorff, the limits in X_0 and X_1 of $(x_n)_{n\geqslant 1}$ agree as elements of $\mathscr{X}$. Denoting this common limit by x we have $\lim_{n\to\infty} x_n = x$ in both X_0 and X_1, hence also in $X_0 \cap X_1$.

To show that $\|\cdot\|_{X_0+X_1}$ is a norm, we only check that $\|x\|_{X_0+X_1} = 0$ implies $x = 0$, the remaining defining properties of a norm being evident. Choose vectors $y_n^0 \in X_0$ and $y_n^1 \in X_1$ such that $x = y_n^0 + y_n^1$ and $\|y_n^0\|_{X_0} + \|y_n^1\|_{X_1} \to 0$ as $n \to \infty$. Then $x = y_n^0 + y_n^1 \to 0$ in $\mathscr{X}$, so $x = 0$.

To prove completeness of $X_0 + X_1$, we use that completeness is equivalent to the convergence of every absolutely convergent series. Suppose therefore that $\sum_{n\geqslant 1} \|x_n\|_{X_0+X_1} < \infty$. Write $x_n = y_n^0 + y_n^1$ with $y_n^0 \in X_0$, $y_n^1 \in X_1$, and $\|y_n^0\|_{X_0} + \|y_n^1\|_{X_1} < \|x_n\|_{X_0+X_1} + 2^{-n}$. The sums $\sum_{j=1}^\infty y_n^0$ and $\sum_{j=1}^\infty y_n^1$ converge absolutely in X_0 and X_1, respectively, say with sums y^0 and y^1. Set $x := y^0 + y^1$. From $x_n = y_n^0 + y_n^1$ we see that

$$\Big\| x - \sum_{n=1}^N x_n \Big\|_{X_0+X_1} = \Big\| \sum_{n=N+1}^\infty (y_n^0 + y_n^1) \Big\|_{X_0+X_1} \leqslant 2^{-N} + \sum_{n=N+1}^\infty \|x_n\|_{X_0+X_1},$$

which tends to 0 as $N \to \infty$. This proves that $\sum_{n=1}^\infty x_n = x$ in $X_0 + X_1$. □

Observe that $X_0 \cap X_1$ and $X_0 + X_1$ are independent of the particular choice of ambient space $\mathscr{X}$.

It is immediate from the definitions that we have contractive embeddings $X_0 \cap X_1 \hookrightarrow X_0$ and $X_0 \cap X_1 \hookrightarrow X_1$. From the trivial decompositions

$$x_0 = x_0 + 0, \qquad x_1 = 0 + x_1$$

one sees that also $X_0 \hookrightarrow X_0 + X_1$ and $X_1 \hookrightarrow X_0 + X_1$ contractively.

Example C.1.4. If (X_0, X_1) is an interpolation couple with $X_0 \cap X_1$ dense in both X_0 and X_1, then the pair (X_0^*, X_1^*) is an interpolation couple with respect to $\mathscr{X} = (X_0 \cap X_1)^*$.

C.2 Complex interpolation

For a topological space T we denote y $C_{\mathrm{b}}(T; X)$ the Banach space of all bounded continuous functions $f : T \to X$ with respect to the supremum norm.

In what follows we fix an interpolation couple (X_0, X_1) of complex Banach spaces. We shall use the theory of vector-valued holomorphic functions (see Section B.3) to define a scale of complex Banach spaces $[X_0, X_1]_\theta$, $0 < \theta < 1$, with continuous inclusions

$$X_0 \cap X_1 \hookrightarrow [X_0, X_1]_\theta \hookrightarrow X_0 + X_1.$$

Consider the open strip

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

and let $\overline{\mathbb{S}}$ be its closure.

Definition C.2.1. *We denote by $\mathscr{H}(X_0, X_1)$ the complex vector space of all continuous functions $f : \overline{\mathbb{S}} \to X_0 + X_1$ with the following properties:*

(i) *f is holomorphic as an $(X_0 + X_1)$-valued function on $\mathbb{S}$;*
(ii) *the function $v \mapsto f(iv)$ belongs to $C_{\rm b}(\mathbb{R}; X_0)$;*
(iii) *the function $v \mapsto f(1 + iv)$ belongs to $C_{\rm b}(\mathbb{R}; X_1)$.*

The space $\mathscr{H}_0(X_0, X_1)$ is defined as the subspace of $\mathscr{H}(X_0, X_1)$ for which

$$\lim_{|v|\to\infty} \|f(iv)\|_{X_0} = \lim_{|v|\to\infty} \|f(1 + iv)\|_{X_1} = 0.$$

Proposition C.2.2. *$\mathscr{H}(X_0, X_1)$ is a Banach space with respect to the norm*

$$\|f\|_{\mathscr{H}(X_0,X_1)} := \max\Big\{ \sup_{v\in\mathbb{R}} \|f(iv)\|_{X_0}, \ \sup_{v\in\mathbb{R}} \|f(1 + iv)\|_{X_1} \Big\}$$

and $\mathscr{H}_0(X_0, X_1)$ is a closed subspace of $\mathscr{H}(X_0, X_1)$.

Proof. It suffices to prove the first assertion, the second being an immediate consequence of the definition of the norm.

If $(f_n)_{n\geqslant 1}$ is a Cauchy sequence in $\mathscr{H}(X_0, X_1)$, then $\lim_{n\to\infty} f_n(i\cdot)$ and $\lim_{n\to\infty} f_n(1+i\cdot)$ exist in $C_{\rm b}(\mathbb{R}; X_0)$ and $C_{\rm b}(\mathbb{R}; X_1)$, respectively. In particular, both limits exist in $C_{\rm b}(\mathbb{R}; X_0 + X_1)$. By the three lines lemma, the sequence $(f_n)_{n\geqslant 1}$ is Cauchy in $C_{\rm b}(\overline{\mathbb{S}}; X_0 + X_1)$ and therefore converges to a limit $f \in C_{\rm b}(\overline{\mathbb{S}}; X_0 + X_1)$. Its restriction to $\mathbb{S}$ is holomorphic, it being the uniform limit of the holomorphic functions $f_n : \mathbb{S} \to X_0 + X_1$. □

Definition C.2.3. *For $0 < \theta < 1$ the* complex interpolation space *$[X_0, X_1]_\theta$ is defined as the complex vector space of all $x \in X_0 + X_1$ such that $f(\theta) = x$ for some $f \in \mathscr{H}(X_0, X_1)$.*

Proposition C.2.4. *Let $0 < \theta < 1$. An element $x \in X_0 + X_1$ belongs to $[X_0, X_1]_\theta$ if and only if $f(\theta) = x$ for some $f \in \mathscr{H}_0(X_0, X_1)$. Furthermore, $[X_0, X_1]_\theta$ is a Banach space with respect to the norm*

$$\begin{aligned}\|x\|_{[X_0,X_1]_\theta} &:= \inf\big\{\|f\|_{\mathscr{H}(X_0,X_1)} : \ f(\theta) = x\big\} \\ &= \inf\big\{\|f\|_{\mathscr{H}_0(X_0,X_1)} : \ f(\theta) = x\big\}.\end{aligned}$$

Proof. If $x = f(\theta)$ with $f \in \mathscr{H}(X_0, X_1)$, then for $\delta > 0$ we have $x = f_\delta(\theta)$ with $f_\delta \in \mathscr{H}_0(X_0, X_1)$ given by $f_\delta(z) = \exp(\delta(z^2 - \theta^2))f(z)$. Moreover,

$$\lim_{\delta\downarrow 0} \|f_\delta\|_{\mathscr{H}(X_0,X_1)} = \|f\|_{\mathscr{H}(X_0,X_1)}.$$

It remains to be proved the completeness of $\mathscr{H}(X_0, X_1)$, for which it suffices to prove that every absolutely convergent series is convergent. Suppose that $\sum_{n\geqslant 1} \|x_n\|_{[X_0,X_1]_\theta} < \infty$. Choose functions $g_n \in [X_0, X_1]_\theta$ such that $g_n(\theta) = x_n$ and $\|g_n\|_{\mathscr{H}(X_0,X_1)} \leqslant \|x_n\|_{[X_0,X_1]_\theta} + 2^{-n}$. The sum $\sum_{n=1}^\infty g_n$ converges absolutely in $\mathscr{H}(X_0, X_1)$, say to the function g. Put $x := g(\theta)$. Then

$$\Big\|x - \sum_{n=1}^N x_n\Big\|_{[X_0,X_1]_\theta} \leqslant \Big\| \sum_{n=N+1}^\infty g_n\Big\|_{\mathscr{H}(X_0,X_1)} \leqslant 2^{-N} + \sum_{n=N+1}^\infty \|x_n\|_{[X_0,X_1]_\theta},$$

which tends to 0 as $N \to \infty$. This proves that $x = \sum_{n\geqslant 1} x_n$ in $[X_0, X_1]_\theta$. $\square$

It is immediate from the definitions that

$$[X, X]_\theta = X$$

and

$$[X_0, X_1]_\theta = [X_1, X_0]_{1-\theta}$$

isometrically. We leave the easy proofs to the reader.

Lemma C.2.5. *For all* $0 < \theta < 1$ *we have contractive embeddings*

$$X_0 \cap X_1 \hookrightarrow [X_0, X_1]_\theta \hookrightarrow X_0 + X_1.$$

Proof. For any $x \in X_0 \cap X_1$ the constant function $f_x :\equiv x$ is in $\mathscr{H}(X_0, X_1)$ and satisfies $f_x(\theta) = x$. Moreover,

$$\|x\|_{[X_0,X_1]_\theta} \leqslant \|f_x\|_{\mathscr{H}(X_0,X_1)} \leqslant \max\{\|x\|_{X_0}, \|x_1\|_{X_1}\} = \|x\|_{X_0\cap X_1}.$$

Suppose next that $x \in [X_0, X_1]_\theta$ and let $f \in \mathscr{H}(X_0, X_1)$ be such that $f(\theta) = x$. Then $x \in X_0 + X_1$, and by the maximum principle,

$$\begin{aligned}\|x\|_{X_0+X_1} = \|f(\theta)\|_{X_0+X_1} &\leqslant \max\Big\{\sup_{v\in\mathbb{R}} \|f(iv)\|_{X_0+X_1}, \sup_{v\in\mathbb{R}} \|f(1 + iv)\|_{X_0+X_1}\Big\} \\ &\leqslant \max\Big\{\sup_{v\in\mathbb{R}} \|f(iv)\|_{X_0}, \sup_{v\in\mathbb{R}} \|f(1 + iv)\|_{X_1}\Big\} \\ &= \|f\|_{\mathscr{H}(X_0,X_1)}.\end{aligned}$$

Taking the infimum over all admissible f gives the inequality $\|x\|_{X_0+X_1} \leqslant \|x\|_{[X_0,X_1]_\theta}$. $\square$

The interest of the spaces $[X_0, X_1]_\theta$ is explained by the next result.

Theorem C.2.6 (Complex interpolation of operators). *Let* (X_0, X_1) *and* (Y_0, Y_1) *be interpolation couples. Suppose* $T : X_0 + X_1 \to Y_0 + Y_1$ *is a linear operator which maps* X_0 *into* Y_0 *and* X_1 *into* Y_1 *with norms*

$$\|T\|_{\mathscr{L}(X_0,Y_0)} = A_0, \qquad \|T\|_{\mathscr{L}(X_1,Y_1)} = A_1.$$

Then for each $0 < \theta < 1$ *the operator* T *maps* $[X_0, X_1]_\theta$ *into* $[Y_0, Y_1]_\theta$ *and we have*

$$\|T\|_{\mathscr{L}([X_0,X_1]_\theta,[Y_0,Y_1]_\theta)} \leqslant A_0^{1-\theta} A_1^\theta.$$

Proof. The assumptions imply that T is bounded as an operator from X_0+X_1 to Y_0+Y_1 of norm at most $\max\{A_0, A_1\}$.

First assume that $T \neq 0$ both in $\mathscr{L}(X_0,Y_0)$ and $\mathscr{L}(X_1,Y_1)$. Fix $x \in [X_0,X_1]_\theta$ and let $f \in \mathscr{H}(X_0,X_1)$ satisfy $f(\theta) = x$. Define the function $g : \overline{\mathbb{S}} \to Y_0+Y_1$ by

$$g(z) := \Big(\frac{A_0}{A_1}\Big)^{z-\theta} Tf(z), \quad z \in \overline{\mathbb{S}}.$$

Clearly g is continuous on $\overline{\mathbb{S}}$, holomorphic on $\mathbb{S}$, $g(\theta) = Tf(\theta) = Tx$, and

$$\|g(iv)\|_{Y_0} \leqslant A_0^{1-\theta}A_1^{\theta}\|f(iv)\|_{X_0},$$
$$\|g(1+iv)\|_{Y_1} \leqslant A_0^{1-\theta}A_1^{\theta}\|f(1+iv)\|_{X_1}.$$

It follows that $g \in \mathscr{H}(Y_0,Y_1)$ and $\|g\|_{\mathscr{H}(Y_0,Y_1)} \leqslant A_0^{1-\theta}A_1^{\theta}\|f\|_{\mathscr{H}(X_0,X_1)}$. Hence, $Tx \in [Y_0,Y_1]_\theta$ and $\|Tx\|_{[Y_0,Y_1]_\theta} \leqslant A_0^{1-\theta}A_1^{\theta}\|f\|_{\mathscr{H}(X_0,X_1)}$. Now the result follows by taking the infimum over all admissible f.

If $A_0 = 0$, the choice

$$g_\varepsilon(z) = \varepsilon^{z-\theta}Tf(z)$$

gives the estimates

$$\|g(iv)\|_{Y_0} = 0,$$
$$\|g(1+iv)\|_{Y_1} \leqslant \varepsilon^{1-\theta}A_1\|f(1+iv)\|_{X_1},$$

and consequently $\|g\|_{\mathscr{H}(Y_0,Y_1)} \leqslant \varepsilon^{1-\theta}A_1\|f\|_{\mathscr{H}(X_0,X_1)}$ and $\|Tx\|_{[Y_0,Y_1]_\theta} \leqslant \varepsilon^{1-\theta}A_1\|f\|_{\mathscr{H}(X_0,X_1)}$. Letting $\varepsilon \downarrow 0$ gives $\|Tx\|_{[Y_0,Y_1]_\theta} = 0$.

If $A_1 = 0$ we reverse the roles of the indices $j = 0,1$ and replace θ by $1-\theta$. Since $[Y_0,Y_1]_\theta = [Y_1,Y_0]_{1-\theta}$ isometrically, it then follows that $\|Tx\|_{[Y_0,Y_1]_\theta} = \|Tx\|_{[Y_1,Y_0]_{1-\theta}} = 0$. $\square$

As a simple application, fix an $x \in X_0 \cap X_1$ and consider the mapping $c \mapsto cx$ as operators from $\mathbb{K}$ into X_0 and into X_1, respectively. Interpolation of these operators gives the estimate

$$\|x\|_{[X_0,X_1]_\theta} \leqslant \|x\|_{X_0}^{1-\theta}\|x\|_{X_1}^{\theta}, \qquad x \in X_0 \cap X_1. \tag{C.1}$$

We now introduce the space

$$\mathscr{H}_0(X_0,X_1;X_0\cap X_1)$$

consisting of all $f \in \mathscr{H}_0(X_0,X_1)$ such that $f(z) \in X_0 \cap X_1$ for all $z \in \mathbb{S}$ and $z \mapsto f(z)$ is holomorphic as a function from $\mathbb{S}$ into $X_0 \cap X_1$.

Lemma C.2.7. *$\mathscr{H}_0(X_0,X_1;X_0\cap X_1)$ is dense in $\mathscr{H}_0(X_0,X_1)$.*

Proof. Fix an arbitrary $f \in \mathscr{H}_0(X_0, X_1)$ and let $f_\delta(z) := e^{\delta z^2} f(z)$. By the decay properties of f one has $\|f - f_\delta\|_{\mathscr{H}(X_0,X_1)} \to 0$ as $\delta \to 0$.

Fix $\delta > 0$ and write $g = f_\delta$. It remains to be shown that there exists a function $h \in \mathscr{H}_0(X_0, X_1; X_0 \cap X_1)$ such that

$$\|g - h\|_{\mathscr{H}(X_0,X_1)} < \varepsilon. \tag{C.2}$$

In view of the estimate $|\exp(\delta z^2)| \leqslant M \exp(-\delta|\Im z|^2)$, where M is a constant independent of $z \in \overline{\mathbb{S}}$, for every integer $n \geqslant 1$ the sum

$$G_n(z) := \sum_{m\in\mathbb{Z}} g(z + imn)$$

converges absolutely in $X_0 + X_1$, locally uniformly on $\overline{\mathbb{S}}$. The resulting continuous function $G_n : \overline{\mathbb{S}} \to X_0 + X_1$ is holomorphic in $\mathbb{S}$, its restrictions to $i\mathbb{R}$ and $1 + i\mathbb{R}$ belong to $C_{\rm b}(\mathbb{R}; X_0)$ and $C_{\rm b}(\mathbb{R}; X_1)$ respectively, and G_n satisfies the periodicity relation

$$G_n(z) = G_n(z + in).$$

Another consequence of the rapid decay of g is that we can find $r > 0$ and $n \geqslant 1$ such that

$$\|z \mapsto \exp(rz^2)G_n(z) - g(z)\|_{\mathscr{H}(X_0,X_1)} < \varepsilon.$$

We sketch the steps and leave the somewhat tedious details to the reader. First, $|g(u + iv)|$ is small for $v \notin [-R, R]$ for R large enough. Then, both $|\exp(r(iv)^2) - 1|$ and $|\exp(r(1 + iv)^2) - 1|$ are small for $v \in [-R, R]$ if we take $r > 0$ small enough. The functions $|\exp(r(iv)^2)|$ and $|\exp(r(1 + iv)^2)|$ are small outside some large enough interval $[-R', R']$ which may be assumed to contain $[-R, R]$. By taking n large enough, we may simultaneously achieve that $|G_n(iv) - g(iv)|$ and $|G_n(1+iv) - g(1+iv)|$ are small on $[-R', R']$. Then $|G_n(iv)|$ and $|G_n(1+iv)|$ are small on $[-R', -R]$ and $[R, R']$, using that $|g(u+iv)|$ is small for $v \in [-R', -R] \cup [R, R']$. Then also $|\exp(r(iv)^2)G_n(iv) - g(iv)|$ and $|\exp(r(1+iv)^2)G_n(1+iv) - g(1+iv)|$ are small on $[-R, R]$, on $[-R', -R]$ and $[R, R']$, and outside $[-R', R']$.

It therefore suffices to approximate the function $z \mapsto \exp(rz^2)G_n(z)$ in $\mathscr{H}(X_0, X_1)$ by functions in $\mathscr{H}_0(X_0, X_1; X_0 \cap X_1)$.

For every positive integer m let $\Gamma_{m,n}$ denote the counterclockwise oriented boundary of the rectangle $0 \leqslant \Re z \leqslant 1$, $0 \leqslant |\Im z| \leqslant mn$. By Cauchy's integral theorem, for all $k \in \mathbb{Z}$ we have

$$\int_{\Gamma_{m,n}} \exp(2\pi kz/n)G_n(z)\,\mathrm{d}z = 0.$$

Therefore, given any $\delta > 0$, for all large enough m we have

$$\frac{1}{m}\Big\| \int_{-mn}^{mn} e^{-\pi ivk/n} G_n(iv)\,\mathrm{d}v - \int_{-mn}^{mn} e^{-\pi(1+iv)k/n} G_n(1 + iv)\,\mathrm{d}v\Big\|_{X_0+X_1} < \delta$$

since the contributions along the horizontal segments are uniformly bounded and get multiplied by the factor $1/m$. But then the periodicity of G_n implies

$$\Big\| \int_{-n}^{n} e^{-\pi i v k/n} G_n(iv)\,\mathrm{d}v - \int_{-n}^{n} e^{-\pi(1+iv)k/n} G_n(1+iv)\,\mathrm{d}v \Big\| < \delta.$$

Since $\delta > 0$ was arbitrary, this proves the identity

$$\int_{-n}^{n} e^{-\pi i v k/n} G_n(iv)\,\mathrm{d}v = \int_{-n}^{n} e^{-\pi(1+iv)k/n} G_n(1+iv)\,\mathrm{d}v.$$

The integral on the left is in X_0, while the integral on the right is in X_1. It follows that both integrals are in $X_0 \cap X_1$. Moreover, they can be interpreted as Fourier integrals on the interval $[-n,n]$, and as such the above relation shows that both the Fourier coefficients of $v \mapsto G_n(iv)$ and $v \mapsto G_n(1+iv)$ belong to $X_0 \cap X_1$. These functions can be uniformly approximated, in $C([-n,n];X_0)$ and $C([-n,n];X_1)$, by the Césaro means of their Fourier series; these approximations belong to $C([-n,n];X_0 \cap X_1)$. This defines corresponding Césaro approximations for $z \mapsto G_n(z)$ on the rectangle $0 \leqslant \Re z \leqslant 1$, $|\Im z| \leqslant n$. Extending these approximations by periodicity to $\overline{\mathbb{S}}$, and multiplying with $\exp(rz^2)$, we obtain the desired approximation for $z \mapsto \exp(rz^2)G_n(z)$. $\square$

Corollary C.2.8. *For all $0 < \theta < 1$, $x \in [X_0, X_1]_\theta$, and $\varepsilon > 0$, there exists a function $h \in \mathscr{H}_0(X_0, X_1; X_0 \cap X_1)$ such that*

$$\|h(\theta) - x\|_{[X_0,X_1]_\theta} \leqslant \varepsilon \|x\|_{[X_0,X_1]_\theta}, \quad \|h\|_{\mathscr{H}(X_0,X_1)} \leqslant (1+2\varepsilon)\|x\|_{[X_0,X_1]_\theta}.$$

In particular, $X_0 \cap X_1$ is dense in $[X_0, X_1]_\theta$.

Proof. By Proposition C.2.4 we can find $f \in \mathscr{H}_0(X_0, X_1)$ such that $f(\theta) = x$ and $\|f\|_{\mathscr{H}(X_0,X_1)} \leqslant (1+\varepsilon)\|x\|_{[X_0,X_1]_\theta}$. Using Lemma C.2.7 we can find $h \in \mathscr{H}_0(X_0, X_1; X_0 \cap X_1)$ such that $\|h - f\|_{\mathscr{H}(X_0,X_1)} \leqslant \varepsilon \|x\|_{[X_0,X_1]_\theta}$. Clearly, the required bound for the norm of h follows by combining the estimates. Moreover,

$$\|h(\theta) - x\|_{[X_0,X_1]_\theta} \leqslant \|h - f\|_{\mathscr{H}(X_0,X_1)} \leqslant \varepsilon \|x\|_{[X_0,X_1]_\theta}.$$

$\square$

In the proof of Theorem C.4.1 we will need a somewhat technical result (Corollary C.2.11 below), the proof of which requires some preparation.

The Poisson kernel for the strip $\mathbb{S}$ is given by

$$P_j(u+iv;t) = \frac{\sin(\pi u)\exp(\pi(v-t))}{\sin^2(\pi u) + (\cos(\pi u) - (-1)^j \exp(\pi(v-t)))^2} \quad (j = 0,1).$$

This can be derived by routine computation from the formula for the Poisson kernel for the unit disc by means of the conformal mapping $z \mapsto (\exp(\pi i z) -$

$i)/(\exp(\pi iz)+i)$. As a consequence, every bounded harmonic function $\phi : \mathbb{S} \to \mathbb{R}$ which is continuous on $\overline{\mathbb{S}}$ is given as

$$\phi(u+iv) = \sum_{j=0,1} \int_{-\infty}^{\infty} \phi(j+it) P_j(u+iv;t)\,\mathrm{d}t, \qquad u+iv \in \mathbb{S}. \tag{C.3}$$

From this identity we obtain the following representation for holomorphic functions.

Lemma C.2.9 (Poisson formula for the strip). *Suppose $f : \overline{\mathbb{S}} \to X$ is a bounded continuous function, holomorphic on $\mathbb{S}$. Then for all $z \in \mathbb{S}$ we have $f(z) = f_0(z) + f_1(z)$, where*

$$f_j(z) = \int_{-\infty}^{\infty} f(j+it) P_j(z;t)\,\mathrm{d}t \qquad (j=0,1).$$

Applying (C.3) to the harmonic functions $\phi(u+iv) = 1-u$ and $\phi(u+iv) = u$ and evaluating at θ, we obtain the identities

$$\int_{-\infty}^{\infty} P_0(\theta;t)\,\mathrm{d}t = 1-\theta, \quad \int_{-\infty}^{\infty} P_1(\theta;t)\,\mathrm{d}t = \theta, \qquad 0<\theta<1. \tag{C.4}$$

Lemma C.2.10. *For all $0<\theta<1$ and $f \in \mathscr{H}(X_0,X_1)$,*

(1) $\log \|f(\theta)\|_{[X_0,X_1]_\theta} \leqslant \displaystyle\sum_{j=0,1} \int_{-\infty}^{\infty} \log \|f(j+it)\|_{X_j} P_j(\theta;t)\,\mathrm{d}t;$

(2) $$\|f(\theta)\|_{[X_0,X_1]_\theta} \leqslant \Big[\frac{1}{1-\theta}\int_{-\infty}^{\infty} \|f(it)\|_{X_0} P_0(\theta,t)\,\mathrm{d}t\Big]^{1-\theta} \times \Big[\frac{1}{\theta}\int_{-\infty}^{\infty} \|f(1+it)\|_{X_1} P_1(\theta,t)\,\mathrm{d}t\Big]^{\theta}.$$

Proof. Let $\phi_j : \mathbb{R} \to \mathbb{R}$ be bounded smooth functions which satisfy

$$\phi_0(v) \geqslant \log \|f(iv)\|_{X_0}, \quad \phi_1(v) \geqslant \log \|f(1+iv)\|_{X_1}; \qquad v \in \mathbb{R}. \tag{C.5}$$

Let $\varPhi : \overline{\mathbb{S}} \to \mathbb{C}$ be a continuous function, holomorphic on $\mathbb{S}$, whose real part equals the harmonic function

$$\Re\varPhi(u+iv) = \sum_{j=0,1} \int_{-\infty}^{\infty} \phi_j(t) P_j(u+iv;t)\,\mathrm{d}t.$$

Note that $\Re\varPhi(iv) = \phi_0(v)$ and $\Re\varPhi(1+iv) = \phi_1(v)$. The function $z \mapsto \exp(-\varPhi(z))f(z)$ belongs to $\mathscr{H}(X_0,X_0)$ and satisfies

$$\begin{aligned} \|\exp(-\varPhi(iv))f(iv)\|_{X_0} &= \exp(-\phi_0(z))\|f(iv)\|_{X_0} \leqslant 1, \\ \|\exp(-\varPhi(1+iv))f(iv)\|_{X_1} &= \exp(-\phi_1(z))\|f(1+iv)\|_{X_1} \leqslant 1, \end{aligned}$$

and therefore we have

$$\|\exp(-\Phi(\theta))f(\theta)\|_{[X_0,X_1]_\theta} \leqslant 1.$$

It follows that $\|f(\theta)\|_{[X_0,X_1]_\theta} \leqslant \exp(\Re\Phi(\theta))$ and therefore

$$\begin{aligned}\log\|f(\theta)\|_{[X_0,X_1]_\theta} &\leqslant \Re\Phi(\theta)\\ &= \sum_{j=0,1}\int_{-\infty}^{\infty}\phi_j(t)P_j(\theta;t)\,\mathrm{d}t = \sum_{j=0,1}\int_{-\infty}^{\infty}\phi_j(t)P_j(\theta;t)\,\mathrm{d}t.\end{aligned}$$

The first part of the lemma now follows by taking the infimum over all admissible ϕ_0 and ϕ_1 satisfying (C.5).

The second part follows from the first via Jensen's inequality and (C.4). Indeed, for $j=0$ we have

$$\begin{aligned}\int_{-\infty}^{\infty}\log\|f(j+it)\|_{X_j}P_0(\theta;t)\,\mathrm{d}t &= (1-\theta)\int_{-\infty}^{\infty}\log\|f(j+it)\|_{X_j}\frac{P_0(\theta;t)}{1-\theta}\,\mathrm{d}t\\ &\leqslant \log\Big[\Big(\int_{-\infty}^{\infty}\|f(j+it)\|_{X_j}\frac{P_j(\theta;t)}{1-\theta}\,\mathrm{d}t\Big)^{1-\theta}\Big].\end{aligned}$$

A similar estimate holds for $j=1$ with $1-\theta$ replaced by θ. Now we apply the exponential function on both sides and apply Jensen's inequality. $\square$

Combining the second part of the lemma with Hölder's inequality (note that $P_j(\theta;\cdot)\in L^p(\mathbb{R})$ for any $p\in[1,\infty]$), and using the trivial inequality $A^{1-\theta}B^\theta \leqslant (1-\theta)A+\theta B$, we arrive at:

Corollary C.2.11. *Let $p_0,p_1\in[1,\infty]$ and $\theta\in(0,1)$. If $f\in\mathscr{H}(X_0,X_1)$ satisfies*

$$f(i\cdot)\in L^{p_0}(\mathbb{R};X_0),\quad f(1+i\cdot)\in L^{p_1}(\mathbb{R};X_1),$$

then $f(\theta)\in[X_0,X_1]_\theta$ and

$$\|f(\theta)\|_{[X_0,X_1]_\theta} \leqslant C_{\theta,p_0',p_1'}\big(\|f(i\cdot)\|_{L^{p_0}(\mathbb{R};X_0)} + \|f(1+i\cdot)\|_{L^{p_1}(\mathbb{R};X_1)}\big),$$

where $C_{\theta,p_0',p_1'} = \max\{\|P_0(\theta,\cdot)\|_{L^{p_0'}(\mathbb{R})}, \|P_1(1-\theta,\cdot)\|_{L^{p_1'}(\mathbb{R})}\}$.

C.3 Real interpolation

We will define the real interpolation spaces through the so-called K-*method*, which assigns to each interpolation couple (X_0,X_1) and each $0<\theta<1$ and $1\leqslant p\leqslant\infty$ an interpolation space $X_{\theta,p}$.

For $t>0$ and $x\in X_0+X_1$, we define the K-*functional* $K(t,x;X_0,X_1)$ by

$$K(t,x;X_0,X_1) := \inf\{\|x_0\|_{X_0}+t\|x_1\|_{X_1} :\ x_0\in X_0,\ x_1\in X_1,\ x=x_0+x_1\}.$$

For each $x\in X_0+X_1$ the function $t\mapsto K(t,x;X_0,X_1)$ is non-decreasing and continuous, and for each $t>0$ the function $x\mapsto K(t,x;X_0,X_1)$ is subadditive. Moreover, the following identity holds:

$$tK(t^{-1},x;X_0,X_1) = K(t,x;X_1,X_0).$$

Definition C.3.1 (Real interpolation spaces via the K-method). *For* $0<\theta<1$ *and* $1 \leqslant p \leqslant \infty$ *we define*

$$(X_0, X_1)_{\theta,p} := \big\{x \in X_0 + X_1 : \|x\|_{\theta,p} < \infty\big\},$$

where

$$\begin{aligned} \|x\|_{\theta,p} &:= \Big(\int_0^\infty [t^{-\theta} K(t,x;X_0,X_1)]^p \,\frac{\mathrm{d}t}{t}\Big)^{1/p}, && 1 \leqslant p < \infty; \\ \|x\|_{\theta,\infty} &:= \operatorname*{ess\,sup}_{t>0} t^{-\theta} K(t,x;X_0,X_1), && p=\infty. \end{aligned}$$

It is easy to verify that $(X,X)_{\theta,p} = X$ and

$$(X_0,X_1)_{\theta,p} = (X_1,X_0)_{1-\theta,p}, \quad 1 \leqslant p \leqslant \infty,\ 0<\theta<1,$$

with identical norms, the latter being immediate by substituting $t = 1/s$.

Proposition C.3.2. *Let* $0<\theta<1$ *and* $1 \leqslant p \leqslant \infty$*. Endowed with the above norm the space* $(X_0,X_1)_{\theta,p}$ *is a Banach space. Moreover, for* $1 \leqslant p_0 \leqslant p_1 \leqslant \infty$ *we have continuous embeddings*

$$X_0 \cap X_1 \hookrightarrow (X_0,X_1)_{\theta,p_0} \hookrightarrow (X_0,X_1)_{\theta,p_1} \hookrightarrow X_0 + X_1. \tag{C.6}$$

Proof. Let us write $K(t,x) := K(t,x;X_0,X_1)$ for brevity.

The continuity of the first and third embedding in (C.6) follow from

$$\min\{1,t\}\|x\|_{X_0+X_1} \leqslant K(t,x) \leqslant \min\{1,t\}\|x\|_{X_0\cap X_1}. \tag{C.7}$$

To prove the continuity of the middle embedding note that

$$\begin{aligned} K(t,x) &\eqsim_{\theta,p} t^\theta K(t,x)\Big(\int_t^\infty s^{-\theta p}\,\frac{\mathrm{d}s}{s}\Big)^{1/p} \\ &\leqslant t^\theta \Big(\int_t^\infty [s^{-\theta}K(s,x)]^p \,\frac{\mathrm{d}s}{s}\Big)^{1/p} \leqslant t^\theta \|x\|_{\theta,p} \end{aligned} \tag{C.8}$$

if $1 \leqslant p < \infty$, the case $p=\infty$ being easier. By (C.8),

$$\begin{aligned} \|x\|_{\theta,p_1} &\leqslant \Big(\int_0^\infty [t^{-\theta}K(t,x)]^{p_0}\,\frac{\mathrm{d}t}{t}\Big)^{1/p_1} \cdot \operatorname*{ess\,sup}_{t>0} \big(t^{-\theta}K(t,x)\big)^{1-p_0/p_1} \\ &\leqslant \|x\|_{\theta,p_0}^{p_0/p_1} \cdot \|x\|_{\theta,p_0}^{1-p_0/p_1} = \|x\|_{\theta,p_0} \end{aligned}$$

if $1 \leqslant p_0 \leqslant p_1 < \infty$, the remaining cases being easier.

We will give the proof of the completeness of $(X_0,X_1)_{\theta,p}$ for $1 \leqslant p < \infty$, the case $p=\infty$ being easier. Suppose that $(x_n)_{n\geqslant 1}$ is a Cauchy sequence in $(X_0,X_1)_{\theta,p}$. Then $(x_n)_{n\geqslant 1}$ converges in X_0+X_1, say to the limit $x \in X_0+X_1$. Given $\varepsilon>0$ choose $N \geqslant 1$ so large that $\|x_n - x_m\|_{\theta,p} < \varepsilon$ for all $n,m \geqslant N$. By the triangle inequality, for all $0<r<R<\infty$ and $m \geqslant n \geqslant N$ we have

$$\Big(\int_r^R [t^{-\theta} K(t, x - x_n)]^p \,\frac{\mathrm{d}t}{t}\Big)^{1/p} \leqslant \varepsilon + \Big(\int_r^R [t^{-\theta} K(t, x - x_m)]^p \,\frac{\mathrm{d}t}{t}\Big)^{1/p}. \quad \text{(C.9)}$$

From the inequality

$$K(t, y) \leqslant \max\{1, t\} \|y\|_{X_0+X_1}$$

and the fact that $x_m \to x$ in $X_0 + X_1$ we infer that the right-hand side integral in (C.9) tends to 0 as $m \to \infty$. Passing to the limits $m \to \infty$, $r \downarrow 0$ and $R \to \infty$, (C.9) shows that $\|x - x_n\|_{\theta,p} \leqslant \varepsilon$ for all $n \geqslant N$. From this we infer that $x \in (X_0, X_1)_{\theta,p}$ and that $(x_n)_{n\geqslant 1}$ converges to x in $(X_0, X_1)_{\theta,p}$. □

The spaces $(X_0, X_1)_{\theta,p}$ are interpolation spaces between X_0 and X_1:

Theorem C.3.3 (Real interpolation of operators). *Let (X_0, X_1) and (Y_0, Y_1) be interpolation couples. Suppose $T : X_0 + X_1 \to Y_0 + Y_1$ is a linear operator which maps X_0 into Y_0 and X_1 into Y_1 with norms*

$$\|T\|_{\mathscr{L}(X_0,Y_0)} = A_0, \qquad \|T\|_{\mathscr{L}(X_1,Y_1)} = A_1.$$

Then for all $0 < \theta < 1$ and $1 \leqslant p \leqslant \infty$ the operator T maps $(X_0, X_1)_{\theta,p}$ into $(Y_0, Y_1)_{\theta,p}$ and we have

$$\|T\|_{\mathscr{L}((X_0,X_1)_{\theta,p},(Y_0,Y_1)_{\theta,p})} \leqslant A_0^{1-\theta} A_1^{\theta}.$$

Proof. Let us first assume that $A_0 \neq 0$. Fix $x \in (X_0, X_1)_{\theta,p}$ and write $x = x_0 + x_1$ with $x_0 \in X_0$ and $x_1 \in X_1$. Then for every $t > 0$,

$$\|Tx_0\|_{Y_0} + t\|Tx_1\|_{Y_1} \leqslant A_0\Big(\|x_0\|_{X_0} + t\frac{A_1}{A_0}\|x_1\|_{X_1}\Big)$$

and therefore

$$K(Tx, t; Y_0, Y_1) \leqslant A_0 K\Big(x, t\frac{A_1}{A_0}; X_0, X_1\Big).$$

With the substitution $s = tA_1/A_0$ we obtain, if $1 \leqslant p < \infty$,

$$\begin{aligned}
\|Tx\|_{\theta,p}^p &= \int_0^\infty [t^{-\theta} K(Tx, t; Y_0, Y_1)]^p \,\frac{\mathrm{d}t}{t} \\
&\leqslant \int_0^\infty \Big[t^{-\theta} A_0 K\Big(x, t\frac{A_1}{A_0}; X_0, X_1\Big)\Big]^p \,\frac{\mathrm{d}t}{t} \\
&= A_0^p \Big(\frac{A_1}{A_0}\Big)^{\theta p} \int_0^\infty [s^{-\theta} K(x, s; X_0, X_1)]^p \,\frac{\mathrm{d}s}{s} = A_0^{(1-\theta)p} A_1^{\theta p} \|x\|_{\theta,p}^p.
\end{aligned}$$

The proof for $p = \infty$ requires obvious modifications.

This proves the theorem under the assumption $A_0 \neq 0$. If $A_0 = 0$ and $A_1 \neq 0$ we reverse the roles of the indices 0 and 1. Finally, if $A_0 = A_1 = 0$ then $T = 0$ and there is nothing to prove. □

For certain applications it is desirable to have a characterisation of the real interpolation method in terms of certain L^p norms. To this end we first introduce the so-called "first mean method".

Definition C.3.4 (First mean method). *Let (X_0, X_1) be an interpolation couple of Banach spaces. For $0 < \theta < 1$ and $1 \leqslant p_0, p_1 \leqslant \infty$ the space $(X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$ is defined as the set of all $x \in X_0 + X_1$ for which there exist strongly measurable functions $u_0 : (0,\infty) \to X_0$ and $u_1 : (0,\infty) \to X_1$ with the following properties:*

(i) *$x = u_0(t) + u_1(t)$ for almost all $t > 0$;*
(ii) *$t \mapsto t^{j-\theta} u_j(t)$ belongs to $L^{p_j}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_j)$ $(j = 0, 1)$.*

On $(X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$ we define a norm by

$$\|x\|^{(\mathrm{m})}_{\theta,p_0,p_1} := \inf_{x\equiv u_0+u_1} \max\Big\{\|t \mapsto t^{-\theta} u_0(t)\|_{L^{p_0}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)}, \|t \mapsto t^{1-\theta} u_1(t)\|_{L^{p_1}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)}\Big\},$$

where the infimum is taken over all decompositions $x \equiv u_0 + u_1$ with the above properties.

Lemma C.3.5. *The space $(X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$ is a Banach space and we have continuous embeddings*

$$X_0 \cap X_1 \hookrightarrow (X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1} \hookrightarrow X_0 + X_1.$$

Proof. We leave it to the reader to check the first assertion, the proof of which is similar to that of Proposition C.1.3.

Fix $x \in X_0 \cap X_1$. To see that x is contained in $(X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$, note that the functions $u_0 = \mathbf{1}_{[1,\infty)} \otimes x$ and $u_1 = \mathbf{1}_{(0,1)} \otimes x$ satisfy $u_0 + u_1 \equiv x$ on $(0, \infty)$ and $t \mapsto t^{-\theta} u_0(t) \in L^{p_0}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_0)$ and $t \mapsto t^{1-\theta} u_1(t) \in L^{p_1}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_1)$, respectively, with norms

$$\begin{aligned} \|t \mapsto t^{-\theta} u_0(t)\|_{L^{p_0}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)} &= (\theta p_0)^{-1/p_0} \|x\|_{X_0}, \\ \|t \mapsto t^{1-\theta} u_1(t)\|_{L^{p_1}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)} &= ((1-\theta)p_1)^{-1/p_1} \|x\|_{X_1}. \end{aligned}$$

It follows that $x \in (X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$ and

$$\|x\|_{(X_0,X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}} \leqslant \max\Big\{(\theta p_0)^{-1/p_0}, ((1-\theta)p_1)^{-1/p_1}\Big\} \|x\|_{X_0\cap X_1}.$$

This also gives the continuity of the first embedding.

Next let $x \in (X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$ and write $x = u_0 + u_1$ as in the definition. Then $x = \int_1^2 (tu_0(t) + tu_1(t)) \frac{\mathrm{d}t}{t} \in X_0 + X_1$ with norm

$$\begin{aligned}\|x\|_{X_0+X_1} &\leqslant \Big\| \int_1^2 t u_0(t)\,\frac{\mathrm{d}t}{t}\Big\|_{X_0} + \Big\|\int_1^2 t u_1(t)\,\frac{\mathrm{d}t}{t}\Big\|_{X_1} \\ &\leqslant \|t\mapsto t^{-\theta}u_0(t)\|_{L^{p_0}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)}\cdot \|t\mapsto t^{1+\theta}\|_{L^{p_0'}((1,2),\frac{\mathrm{d}t}{t})} \\ &\qquad + \|t\mapsto t^{1-\theta}u_1(t)\|_{L^{p_1}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)}\cdot \|t\mapsto t^{\theta}\|_{L^{p_1'}((1,2),\frac{\mathrm{d}t}{t})}.\end{aligned}$$

Taking the infimum over all admissible representations $x = u_0 + u_1$, it follows that

$$\|x\|_{X_0+X_1} \leqslant 2C_{\theta,p_0,p_1}\|x\|_{(X_0,X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}}$$

where $C_{\theta,p_0,p_1} := \max\{\|t\mapsto t^{1+\theta}\|_{L^{p_0'}((1,2),\frac{\mathrm{d}t}{t})}, \|t\mapsto t^{\theta}\|_{L^{p_1'}((1,2),\frac{\mathrm{d}t}{t})}\}$. This also gives the continuity of the second embedding. □

It will turn out (see Theorem C.3.14) that $(X_0,X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$ coincides with the real interpolation space $(X_0,X_1)_{\theta,p}$ for $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$ up to equivalence of norms. We will first prove this fact for $p_0 = p_1 = p$.

Proposition C.3.6. *Let (X_0,X_1) be an interpolation couple of Banach spaces and let $1 \leqslant p \leqslant \infty$ and $0 < \theta < 1$. Then $(X_0,X_1)_{\theta,p} = (X_0,X_1)^{(\mathrm{m})}_{\theta,p,p}$ with equivalent norms satisfying*

$$\frac{1}{2}\|x\|_{\theta,p} \leqslant \|x\|^{(\mathrm{m})}_{\theta,p,p} \leqslant \|x\|_{\theta,p}. \tag{C.10}$$

Moreover, it suffices to consider continuous functions $u_0 : (0,\infty) \to X_0$ and $u_1 : (0,\infty) \to X_1$ in Definition C.3.4.

Proof. Fix $x \in (X_0,X_1)^{(\mathrm{m})}_{\theta,p,p}$ and let us write $K(t,x) := K(t,x;X_0,X_1)$ for brevity. Suppose $x \equiv u_0 + u_1$ as in Definition C.3.4. Then, for almost all $t > 0$,

$$K(t,x) \leqslant \|u_0(t)\|_{X_0} + t\|u_1(t)\|_{X_1}.$$

Multiplying both sides with $t^{-\theta}$ and taking L^p-norms, we find

$$\begin{aligned}&\|t\mapsto t^{-\theta}K(t,x)\|_{L^p(\mathbb{R}_+,\frac{\mathrm{d}t}{t})} \\ &\quad\leqslant \|t\mapsto t^{-\theta}u_0(t)\|_{L^p(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)} + \|t\mapsto t^{1-\theta}u_1(t)\|_{L^p(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)} \\ &\quad\leqslant 2\max\Big\{\|t\mapsto t^{-\theta}u_0(t)\|_{L^p(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)}, \|t\mapsto t^{1-\theta}u_1(t)\|_{L^p(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)}\Big\}.\end{aligned}$$

We conclude that $x \in (X_0,X_1)_{\theta,p}$. Taking the infimum over all admissible representations $x \equiv u_0 + u_1$, the first inequality in (C.10) is obtained.

In the converse direction, suppose that $x \in (X_0,X_1)_{\theta,p}$ is a non-zero element and let $\varepsilon > 0$. For every $t_0 > 0$ select a decomposition $x = x_{0,t_0} + x_{1,t_0}$ in $X_0 + X_1$ such that

$$\|x_{0,t_0}\|_{X_0} + t_0\|x_{1,t_0}\|_{X_1} < (1+\varepsilon)K(t_0,x).$$

By the continuity of $K(t_0,x)$ as a function of t_0 there is a small interval I_{t_0} about t_0 such that

$$\|x_{0,t_0}\|_{X_0} + t\|x_{1,t_0}\|_{X_1} < (1+\varepsilon)K(t,x)$$

for all $t \in I_{t_0}$. The covering of $(0,\infty)$ with the intervals I_{t_0}, $t_0>0$, has a countable sub-cover. Using this sub-cover, we find piecewise constant functions $u_j : (0,\infty) \to X_j$ $(j=0,1)$ such that $x = u_0(t)+u_1(t)$ for all $t>0$ and

$$\|u_0(t)\|_{X_0} + t\|u_1(t)\|_{X_1} < (1+\varepsilon)K(t,x)$$

for all $t>0$. Multiplying both sides with $t^{-\theta}$ and taking L^p-norms, we find

$$\|t \mapsto t^{-\theta}u_0(t)\|_{L^p(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)} \leqslant (1+\varepsilon)\|t\mapsto t^{-\theta}K(t,x)\|_{L^p(\mathbb{R}_+,\frac{\mathrm{d}t}{t})}$$

and

$$\|t \mapsto t^{1-\theta}u_1(t)\|_{L^p(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)} \leqslant (1+\varepsilon)\|t\mapsto t^{-\theta}K(t,x)\|_{L^p(\mathbb{R}_+,\frac{\mathrm{d}t}{t})},$$

and therefore the maximum of these two expressions is bounded above by $(1+\varepsilon)\|x\|_{\theta,p}$. This shows that the decomposition $x = u_0+u_1$ has the properties (i) and (ii). Since $\varepsilon>0$ was arbitrary, this proves the second inequality in (C.10). Finally, by linear interpolation on small intervals around the jump points, the functions u_0 and u_1 can be taken to be continuous. □

In the next proposition we continue with an estimate for the norm of $(X_0,X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$.

Proposition C.3.7. *Let $1\leqslant p_0 \leqslant p_1 \leqslant \infty$ and $0<\theta<1$, and suppose that $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$. Then for all $x \in X_0\cap X_1$ we have*

$$\frac{1}{4}\|x\|_{\theta,p} \leqslant \|x\|^{(\mathrm{m})}_{\theta,p_0,p_1} \leqslant 2\|x\|_{\theta,p}. \tag{C.11}$$

The special case $p_0=p_1$ was already considered in Proposition C.3.6, so we may assume $1\leqslant p_0<p_1\leqslant\infty$ below.

For the proof of Proposition C.3.7 we need some preparation. Given an element $x\in X_0\cap X_1$, for $t\geqslant 0$ we set

$$f(x,t) := \inf\big\{\|x_1\|_{X_1} :\ x = x_0+x_1,\ \|x_0\|_{X_0}\leqslant t\big\}.$$

As a function of t, this function is non-negative, non-increasing, and convex, and it satisfies $f(x,t)=0$ for $t\geqslant \|x\|_{X_0}$. We claim that

$$\|x\|^{(\mathrm{m})}_{\theta,p_0,p_1} = \inf_w \max\Big\{\|t\mapsto t^{-\theta}w(t)\|_{L^{p_0}(\mathbb{R}_+,\frac{\mathrm{d}t}{t})}, \|t\mapsto t^{1-\theta}f(x,w(t))\|_{L^{p_1}(\mathbb{R}_+,\frac{\mathrm{d}t}{t})}\Big\}, \tag{C.12}$$

where the infimum is taken over all non-negative continuous functions w. Indeed, the inequality "$\geqslant$" follows by taking $w(t) = \|u_0(t)\|_{X_0}$ and $x_0 = u_0(t)$, $x_1 = u_1(t)$ in the definition of f. The inequality "$\leqslant$" follows using the definition of f and a selection argument similar to the one in Proposition C.3.6.

First assume $p_1 < \infty$. Fix a non-zero $x \in X_0 \cap X_1$ and write $f(t) := f(x,t)$. Let $s_0 := \min\{s \geqslant 0 : \ f(s) = 0\}$. Then $0 < s_0 \leqslant \|x\|_{X_0} < \infty$ and f is strictly decreasing on the interval $[0, s_0)$ (since f is non-increasing and convex on $[0,\infty)$ and vanishes on $[s_0,\infty)$). Set

$$\alpha := \alpha(x) := \theta^{(p_0-p_1)/p} \|t \mapsto t^{1-\theta} f(t)^\theta\|^{p_1-p_0}_{L^p(\mathbb{R}_+, \frac{\mathrm{d}t}{t})}.$$

The continuous function $v^{-1} : [0, s_0) \to [0,\infty)$ given by

$$v^{-1}(s) := \alpha^{p/(p_0p_1)} s^{p/p_1} f(s)^{-p/p_0} \tag{C.13}$$

is strictly increasing and satisfies $v^{-1}(0) = 0$ and $\lim_{s\to s_0} v^{-1}(s) = \infty$. Accordingly v^{-1} has an inverse $v : [0,\infty) \to [0, s_0)$. It satisfies

$$t = \alpha^{p/(p_0p_1)} v(t)^{p/p_1} f(v(t))^{-p/p_0},$$

or equivalently, for all $t > 0$

$$\alpha t^{-\theta p_0} v(t)^{p_0} = t^{(1-\theta)p_1} f(v(t))^{p_1}. \tag{C.14}$$

In the case $p_1 = \infty$ we let $v^{-1} : [0, s_0) \to [0,\infty)$ be given by

$$v^{-1}(s) := f(s)^{-p/p_0} \theta^{-1/p_0} \|t \mapsto t^{1-\theta} f(t)^\theta\|^{p/p_0}_{L^p(\mathbb{R}_+, \frac{\mathrm{d}t}{t})}. \tag{C.15}$$

Then, using that $p_0 = (1-\theta)p$ we find that for all $\tau > 0$,

$$\tau^{1-\theta} f(v(\tau)) = \theta^{-1/p} \|t \mapsto t^{1-\theta} f(t)^\theta\|_{L^p(\mathbb{R}_+, \frac{\mathrm{d}t}{t})}. \tag{C.16}$$

Lemma C.3.8. *Under the assumptions of Proposition C.3.7, if in addition $p_0 < \infty$ we have*

$$\begin{aligned}\|t \mapsto t^{-\theta} v(t)\|_{L^{p_0}(\mathbb{R}_+, \frac{\mathrm{d}t}{t})} &= \|t \mapsto t^{1-\theta} f(v(t))\|_{L^{p_1}(\mathbb{R}_+, \frac{\mathrm{d}t}{t})} \\ &= \theta^{-1/p} \|t \mapsto t^{1-\theta} f(t)^\theta\|_{L^p(\mathbb{R}_+, \frac{\mathrm{d}t}{t})}.\end{aligned}$$

These identities are to be understood in the sense that either all three quantities are infinite, or else if one of them is finite, then so are the other two, and in that case all three have the same finite value.

Proof. Substituting $t = v^{-1}(s)$ and integrating by parts,

$$\begin{aligned}\int_0^\infty t^{-\theta p_0} v(t)^{p_0} \frac{\mathrm{d}t}{t} &= \int_0^{s_0} v^{-1}(s)^{-\theta p_0 - 1} s^{p_0} \frac{\mathrm{d}}{\mathrm{d}s}(v^{-1}(s)) \,\mathrm{d}s \\ &= -\frac{1}{\theta p_0} \int_0^{s_0} s^{p_0} \frac{\mathrm{d}}{\mathrm{d}s}(v^{-1}(s)^{-\theta p_0}) \,\mathrm{d}s \\ &= \frac{1}{\theta} \int_0^{s_0} s^{p_0 - 1} v^{-1}(s)^{-\theta p_0} \,\mathrm{d}s,\end{aligned}$$

using that if the integral $\int_0^\infty t^{-\theta p_0} v(t)^{p_0} \frac{\mathrm{d}t}{t}$ converges, then $\lim_{s\downarrow 0} sv^{-1}(s)^{-\theta} = 0$ and $\lim_{s\to s_0} sv^{-1}(s)^{-\theta} = 0$ and we claim that the boundary terms vanish in the integration by parts. Indeed, suppose for a contradiction that the former fails. Substituting $t = v(s)$ then gives $\limsup_{t\to\infty} t^{-\theta} v(t) > 0$. Thus we can find an $\varepsilon > 0$ and a sequence $t_n \uparrow \infty$ with $t_{n+1} - t_n \geqslant 1$ and $t_n^{-\theta} v(t_n) \geqslant \varepsilon$. For large enough n we have $(t_n + 1)^{-\theta} \geqslant \frac{1}{2} t_n^{-\theta}$ and therefore $t^{-\theta} v(t) \geqslant \frac{1}{2}\varepsilon$ for all $t \in [t_n, t_n + 1]$. This is the desired contradiction. The vanishing of $\lim_{s\to s_0} sv^{-1}(s)^{-\theta} = 0$ is immediate from the fact that $\lim_{s\to s_0} v^{-1}(s) = \infty$ and the claim follows.

Assume first that $p_1 < \infty$. Using the definitions of α and v^{-1} in (C.13),

$$\begin{aligned}\frac{1}{\theta}\int_0^{s_0} v^{-1}(s)^{-\theta p_0} s^{p_0 - 1}\,\mathrm{d}s &= \frac{1}{\theta}\alpha^{-\theta p/p_1}\int_0^\infty f(s)^{\theta p} s^{-\theta p_0 p/p_1 + p_0 - 1}\,\mathrm{d}s \\ &= \frac{1}{\theta}\alpha^{-\theta p/p_1}\int_0^\infty s^{(1-\theta)p} f(s)^{\theta p}\,\frac{\mathrm{d}s}{s} \\ &= \theta^{-p_0/p}\Big(\int_0^\infty s^{(1-\theta)p} f(s)^{\theta p}\,\frac{\mathrm{d}s}{s}\Big)^{p_0/p}.\end{aligned}$$

This shows that if the first quantity in the statement of the lemma is finite, then so is the third and it is equal to the first. Running the argument backwards, the finiteness of the third quantity also implies the finiteness of the first.

The proof of the other equality is similar and is left to the reader.

When $p_1 = \infty$ we use the function v^{-1} defined by (C.15) to prove the equality of the first and third terms. The proof of the other equality is now a trivial consequence of (C.16). $\square$

Lemma C.3.9. *Under the assumptions of Proposition C.3.7 and using the preceding notation, if $1 \leqslant p_0 < \infty$ and $p_0 \leqslant p_1 \leqslant \infty$, then for all $x \in X_0 \cap X_1$ the three quantities of Lemma C.3.8 are finite and*

$$\|x\|_{\theta,p_0,p_1}^{(\mathrm{m})} \leqslant \theta^{-1/p}\|t \mapsto t^{1-\theta} f(t)^\theta\|_{L^p(\mathbb{R}_+, \frac{\mathrm{d}t}{t})} \leqslant 2\|x\|_{\theta,p_0,p_1}^{(\mathrm{m})}.$$

Proof. The first inequality follows from Lemma C.3.8 and (C.12) with $w = v$.

For the proof of the second inequality, let $w : (0,\infty) \to \mathbb{R}$ be a non-negative continuous function for which $t \mapsto t^{-\theta} w(t)$ belongs to $L^{p_0}(\mathbb{R}_+, \frac{\mathrm{d}t}{t})$ and $t \mapsto t^{1-\theta} f(x, w(t))$ belongs to $L^{p_1}(\mathbb{R}_+, \frac{\mathrm{d}t}{t})$. Put

$$M := M(w) := \{t > 0 :\ v(t) \leqslant w(t)\}.$$

First assume $p_1 < \infty$. By Lemma C.3.8, writing $\int_0^\infty = \int_M + \int_{\complement M}$, using the triangle inequality, (C.14), and that f is non-increasing,

$$\begin{aligned}I &:= \theta^{-1/p}\|t^{1-\theta} f(t)^\theta\|_{L^p(\mathbb{R}_+, \frac{\mathrm{d}t}{t})} = \|t^{-\theta} v(t)\|_{L^{p_0}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_0)} \\ &\leqslant \Big(\int_M t^{-\theta p_0} v(t)^{p_0}\,\frac{\mathrm{d}t}{t}\Big)^{1/p_0} + \alpha^{-1/p_0}\Big(\int_{\complement M} t^{(1-\theta)p_1} f(v(t))^{p_1}\,\frac{\mathrm{d}t}{t}\Big)^{1/p_0}\end{aligned}$$

$$\leqslant \Big(\int_M t^{-\theta p_0} w(t)^{p_0} \frac{\mathrm{d}t}{t}\Big)^{1/p_0} + \alpha^{-1/p_0}\Big(\int_{\complement M} t^{(1-\theta)p_1} f(w(t))^{p_1} \frac{\mathrm{d}t}{t}\Big)^{1/p_0}$$

which is finite. Putting

$$I_0 := \|t \mapsto t^{-\theta} w(t)\|_{L^{p_0}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_0)}, \quad I_1 := \|t \mapsto t^{1-\theta} f(w(t))\|_{L^{p_1}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_1)}$$

and using the definition of α, we obtain

$$I \leqslant (I_0 + I^{-(p_1-p_0)/p_0} I_1^{p_1/p_0}).$$

With $K_0 := I_0/I$, $K_1 := I_1/I$, $q := p_1/p_0 \geqslant 1$, this can be rewritten as

$$1 \leqslant K_0 + K_1^q.$$

This implies that $\max\{K_0, K_1\} \geqslant \frac{1}{2}$, i.e.,

$$\begin{aligned} I &\leqslant 2\max\{I_0, I_1\} \\ &\leqslant 2\max\Big\{\|t \mapsto t^{-\theta} w(t)\|_{L^p(\mathbb{R}_+, \frac{\mathrm{d}t}{t})}, \|t \mapsto t^{1-\theta} f(w(t))\|_{L^{p_1}(\mathbb{R}_+, \frac{\mathrm{d}t}{t})}\Big\}. \end{aligned}$$

Taking the infimum over all admissible functions w, we obtain the second inequality of the lemma.

For $p_1 = \infty$ we proceed similarly. First assume that M is a proper subset of $(0, \infty)$. Then by (C.16) we have $I = t^{(1-\theta)} f(v(t))$

$$I = \sup_{t \in \complement M} t^{(1-\theta)} f(v(t)) \leqslant \sup_{t \in \complement M} t^{(1-\theta)} f(w(t)) \leqslant I_1 \leqslant \max\{I_0, I_1\},$$

with I, I_0, I_1 as before; we used that M is closed and therefore the open set $\complement M$ has positive measure. If $M = (0, \infty)$, then by Lemma C.3.8

$$I = \|t \mapsto t^{-\theta} v(t)\|_{L^{p_0}(\mathbb{R}_+, \frac{\mathrm{d}t}{t})} \leqslant \|t \mapsto t^{-\theta} w(t)\|_{L^{p_0}(\mathbb{R}_+, \frac{\mathrm{d}t}{t})} = I_0 \leqslant \max\{I_0, I_1\},$$

and the result follows as before. □

Proof of Proposition C.3.7. The case $p_0 = p_1$ has already been dealt with in Proposition C.3.6 so it suffices to consider $p_0 < p_1$.

Put

$$J := \theta^{-1/p} \|t \mapsto t^{1-\theta} f(t)^{\theta}\|_{L^p(\mathbb{R}_+, \frac{\mathrm{d}t}{t})}.$$

By Lemma C.3.9, J is finite if and only if the infimum defining $\|x\|^{(\mathrm{m})}_{\theta, p_0, p_1}$ is finite, with a two-sided estimate

$$\|x\|^{(\mathrm{m})}_{\theta, p_0, p_1} \leqslant J \leqslant 2\|x\|^{(\mathrm{m})}_{\theta, p_0, p_1}.$$

Similarly, applying Lemma C.3.9 with $p_0 = p_1 = p$, we find that J is finite if and only if the infimum defining $\|x\|^{(\mathrm{m})}_{\theta, p_0, p_1}$ is finite, with a two-sided estimate

$$\|x\|^{(\mathrm{m})}_{\theta, p, p} \leqslant J \leqslant 2\|x\|^{(\mathrm{m})}_{\theta, p, p}.$$

Now the result follows from Proposition C.3.6, where it has been shown that $\frac{1}{2}\|x\|_{\theta, p} \leqslant \|x\|^{(\mathrm{m})}_{\theta, p, p} \leqslant \|x\|_{\theta, p}$. □

In order to state and prove the next result we employ the so-called "second mean method". Later on it will turn out that the space $(X_0, X_1)_{\theta,p_0,p_1}$ we are about to introduce coincides with the real interpolation space $(X_0, X_1)_{\theta,p}$ with $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$.

Definition C.3.10 (Lions–Peetre method). *For $1 \leqslant p_0 \leqslant p_1 \leqslant \infty$ and $0 < \theta < 1$, the space $(X_0, X_1)_{\theta,p_0,p_1}$ is defined as the set of all $x \in X_0 + X_1$ for which there exists a strongly measurable function $u : (0, \infty) \to X_0 \cap X_1$ with the following properties:*

(i) $x = \int_0^\infty u(t) \frac{\mathrm{d}t}{t}$ *with convergence of the improper integral in $X_0 + X_1$;*
(ii) $t \mapsto t^{j-\theta} u(t)$ *belongs to $L^{p_j}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_j)$ $(j = 0, 1)$.*

On $(X_0, X_1)_{\theta,p_0,p_1}$ we define a norm by

$$\begin{aligned}\|x\|_{\theta,p_0,p_1} \\ := \inf_u \max\Big\{\|t \mapsto t^{-\theta}u(t)\|_{L^{p_0}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)}, \|t \mapsto t^{1-\theta}u(t)\|_{L^{p_1}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)}\Big\},\end{aligned}$$

where the infimum is taken over all functions u having the properties listed in the definition. It will turn out in a moment that

$$(X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1} = (X_0, X_1)_{\theta,p_0,p_1}$$

with equivalent norms. In particular this implies that $(X_0, X_1)_{\theta,p_0,p_1}$ is a Banach space.

Note that the convergence of the integral in $X_0 + X_1$ in (i) follows from (ii) since for $\tau > 0$,

$$\begin{aligned}
&\int_0^\infty \|u(t)\|_{X_0+X_1} \frac{\mathrm{d}t}{t} \\
&\quad\leqslant \int_0^\tau \|u(t)\|_{X_0} \frac{\mathrm{d}t}{t} + \int_\tau^\infty \|u(t)\|_{X_1} \frac{\mathrm{d}t}{t} \\
&\quad\leqslant \Big(\int_0^\infty t^{-\theta p_0}\|u(t)\|_{X_0}^{p_0} \frac{\mathrm{d}t}{t}\Big)^{1/p_0} \Big(\int_0^\tau t^{\theta p_0'} \frac{\mathrm{d}t}{t}\Big)^{1/p_0'} \\
&\quad\quad + \Big(\int_0^\infty t^{(1-\theta)p_1}\|u(t)\|_{X_1}^{p_1} \frac{\mathrm{d}t}{t}\Big)^{1/p_1} \Big(\int_\tau^\infty t^{-(1-\theta)p_1'} \frac{\mathrm{d}t}{t}\Big)^{1/p_1'} < \infty
\end{aligned} \tag{C.17}$$

with obvious changes if $p_1 = \infty$. Similarly one checks that $\int_a^b u(t) \frac{\mathrm{d}t}{t} \in X_0 \cap X_1$ for all $0 < a < b < \infty$; this will be used later on.

Lemma C.3.11. *For all $x \in (X_0, X_1)_{\theta,p_0,p_1}$ we have*

$$\|x\|_{\theta,p_0,p_1} = \inf_u \Big(\|t \mapsto t^{-\theta}u(t)\|^{1-\theta}_{L^{p_0}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)} \cdot \|t \mapsto t^{1-\theta}u(t)\|^{\theta}_{L^{p_1}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)}\Big),$$

where the infimum is taken over all u as in Definition C.3.10.

Proof. First observe that for all $0 < \theta < 1$ and real numbers $a, b > 0$ we have

$$\inf_{\mu>0} \max\{\mu^\theta a, \mu^{\theta-1} b\} = a^{1-\theta} b^\theta. \tag{C.18}$$

Indeed, the left-hand side of (C.18) divided by $a^{1-\theta}b^\theta$ equals

$$\inf_{\mu>0} \max\Big\{\Big(\frac{\mu a}{b}\Big)^\theta, \Big(\frac{\mu a}{b}\Big)^{\theta-1}\Big\} = \inf_{\lambda>0} \max\big\{\lambda^\theta, \lambda^{\theta-1}\big\} = 1.$$

If $t \mapsto u(t)$ satisfies conditions in Definition C.3.10, then so does $t \mapsto u(\mu t)$ for every $\mu > 0$, and

$$\begin{aligned}
\|t \mapsto t^{-\theta} u(\mu t)\|_{L^{p_0}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_0)} &= \mu^\theta \|t \mapsto t^{-\theta} u(t)\|_{L^{p_0}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_0)}, \\
\|t \mapsto t^{1-\theta} u(\mu t)\|_{L^{p_1}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_1)} &= \mu^{\theta-1} \|t \mapsto t^{1-\theta} u(t)\|_{L^{p_1}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_1)}.
\end{aligned}$$

Therefore, (C.18) applied with

$$a = \|t \mapsto t^{-\theta} u(t)\|_{L^{p_0}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_0)} \quad \text{and} \quad b = \|t \mapsto t^{1-\theta} u(t)\|_{L^{p_1}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_1)}$$

gives the desired identity. □

Lemma C.3.12. *Let $1 \leqslant p_0 < \infty$ and $p_0 \leqslant p_1 \leqslant \infty$, and let $0 < \theta < 1$. Then $X_0 \cap X_1$ is dense in $(X_0, X_1)_{\theta,p_0,p_1}$.*

Proof. Fix $x \in (X_0, X_1)_{\theta,p_0,p_1}$ and let u be any function as in Definition C.3.10. Choose functions $\varphi_n \in C_c(0,\infty)$ subject to the conditions $0 \leqslant \varphi_n \leqslant 1$ and $\varphi_n(t) = 1$ for all $t \in [\frac{1}{n}, n]$. Then $x_n := \int_0^\infty \varphi_n(t) u(t) \frac{\mathrm{d}t}{t}$ belongs to $X_0 \cap X_1$. Moreover,

$$x - x_n = \int_0^\infty (1 - \varphi_n(t)) u(t) \, \frac{\mathrm{d}t}{t}$$

(with convergence of the improper integral in $X_0 + X_1$). Since $1 - \varphi_n(t) = 0$ on the interval $I_n = [1/n, n]$, by Lemma C.3.11 we find

$$\|x - x_n\|_{\theta,p_0,p_1} \leqslant \|t \mapsto t^{-\theta} u(t)\|^{1-\theta}_{L^{p_0}(\complement I_n, \frac{\mathrm{d}t}{t}; X_0)} \|t \mapsto t^{1-\theta} u(t)\|^{\theta}_{L^{p_1}(\complement I_n, \frac{\mathrm{d}t}{t}; X_1)}.$$

The right-hand side tends to 0 as $n \to \infty$, noting that the corresponding norms in $L^{p_0}(\complement I_n, \frac{\mathrm{d}t}{t}; X_0)$ tend to 0 since we assumed $p_0 < \infty$. □

Proposition C.3.13 (Equivalence of mean methods). *Let (X_0, X_1) be an interpolation couple of Banach spaces. Let $1 \leqslant p_0 \leqslant p_1 \leqslant \infty$, $0 < \theta < 1$, and $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$. Then*

$$(X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1} = (X_0, X_1)_{\theta,p_0,p_1}$$

with equivalent norms satisfying

$$\frac{1}{6}\|x\|_{\theta,p_0,p_1} \leqslant \|x\|^{(\mathrm{m})}_{\theta,p_0,p_1} \leqslant C_\theta \|x\|_{\theta,p_0,p_1}, \tag{C.19}$$

where $C_\theta = \max\{\theta^{-1}, (1-\theta)^{-1}\}$.

The proof will show also that in Definition C.3.10 it suffices to consider continuous functions u.

Proof. Let $x \in (X_0, X_1)_{\theta,p_0,p_1}$ and write $x = \int_0^\infty u(t)\frac{dt}{t}$, where u is as in Definition C.3.10. Set

$$u_0(t) := \int_0^t u(s)\,\frac{ds}{s}, \quad u_1(t) := \int_t^\infty u(s)\,\frac{ds}{s}.$$

Then $u_0(t) + u_1(t) = x$ for all $t > 0$. By (C.17) the integrals defining $u_0(t)$ and $u_1(t)$ exist as a Bochner integral in X_0 and X_1, respectively.

We estimate the $L^{p_0}(\mathbb{R}_+, \frac{dt}{t}; X_0)$-norm of $t \mapsto t^{-\theta}u_0(t)$. By Young's inequality for convolutions on the multiplicative group $(0,\infty)$ with Haar measure dt/t,

$$\begin{aligned}
&\|t \mapsto t^{-\theta}u_0(t)\|_{L^{p_0}(\mathbb{R}_+,\frac{dt}{t};X_0)} \\
&= \Big(\int_0^\infty \Big\|\int_0^t (t/s)^{-\theta}\mathbf{1}_{(1,\infty)}(t/s)\cdot s^{-\theta}u(s)\,\frac{ds}{s}\Big\|_{X_0}^{p_0}\frac{dt}{t}\Big)^{1/p_0} \\
&\leqslant \big\|t \mapsto t^{-\theta}\mathbf{1}_{(1,\infty)}(t)\big\|_{L^1(\mathbb{R}_+,\frac{dt}{t})}\big\|t \mapsto t^{-\theta}u(t)\big\|_{L^{p_0}(\mathbb{R}_+,\frac{dt}{t};X_0)} \\
&= \frac{1}{\theta}\big\|t \mapsto t^{-\theta}u(t)\big\|_{L^{p_0}(\mathbb{R}_+,\frac{dt}{t};X_0)} < \infty.
\end{aligned}$$

In the same way one checks that

$$\|t \mapsto t^{1-\theta}u_1(t)\|_{L^{p_1}(\mathbb{R}_+,\frac{dt}{t};X_1)} \leqslant \frac{1}{1-\theta}\big\|t \mapsto t^{1-\theta}u(t)\big\|_{L^{p_1}(\mathbb{R}_+,\frac{dt}{t};X_1)} < \infty.$$

This shows that $x \in (X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$, and by combining the above estimates we obtain the second estimate in (C.19).

In the converse direction, let $x \in (X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$ be given. Let the functions $u_j : \mathbb{R}_+ \to X_j$ $(j = 0, 1)$ be as in Definition C.3.4. Let $g : [0,\infty) \to \mathbb{R}$ be the 'tent shaped' piecewise linear function which is identically zero on $[0, 1]$ and $[3,\infty)$ and whose graph connects the points $(1, 0)$, $(2, 1)$ and $(3, 0)$ linearly. Let $\varphi(t) := tg(t)$. Then

$$\int_0^\infty \varphi(t)\,\frac{dt}{t} = \int_0^\infty g(t)\,dt = 1.$$

Define the functions $v_j \in C^1(0,\infty; X_j)$ $(j = 0, 1)$ by

$$v_j(t) := \int_0^\infty \varphi(t/\tau)u_j(\tau)\,\frac{d\tau}{\tau}, \quad t > 0.$$

Then $u_0 + u_1 \equiv x$ implies $v_0 + v_1 \equiv x$, and for all $r > 0$ we have

$$\|v_0(r)\|_{X_0} = \Big\|\int_{r/3}^r \varphi(r/\tau)u_0(\tau)\,\frac{d\tau}{\tau}\Big\|_{X_0} \leqslant \int_{r/3}^r \|u_0(\tau)\|_{X_0}\,\frac{d\tau}{\tau}$$

$$\leqslant \Big(\int_{r/3}^{r} \tau^{\theta p_0'}\,\frac{\mathrm{d}\tau}{\tau}\Big)^{1/p_0'}\Big(\int_0^\infty \tau^{-\theta p_0}\|u_0(\tau)\|_{X_0}^{p_0}\,\frac{\mathrm{d}\tau}{\tau}\Big)^{1/p_0},$$

from which it follows that $\lim_{r\downarrow 0}\|v_0(r)\|_{X_0} = 0$ if $p_0 > 1$. If $p_0 = 1$, this can be seen more directly. In the same way one sees that $\lim_{R\to\infty}\|v_1(R)\|_{X_1} = 0$.

Define $u \in C((0,\infty);X_0)\cap C((0,\infty);X_1) = C((0,\infty);X_0\cap X_1)$ by

$$u(t) := t v_0'(t) = -t v_1'(t), \quad t > 0.$$

From the above convergence properties it follows that

$$\begin{aligned}\int_0^\infty u(t)\,\frac{\mathrm{d}t}{t} &= \lim_{R\to\infty}\int_{1/R}^{R} v_0'(t)\,\mathrm{d}t = \lim_{R\to\infty}(v_0(R) - v_0(1/R))\\ &= \lim_{R\to\infty}(x - v_1(R) - v_0(1/R)) = x\end{aligned}$$

with convergence in $X_0 + X_1$.

For $j = 0,1$ we have

$$t^{j-\theta}\|u(t)\|_{X_j} = t^{1+j-\theta}\|v_j'(t)\|_{X_j} \leqslant \int_0^\infty (t/\tau)^{1+j-\theta}|\varphi'(t/\tau)|\tau^{j-\theta}\|u_j(\tau)\|_{X_j}\frac{\mathrm{d}\tau}{\tau}.$$

The integral on the right-hand side is a convolution as before. Since $|\varphi'(t)| = |g(t) + tg'(t)|$ is supported on $[1,3]$ and equals $2t-1$ on $[1,2]$ and $2t-3$ on $[2,3]$, an elementary computation gives

$$c_{0,\varphi} := \int_0^\infty t^{1-\theta}|\varphi'(t)|\frac{\mathrm{d}t}{t} \leqslant 2 \ \text{ and } \ c_{1,\varphi} := \int_0^\infty t^{2-\theta}|\varphi'(t)|\frac{\mathrm{d}t}{t} \leqslant 6.$$

Applying Young's inequality as before, we find

$$\Big(\int_0^\infty t^{(j-\theta)p_j}\|u(t)\|_{X_j}^{p_j}\,\frac{\mathrm{d}t}{t}\Big)^{1/p_j} \leqslant c_{j,\varphi}\Big(\int_0^\infty t^{(j-\theta)p_j}\|u_j(t)\|_{X_j}^{p_j}\Big)^{1/p_j}\frac{\mathrm{d}t}{t}.$$

If follows that $x \in (X_0,X_1)_{\theta,p_0,p_1}$, and by the above estimates the first estimate in (C.19) holds. □

We are now in a position to prove the following fundamental result.

Theorem C.3.14 (Lions–Peetre). *Let (X_0,X_1) be an interpolation couple of Banach spaces. Let $1 \leqslant p_0 \leqslant p_1 \leqslant \infty$, $0 < \theta < 1$, and $\frac{1}{p} := \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$. Then*

$$(X_0,X_1)_{\theta,p} = (X_0,X_1)^{(\mathrm{m})}_{\theta,p_0,p_1} = (X_0,X_1)_{\theta,p_0,p_1}$$

with equivalent norms satisfying

$$\frac{1}{12}\|x\|_{\theta,p_0,p_1} \leqslant \|x\|_{\theta,p} \leqslant 4C_\theta\|x\|_{\theta,p_0,p_1} \tag{C.20}$$

where $C_\theta = \max\{\theta^{-1},(1-\theta)^{-1}\}$, and

$$\frac{1}{2}\|x\|^{(\mathrm{m})}_{\theta,p_0,p_1} \leqslant \|x\|_{\theta,p} \leqslant 4\|x\|^{(\mathrm{m})}_{\theta,p_0,p_1}. \tag{C.21}$$

Proof. *Step 1* – First consider the case $1 \leqslant p_0 = p_1 = p \leqslant \infty$. By Propositions C.3.6 and C.3.13 we obtain that $(X_0, X_1)_{\theta,p} = (X_0, X_1)^{(\mathrm{m})}_{\theta,p,p} = (X_0, X_1)_{\theta,p,p}$, with equivalent norms satisfying (C.20).

Step 2 – The case $1 \leqslant p_0 \leqslant p_1 \leqslant \infty$ with $p_0 < \infty$ remains to be settled. We begin by noting that for $x \in X_0 \cap X_1$, the estimate (C.20) follows by combining Propositions C.3.7 and C.3.13. By Step 1 and Lemma C.3.12, $X_0 \cap X_1$ is dense in both $(X_0, X_1)_{\theta,p_0,p_1}$ and $(X_0, X_1)_{\theta,p}$. Thus the proof can be completed by a standard approximation argument.

Step 3 – Since $(X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1} = (X_0, X_1)_{\theta,p_0,p_1}$ by Proposition C.3.13 we can deduce from Steps 1 and 2 that also $(X_0, X_1)_{\theta,p} = (X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$ with equivalent norms, and then (C.11) and a density argument as below tells us that for all $x \in (X_0, X_1)_{\theta,p}$

$$\frac{1}{4}\|x\|_{\theta,p} \leqslant \|x\|^{(\mathrm{m})}_{\theta,p_0,p_1} \leqslant 2\|x\|_{\theta,p}$$

if $p_0 \leqslant p_1 \leqslant \infty$ and $p_0 < \infty$. The same holds if $p_0 = p_1 = \infty$ by Proposition C.3.6, and (C.21) follows. $\square$

From the above proof we infer:

Corollary C.3.15. *For all $1 \leqslant p < \infty$ and $0 < \theta < 1$, $X_0 \cap X_1$ is dense in $(X_0, X_1)_{\theta,p}$.*

Theorem C.3.16 (Real interpolation of operators, second version). *Let (X_0, X_1) and (Y_0, Y_1) be interpolation couples of Banach spaces. Suppose $T : X_0 + X_1 \to Y_0 + Y_1$ is a linear operator which maps X_0 into Y_0 and X_1 into Y_1 with norms*

$$\|T\|_{\mathscr{L}(X_0,Y_0)} = A_0, \qquad \|T\|_{\mathscr{L}(X_1,Y_1)} = A_1.$$

Then:

(1) *For all $1 \leqslant p \leqslant \infty$ and $0 < \theta < 1$ the operator T maps $(X_0, X_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$ into $(Y_0, Y_1)^{(\mathrm{m})}_{\theta,p_0,p_1}$ and we have*

$$\|T\|_{\mathscr{L}((X_0,X_1)^{(\mathrm{m})}_{\theta,p_0,p_1},(Y_0,Y_1)^{(\mathrm{m})}_{\theta,p_0,p_1})} \leqslant A_0^{1-\theta} A_1^{\theta}.$$

(2) *For all $1 \leqslant p \leqslant \infty$ and $0 < \theta < 1$ the operator T maps $(X_0, X_1)_{\theta,p_0,p_1}$ into $(Y_0, Y_1)_{\theta,p_0,p_1}$ and we have*

$$\|T\|_{\mathscr{L}((X_0,X_1)_{\theta,p_0,p_1},(Y_0,Y_1)_{\theta,p_0,p_1})} \leqslant A_0^{1-\theta} A_1^{\theta}.$$

Proof. We prove (1), the proof of (2) being entirely similar. First observe that a version of Lemma C.3.11 holds for the first mean method as well. Let us first assume that $A_0 \neq 0$. Fix $x \in (X_0, X_1)^{(\mathrm{m})}_{\theta,p}$ and write $x \equiv u_0 + u_1$ with $t^{j-\theta}u_j(t) \in L^{p_j}(\mathbb{R}_+, \frac{\mathrm{d}t}{t}; X_j)$. Then

$$\|t \mapsto t^{-\theta} T u_0(t)\|^{1-\theta}_{L^{p_j}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)} \cdot \|t \mapsto t^{1-\theta} T u_1(t)\|^{\theta}_{L^{p_1}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)}$$
$$\leqslant A_0^{1-\theta} A_1^{\theta} \|t \mapsto t^{-\theta} u_0(t)\|^{1-\theta}_{L^{p_0}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)} \cdot \|t \mapsto t^{1-\theta} u_1(t)\|^{\theta}_{L^{p_1}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)}.$$

Taking infima over all admissible u_0, u_1, the result follows from the above observation. □

C.4 Complex versus real

In this section we shall prove the inclusions

$$(X_0, X_1)_{\theta,1} \hookrightarrow [X_0, X_1]_\theta \hookrightarrow (X_0, X_1)_{\theta,\infty}$$

valid for arbitrary interpolation couples (X_0, X_1) of complex Banach spaces. With only a bit of additional effort a much sharper version of this result can be proved which takes into account the Fourier types of X_0 and X_1. Let $p \in [1, 2]$. A complex Banach space X is said to have *Fourier type* p if the Hausdorff-Young inequality holds for the Fourier transform of X-valued functions, that is, if the Fourier transform $\mathscr{F} : L^p(\mathbb{R}) \to L^{p'}(\mathbb{R})$ extends to a bounded operator $\mathscr{F} : L^p(\mathbb{R}; X) \to L^{p'}(\mathbb{R}; X)$, where $\frac{1}{p} + \frac{1}{p'} = 1$. Every Banach space X has Fourier type 1, and since the Plancherel theorem extends to Hilbert spaces, Hilbert spaces have Fourier type 2. Further examples of Banach spaces with non-trivial Fourier type (among them the L^p-spaces in the range $1 < p < \infty$) are given in Subsection 2.4.b, where the notion of Fourier type is studied in more detail.

Theorem C.4.1 (Peetre). *Let (X_0, X_1) be an interpolation couple of complex Banach spaces. If X_0 and X_1 have Fourier type p_0 and p_1, respectively, then for all $0 < \theta < 1$ we have continuous embeddings*

$$(X_0, X_1)_{\theta,p} \hookrightarrow [X_0, X_1]_\theta \hookrightarrow (X_0, X_1)_{\theta,p'}, \tag{C.22}$$

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

Corollary C.4.2. *For Hilbert couples, for all $0 < \theta < 1$ we have*

$$(H_0, H_1)_{\theta,2} = [H_0, H_1]_\theta \quad \textit{with equivalent norms.}$$

Proof of Theorem C.4.1. First we will show that $(X_0, X_1)_{\theta,p}$ embeds into $[X_0, X_1]_\theta$. By Corollary C.3.15 $X_0 \cap X_1$ is dense in $(X_0, X_1)_{\theta,p}$ since $p < \infty$, so by Theorem C.3.14 it suffices to show that there is a constant C such that for all $x \in X_0 \cap X_1$ we have

$$\|x\|_{[X_0,X_1]_\theta} \leqslant C \|x\|_{\theta,p_0,p_1}. \tag{C.23}$$

Fix a function u as in Definition C.3.10, let $u_k := \mathbf{1}_{(1/k,k)} u$, and set

$$f_k(z) := \int_0^\infty t^{z-\theta} u_k(t)\,\frac{\mathrm{d}t}{t}, \quad z \in \mathbb{S}.$$

This integral converges in $X_0 \cap X_1$ as follows in a similar way as for (C.17). As a function with values in $X_0 \cap X_1$, f is holomorphic on $\mathbb{S}$, the function $v \mapsto f_k(j+iv) \in X_j$ is continuous and it satisfies $\lim_{k\to\infty} f_k(\theta) = x$ in X_0+X_1. Substituting $t = e^{-2\pi s}$ and setting $v_k(s) := u_k(e^{-2\pi s})$ we calculate

$$\begin{aligned}\|t \mapsto t^{-\theta} u_k(t)\|_{L^{p_0}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)} &= \Big(\int_0^\infty t^{-\theta p_0}\|u_k(t)\|_{X_0}^{p_0}\,\frac{\mathrm{d}t}{t}\Big)^{1/p_0}\\ &= (2\pi)^{1/p_0}\Big(\int_{-\infty}^\infty e^{2\pi\theta p_0 s}\|v_k(s)\|_{X_0}^{p_0}\,\mathrm{d}s\Big)^{1/p_0}\\ &= (2\pi)^{1/p_0}\|s \mapsto e^{2\pi\theta s} v_k(s)\|_{L^{p_0}(\mathbb{R};X_0)}\end{aligned}$$

and similarly

$$\|t \mapsto t^{(1-\theta)} u_k(t)\|_{L^{p_1}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)} = (2\pi)^{1/p_1}\|s \mapsto e^{(\theta-1)s} v_k(s)\|_{L^{p_1}(\mathbb{R};X_1)}.$$

Moreover,

$$\frac{1}{2\pi} f_k(z) = \int_{-\infty}^\infty e^{-2\pi z s} e^{2\pi\theta s} v_k(s)\,\mathrm{d}s.$$

Taking $z = iy$ and $z = 1+iy$ with $y \in \mathbb{R}$, the Fourier type assumptions on X_0 and X_1 can be applied and we find, for $j \in \{0,1\}$,

$$\begin{aligned}\|f_k(j+i\cdot)\|_{L^{p_j'}(\mathbb{R};X_j)} &\leqslant 2\pi\varphi_{p_j,X_j}\|s \mapsto e^{2\pi(\theta-j)s} v_k(s)\|_{L^{p_j}(\mathbb{R};X_j)}\\ &= (2\pi)^{1/p_j'}\varphi_{p_j,X_j}\|t \mapsto t^{j-\theta} u_k(t)\|_{L^{p_j}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_j)}\\ &\leqslant (2\pi)^{1/p_j'}\varphi_{p_j,X_j}\|x\|_{\theta,p_0,p_1},\end{aligned} \tag{C.24}$$

with φ_{p_j,X_j} the norm of the Fourier transform $\mathscr{F} : L^{p_j}(\mathbb{R};X_j) \to L^{p_j'}(\mathbb{R};X_j)$. Therefore, by Corollary C.2.11,

$$\|x_k\|_{[X_0,X_1]_\theta} \leqslant C\|x\|_{\theta,p_0,p_1}, \tag{C.25}$$

where $C = C_{\theta,p_0,p_1}\max_{j\in\{0,1\}}(2\pi)^{1/p_j'}\varphi_{p_j,X_j}$.

By (C.24) with f_k replaced by $f_k - f_\ell$, another appeal to Corollary C.2.11 shows that

$$\|x_k - x_\ell\|_{[X_0,X_1]_\theta} \leqslant C\|t \mapsto t^{j-\theta}(u_k(t) - u_\ell(t))\|_{L^{p_j}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_j)}.$$

Since $p_0, p_1 < \infty$, the latter converges to zero as $k, \ell \to \infty$. Therefore, the sequence $(x_k)_{k\geqslant 1}$ is Cauchy in $[X_0,X_1]_\theta$ and therefore converges to a limit y in $[X_0,X_1]_\theta$. Since also $x = \lim_{k\to\infty} x_k = \lim_{k\to\infty} f_k(\theta)$ in X_0+X_1, we find that $\lim_{k\to\infty} x_k = x$ in $[X_0,X_1]_\theta$. Finally, another appeal to Corollary C.2.11 and letting $k \to \infty$ in (C.25) yields (C.23).

Turning to the proof of the second inclusion in (C.22), by Corollary C.2.8 it suffices to show that

$$\|x\|_{(X_0,X_1)_{\theta,p'}} \leqslant C\|f\|_{\mathscr{H}(X_0,X_1)} \tag{C.26}$$

for elements of the form $x = f(\theta) \in X_0 \cap X_1$ with $f \in \mathscr{H}_0(X_0, X_1; X_0 \cap X_1)$.

Fix an $x = f(\theta)$ of this form. The function $g(z) := e^{z^2-\theta^2} f(z)$ has the property that for all $x \in (0,1)$, $y \mapsto g(x+iy)$ is a Schwartz function with values in $X_0 \cap X_1$, it satisfies $g(\theta) = x$, and

$$\|g(j+2\pi i\cdot)\|_{L^{p_j}(\mathbb{R};X_j)} \leqslant C_{p_j}\|f\|_{\mathscr{H}(X_0,X_1)}, \tag{C.27}$$

with $C_{p_j} = \|e^{j-(2\pi\cdot)^2}\|_{L^{p_j}(\mathbb{R})} \leqslant 2$. Put

$$w(t) := \frac{1}{2\pi i}\int_{\theta-i\infty}^{\theta+i\infty} t^{\theta-z} g(z)\,\mathrm{d}z.$$

Then $w : (0,\infty) \to X_0 \cap X_1$ is continuous. Substituting $t = e^s$ and $z = \theta + 2\pi i r$, we find

$$\begin{aligned}\int_0^\infty w(t)\,\frac{\mathrm{d}t}{t} = \int_{\mathbb{R}} w(e^s)\,\mathrm{d}s &= \frac{1}{2\pi i}\int_{\mathbb{R}}\int_{\theta-i\infty}^{\theta+i\infty} e^{(\theta-z)s} g(z)\,\mathrm{d}z\,\mathrm{d}s \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} e^{-2\pi i s r} g(\theta+2\pi i r)\,\mathrm{d}s\,\mathrm{d}r = g(\theta) = x,\end{aligned}$$

where we applied the inversion formula for the Fourier transform for Schwartz functions in the last step (see Subsection 2.4.a). Moreover, by Cauchy's integral theorem, for all $\eta \in (0,1)$ we have

$$w(t) = \frac{1}{2\pi i}\int_{\eta-i\infty}^{\eta+i\infty} t^{\theta-z} g(z)\,\mathrm{d}z = t^\theta \int_{\mathbb{R}} t^{-2\pi i y} g(2\pi i y)\,\mathrm{d}y, \tag{C.28}$$

where in the last step we substituted $z = \eta + 2\pi i r$ and passed to the limit $\eta \downarrow 0$ in the norm of X_0. Substituting $t = e^s$, we find

$$\begin{aligned}\|t \mapsto t^{-\theta} w(t)\|_{L^{p_0'}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)} &= \Big\|t \mapsto \int_{\mathbb{R}} t^{-2\pi i y} g(2\pi i y)\,\mathrm{d}y\Big\|_{L^{p_0'}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_0)} \\ &= \Big\|s \mapsto \int_{\mathbb{R}} e^{-2\pi i y s} g(2\pi i y)\,\mathrm{d}y\Big\|_{L^{p_0'}(\mathbb{R};X_0)} \\ &= \|(g(2\pi i\cdot))\widehat{\ }\|_{p_0'} \\ &\leqslant \varphi_{p_0,X_0}\|g(2\pi i\cdot)\|_{p_0} \leqslant 2\varphi_{p_0,X_0}\|f\|_{\mathscr{H}(X_0,X_1)}\end{aligned}$$

using the Fourier type p_0 of X_0 and (C.27). Similarly, by letting $\eta \uparrow 1$ in (C.28),

$$\|t \mapsto t^{1-\theta} w(t)\|_{L^{p_1'}(\mathbb{R}_+,\frac{\mathrm{d}t}{t};X_1)} \leqslant \varphi_{p_1,X_1}\|g(1+2\pi i\cdot)\|_{p_1} \leqslant 2\varphi_{p_1,X_1}\|f\|_{\mathscr{H}(X_0,X_1)}.$$

Using Definition C.3.10, this implies

$$\|x\|_{\theta,p_0',p_1'} \leqslant 2\max\{\varphi_{p_0,X_0},\varphi_{p_1,X_1}\}\|f\|_{\mathscr{H}(X_0,X_1)}.$$

Now (C.26) follows from Theorem C.3.14, with constant

$$C = 8\max\{\varphi_{p_0,X_0},\varphi_{p_1,X_1}\}\max\big\{\theta^{-1},(1-\theta)^{-1}\big\}.$$

□

C.5 Notes

Most of the material of this appendix is standard and can be found in many textbooks, among which we mention Bergh and Löfström [1976], Kreĭn, Petunīn, and Semënov [1982], Lunardi [2009], Pisier [2016], Triebel [1978]. It barely scratches the surface of this vast subject, and our choices have been dictated by our needs to cover on interpolation of the Bochner spaces $L^p(S;X)$ (see Section 2.2). Important omissions include treatments of duality, reiteration, a fuller treatment of the trace method, as well as many important examples where the interpolation spaces can be explicitly computed. In this book we cover real and complex interpolation between $L^{p_0}(S;X_0)$ and $L^{p_1}(S;X_1)$ in Chapter 2, real interpolation between $L^1(S;X)$ and $W^{1,p}(S;X)$ in Chapter 2, and complex interpolation between $L^{p_0}(S;X)$ and $W^{1,p}(S;X)$ for UMD spaces in Chapter 5). For an arbitrary Banach spaces X and $\theta\in(0,1)$, the real method gives (see König [1986, Proposition 3.b.5])

$$(L^1(\mathbb{R}^d;X),W^{1,p}(\mathbb{R}^d;X))_{\theta,q} = B^{\theta}_{p,q}(\mathbb{R}^d;X), \qquad 1\leqslant p,q\leqslant\infty,$$

where $B^s_{p,q}$ is the Besov scale (see Triebel [1978, Theorem 1.18.6.1])

$$(L^1(S;X),L^\infty(S;X))_{\theta,q} = L^{1/(1-\theta),q}(S;X), \qquad 1\leqslant q\leqslant\infty,$$

where $L^{p,q}$ is the Lorentz scale (see Triebel [1978, Theorem 4.5.2.1])

$$(C_{\rm b}(\mathbb{R}^d;X),C^1_{\rm b}(\mathbb{R}^d;X))_{\theta,\infty} = (BUC(\mathbb{R}^d;X),BUC^1(\mathbb{R}^d;X))_{\theta,\infty} = C^\theta_{\rm b}(\mathbb{R}^d;X),$$

where $C^s_{\rm b}$ is the bounded Hölder scale. The same result holds with $\mathbb{R}^d$ replaced by open domains in $\mathbb{R}^d$ whose boundary is uniformly C^1 (all equalities up to equivalent norms). Interpolation of vector-valued BMO and Hardy spaces is discussed in Blasco and Xu [1991].

Section C.1

The results of this section are standard.

Section C.2

The results of Section C.2 are due to Calderón [1964], whose treatment we follow.

Section C.3

Our treatment of real interpolation follows Triebel [1978]. Proposition C.3.7 is due to Peetre [1963]. Our proof is taken from Holmstedt [1970]. Theorem C.3.14 is due to Lions and Peetre [1964].

Section C.4

Theorem C.4.1 is due to Peetre [1969]. Our presentation is based on García-Cuerva, Kazaryan, Kolyada, and Torrea [1998]. A related version of this theorem has been obtained in Suárez and Weis [2006] by the so-called γ-interpolation method; instead of using Fourier type, this result is stated in terms of the type and cotype of X_0 and X_1. For an introduction to this interpolation method and some of its applications we refer the reader to Kalton, Kunstmann, and Weis [2006].

D

Schatten classes

The Schatten classes $\mathscr{C}^p$, defined and discussed in this appendix, are certain "non-commutative analogues" of the familiar sequence spaces ℓ^p. One could more generally define *non-commutative L^p-spaces*, but since our aim is to use these spaces only as examples rather than developing their theory in a systematic way, we restrict the considerations to the classes $\mathscr{C}^p$.

D.1 Approximation numbers and Schatten classes

Let us consider bounded linear operators on a Hilbert space H. The nth *approximation number* (also called the nth *singular value*) of an operator u is defined as

$$a_n(u) := \inf\{\|u-v\| : \dim \mathsf{R}(v) < n\}.$$

Then clearly

$$\|u\| = a_1(u) \geqslant \ldots \geqslant a_{n-1}(u) \geqslant a_n(u) \geqslant \ldots,$$

and $a_n(u) \to 0$ if and only if u is a compact operator. We will introduce certain subspaces of compact operators by requiring that the sequence $(a_n(u))_{n=1}^\infty$ is in some subspace of c_0. Before doing this, however, it is convenient to know a little about the behaviour of the approximation numbers $a_n(u)$.

Recall that every compact u has the *singular value decomposition*

$$u = \sum_{k=1}^\infty \tau_k (e_k|\cdot) f_k, \qquad \tau_1 \geqslant \ldots \geqslant \tau_{k-1} \geqslant \tau_k \geqslant \ldots \geqslant 0. \tag{D.1}$$

where $(e_k)_{k=1}^\infty$ and $(f_k)_{k=1}^\infty$ are orthonormal sequences. We use this word in a slightly extended sense, meaning that $(e_k|e_\ell) = \delta_{k\ell}\|e_k\|^2$, and $\|e_k\|^2 \in \{0,1\}$, where the value zero is allowed only when also $\tau_k = 0$. In this way, the finite dimensional case (when no infinite orthonormal sequences in the usual sense exist) is conveniently treated at the same time.

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 63, DOI 10.1007/978-3-319-48520-1

Lemma D.1.1. *If u has a singular value decomposition* (D.1)*, then $\tau_k = a_k(u)$.*

Proof. With $u_n := \sum_{k=1}^{n-1} \tau_k (e_k|\cdot) f_k$, one checks that $a_n(u) \leqslant \|u - u_n\| = \tau_n$. Conversely, if v is any operator with $\dim \mathsf{R}(v) < n$, then one can find a unit vector $\xi = \sum_{k=1}^{n} \xi_k e_k \in \mathsf{N}(v)$, and hence

$$\|u - v\| \geqslant \|(u - v)\xi\| = \|u\xi\| = \Big(\sum_{k=1}^{n} \tau_k^2 |\xi_k|^2\Big)^{1/2} \geqslant \tau_n,$$

and taking the infimum over all such v it follows that $a_n(u) \geqslant \tau_n$. ☐

Let us collect some basic properties of the approximation numbers. Since u^* and u^*u have the representations

$$u^* = \sum_{k=1}^{\infty} \tau_k (f_k|\cdot) e_k, \qquad u^*u = \sum_{k=1}^{\infty} \tau_k^2 (e_k|\cdot) e_k,$$

it follows that

$$a_n(u^*) = a_n(u), \quad a_n(u^*u) = a_n(u)^2, \qquad \forall\, n \in \mathbb{Z}_+. \tag{D.2}$$

Since $\dim \mathsf{R}(wv) \leqslant \dim \mathsf{R}(v)$ for any operators v and w, we deduce

$$a_n(wu) \leqslant \inf\{\|wu - wv\| : \dim \mathsf{R}(v) < n\} \leqslant \|w\| a_n(u),$$

and combining the above two results, also

$$a_n(vuw) \leqslant \|v\| a_n(w^*u^*) \leqslant \|v\| \|w^*\| a_n(u^*) = \|v\| a_n(u) \|w\|. \tag{D.3}$$

Finally,

$$a_n(u) = \inf_v \|u - v\| \leqslant \inf_v \|w - v\| + \|u - w\| = a_n(w) + \|u - w\|$$

which, combined with the similar estimate with the roles of u and w reversed, shows that

$$|a_n(u) - a_n(w)| \leqslant \|u - w\|. \tag{D.4}$$

Lemma D.1.2. *For any compact u, there holds*

$$\sum_{j=1}^{n} a_j(u) = \max \Big| \sum_{j=1}^{n} (h_j|ug_j) \Big| = \max \sum_{j=1}^{n} |(h_j|ug_j)|,$$

where the maxima are taken over all orthonormal sequences $(g_j)_{j=1}^n$ and $(h_j)_{j=1}^n$. The existence of the maxima, rather than just suprema, is part of the assertion.

Proof. Let u have the singular value decomposition (D.1). Then $a_j(u) = (f_j|ue_j)$, so the two maxima (or *a priori*, suprema) are at least as large as the left side of the claim. Now consider arbitrary orthonormal sequences g_j and h_j. Then

$$\sum_{j=1}^{n} |(h_j|ug_j)| \leqslant \sum_{k=1}^{\infty} \tau_k \sum_{j=1}^{n} |(e_k|g_j)(h_j|f_k)| =: \sum_{k=1}^{\infty} \tau_k b_k.$$

By Cauchy–Schwarz one checks that

$$0 \leqslant b_k \leqslant \|e_k\|\|f_k\| \leqslant 1, \qquad \sum_{k=1}^{\infty} b_k \leqslant \Big(\sum_{j=1}^{n} \|g_j\|^2\Big)^{1/2} \Big(\sum_{j=1}^{n} \|h_j\|^2\Big)^{1/2} \leqslant n.$$

Writing $\sigma_j := \tau_j - \tau_{j+1} \geqslant 0$, we conclude that

$$\begin{aligned} \sum_{k=1}^{\infty} \tau_k b_k &= \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} \sigma_j b_k = \sum_{j=1}^{\infty} \sigma_j \sum_{k=1}^{j} b_k \\ &\leqslant \sum_{j=1}^{\infty} \sigma_j \sum_{k=1}^{j \wedge n} 1 = \sum_{k=1}^{n} \sum_{j=k}^{\infty} \sigma_j = \sum_{k=1}^{n} \tau_k. \end{aligned}$$

□

Proposition D.1.3. *For compact operators u and v,*

$$\sum_{k=1}^{n} a_k(u+v) \leqslant \sum_{k=1}^{n} \big(a_k(u) + a_k(v)\big).$$

Proof. By Lemma D.1.2, taking the same maxima as there,

$$\begin{aligned} LHS &= \max \sum_{k=1}^{n} |(h_k|(u+v)g_k)| \\ &\leqslant \max \sum_{k=1}^{n} |(h_k|ug_k)| + \max \sum_{k=1}^{n} |(h_k|vg_k)| = RHS. \end{aligned}$$

□

With the following elementary inequality, the estimate of Proposition D.1.3 implies various further inequalities for the approximation numbers:

Lemma D.1.4. *Let $f : \mathbb{R} \to \mathbb{R}$ be a convex non-decreasing function, and let the sequences*

$$a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n, \qquad b_1 \geqslant b_2 \geqslant \ldots \geqslant b_n$$

satisfy

$$\sum_{j=1}^{k} a_j \leqslant \sum_{j=1}^{k} b_j$$

for all $k = 1, \ldots, n$. *Then*

$$\sum_{j=1}^{n} f(a_j) \leqslant \sum_{j=1}^{n} f(b_j).$$

Proof. We may assume that $a_j \neq b_j$ for all j. Namely, both the assumptions and the conclusions remain unchanged on the removal of the equal entries from the sequences.

The claim says that

$$0 \leqslant \sum_{j=1}^{n} (f(b_j) - f(a_j)) = \sum_{j=1}^{n} c_j (b_j - a_j),$$

where

$$c_j := \frac{f(b_j) - f(a_j)}{b_j - a_j}.$$

Since f is increasing, we have $c_j \geqslant 0$, whereas the convexity implies that $c_j \geqslant c_{j+1}$; namely, we have $a_j \geqslant a_{j+1}$ and $b_j \geqslant b_{j+1}$, and the definition of convexity easily implies that the difference quotient $(f(b) - f(a))/(b - a)$ is increasing in both a and b. Denoting $A_k := \sum_{j=1}^{k} a_j$ and $B_k := \sum_{j=1}^{k} b_j$, we have also assumed that $A_k \leqslant B_k$ for all k.

Now the claim follows from a simple calculation, putting together the information just discussed:

$$\begin{aligned}
\sum_{j=1}^{n} c_j (b_j - a_j) &= \sum_{j=1}^{n} c_j ((B_j - B_{j-1}) - (A_j - A_{j-1})) \\
&= \sum_{j=1}^{n} c_j (B_j - A_j) - \sum_{j=1}^{n} c_j (B_{j-1} - A_{j-1}) \\
&= c_n (B_n - A_n) + \sum_{j=1}^{n-1} (c_j - c_{j+1})(B_j - A_j) \geqslant 0,
\end{aligned}$$

since $c_n \geqslant 0$, $c_j - c_{j+1} \geqslant 0$, and $B_j - A_j \geqslant 0$ for all j. □

We are ready to introduce:

Definition D.1.5. *For a Hilbert space* H *and* $p \in [1, \infty)$, *the corresponding Schatten class is*

$$\mathscr{C}^p(H) := \{u \in \mathscr{K}(H) : \|u\|_p := \|(a_n(u))_{n=1}^{\infty}\|_{\ell^p} < \infty\},$$

where $\mathscr{K}(H)$ *is the space of compact linear operators on* H.

We will often abbreviate $\mathscr{C}^p := \mathscr{C}^p(H)$, when the choice of Hilbert space is either clear from the context or irrelevant for a particular statement.

Theorem D.1.6. *For a Hilbert space H and $p \in [1, \infty)$, the Schatten class $\mathscr{C}^p(H)$ is a Banach space.*

Proof. The triangle inequality

$$\begin{aligned} \Big(\sum_{k=1}^{\infty} a_k(u+v)^p \Big)^{1/p} &\leqslant \Big(\sum_{k=1}^{\infty} \big(a_k(u) + a_k(v) \big)^p \Big)^{1/p} \\ &\leqslant \Big(\sum_{k=1}^{\infty} a_k(u)^p \Big)^{1/p} + \Big(\sum_{k=1}^{\infty} a_k(v)^p \Big)^{1/p} \end{aligned}$$

follows from Proposition D.1.3 combined with Lemma D.1.4, and the usual triangle inequality of ℓ^p in the second step.

Let us next prove the completeness. Since $\|u\|_p \geqslant a_1(u) = \|u\|$, the Cauchy condition $\|u_m - u_n\|_p \to 0$ implies $\|u_m - u_n\| \to 0$, and hence $u_n \to u$ (in the operator norm) for some compact operator u. It remains to be checked that $u \in \mathscr{C}^p$ and that the limit is also attained with respect to the corresponding norm.

Writing out the Cauchy condition, for every $\varepsilon > 0$ there exists an $N = N(\varepsilon) \in \mathbb{Z}_+$ such that for all $n, m > N$ and all $k \in \mathbb{Z}_+$ there holds

$$\sum_{j=1}^{k} a_j(u_n - u_m)^p \leqslant \|u_n - u_m\|_p^p \leqslant \varepsilon.$$

Keeping k and n fixed, we pass to the limit $m \to \infty$. Then $u_m \to u$ in the operator norm and hence $a_j(u_n - u_m) \to a_j(u_n - u)$ by (D.4). Thus

$$\sum_{j=1}^{k} a_j(u_n - u)^p \leqslant \varepsilon, \qquad \forall\ k \in \mathbb{Z}_+, \forall\ n > N(\varepsilon).$$

With $k \to \infty$, it follows that $\|u_n - u\|_p^p \leqslant \varepsilon$ for all $n > N(\varepsilon)$; thus $u \in \mathscr{C}^p$ by the triangle inequality and $u_n \to u$ in $\mathscr{C}^p$ as $n \to \infty$. □

Some basic properties of the Schatten norms follow at once from corresponding properties of the approximation numbers. Indeed, (D.3) implies the *ideal property*

$$\|vuw\|_p \leqslant \|v\| \, \|u\|_p \|w\|, \tag{D.5}$$

and from (D.2) we deduce that

$$\|u\|_p = \|u^*\|_p = \|u^*u\|_{p/2}^{1/2}. \tag{D.6}$$

The case $p = 2$ of this identity has an interesting consequence:

Proposition D.1.7. *For any Hilbert space H, the space $\mathscr{C}^2(H)$ is again a Hilbert space with the inner product*

$$(v|u)_{\mathscr{C}^2} := \sum_{i\in I}(vh_i|uh_i),$$

where $(h_i)_{i\in I}$ is any orthonormal basis of H.

Proof. If u has the singular value decomposition (D.1), then $\|u\|_2^2 = \|u^*u\|_1 = \sum_{k=1}^{\infty}\tau_k^2$. On the other hand, if $(h_i)_{i\in I}$ is any orthonormal basis, then

$$\sum_{i\in I}\|uh_i\|^2 = \sum_{i\in I}(h_i|u^*uh_i) = \sum_{i\in I}\sum_{k=1}^{\infty}\tau_k^2|(e_k|h_i)|^2 = \sum_{k=1}^{\infty}\tau_k^2.$$

This shows the convergence of the left side, its independence of the chosen orthonormal basis, and then by polarisation the same conclusions follow for the corresponding inner product as defined in the assertion. □

D.2 Hölder's inequality and duality

Besides the linear structure of the Schatten classes established in Theorem D.1.6, these spaces also exhibit interesting multiplicative phenomena, which are reflections of corresponding properties of the approximation numbers. Let us recall the following fact from linear algebra:

Lemma D.2.1. *If u is an operator on an n-dimensional Hilbert space, then*

$$|\det(u)| = \prod_{k=1}^{n} a_k(u).$$

Proof. First, $|\det(u)|^2 = \overline{\det(u)}\det u = \det u^*\det u = \det(u^*u)$, and u^*u is represented in the basis $(e_k)_{k=1}^n$ by the diagonal matrix with diagonal entries $a_1(u)^2,\dots,a_n(u)^2$. □

In arbitrary Hilbert spaces we obtain the following inequality between finite determinants:

Lemma D.2.2. *For $u\in\mathscr{K}(H)$ and arbitrary vectors $\phi_1,\dots,\phi_n\in H$, there holds*

$$\det(u\phi_i|u\phi_j)_{i,j=1}^n \leqslant \prod_{k=1}^{n} a_k(u)^2\det(\phi_i|\phi_j)_{i,j=1}^n,$$

where equality is attained (e.g., when $\phi_k = e_k$), the vectors from a singular value decomposition of u.

Proof. Notice that both determinants involve non-negative matrices, so the determinants are some non-negative numbers. The case of the equality is clear, since $(ue_i|ue_j)_{i,j=1}^n = \delta_{ij}a_j(u)^2$.

Let some $\phi_1, \dots, \phi_n$ be given, and $h_1, \dots, h_n$ be orthonormal vectors so that $\operatorname{span}\{h_1, \dots, h_n\} \supseteq \operatorname{span}\{\phi_1, \dots, \phi_n\}$. Then

$$(u\phi_i|u\phi_j) = \sum_{\ell,m=1}^{n} (\phi_i|h_\ell)(uh_\ell|uh_m)(h_m|\phi_j)$$

is the (i,j) entry of the product of the three matrices

$$(\phi_i|h_\ell)_{i,\ell=1}^n, \quad (uh_\ell|uh_m)_{\ell,m=1}^n, \quad (h_m|\phi_j)_{m,j=1}^n.$$

By the product rule of determinants, then

$$\det(u\phi_i|u\phi_j)_{i,j} = \det(uh_\ell|uh_m)_{\ell,m} \times \big(\det(\phi_i|h_\ell)_{i,\ell} \times \det(h_m|\phi_j)_{m,j}\big),$$

and the latter product is equal to the determinant of the product matrix

$$\Big(\sum_{\ell=1}^{n}(\phi_i|h_\ell)(h_\ell|\phi_j)\Big)_{i,j=1}^n = (\phi_i|\phi_j)_{i,j=1}^n.$$

On the other hand, $(uh_\ell|uh_m)_{\ell,m=1}^n$ is the representation in the basis $(h_\ell)_{\ell=1}^n$ of the operator $\pi u^* u \pi^*$, where $\pi : H \to \operatorname{span}\{h_1, \dots, h_n\}$ is the orthogonal projection. By Lemma D.2.1, (D.2), and (D.3),

$$\begin{aligned}\det(uh_\ell|uh_m)_{\ell,m=1}^n = \det(\pi u^* u \pi^*) &= \prod_{k=1}^{n} a_k(\pi u^* u \pi^*)\\ &\leqslant \prod_{k=1}^{n} a_k(u^* u) = \prod_{k=1}^{n} a_k(u)^2.\end{aligned}$$

Substituting back, the proof is completed. □

Proposition D.2.3. *For compact operators u and v,*

$$\prod_{k=1}^{n} a_k(uv) \leqslant \prod_{k=1}^{n} a_k(u)a_k(v).$$

Proof. Taking below the maximum over all sequences of vectors $(\phi_k)_{k=1}^n$ with $\det(\phi_i|\phi_j)_{i,j=1}^n \leqslant 1$ and using Lemma D.2.2 three times, we find that

$$\begin{aligned}LHS^2 &= \max \det(uv\phi_i|uv\phi_j)_{i,j=1}^n\\ &\leqslant \prod_{k=1}^{n} a_k(u)^2 \max \det(v\phi_i|v\phi_j)_{i,j=1}^n = RHS^2.\end{aligned}$$

□

Corollary D.2.4. *Let $p, q, r \in [1, \infty)$ with $1/p = 1/q + 1/r$. For $u \in \mathscr{C}^q$, $v \in \mathscr{C}^r$, we have $uv \in \mathscr{C}^p$ with $\|uv\|_p \leqslant \|u\|_q \|v\|_p$. Conversely, every $w \in \mathscr{C}^p$ admits such a factorisation $w = uv$ with equality above.*

Proof. Taking logarithms of the estimate of Proposition D.2.3 and applying Lemma D.1.4 with the convex increasing function $f(t) = e^{pt}$, and then the usual Hölder's inequality, we get

$$\Big(\sum_{j=1}^{n} a_j(uv)^p\Big)^{1/p} \leqslant \Big(\sum_{j=1}^{n} a_j(u)^p a_j(v)^p\Big)^{1/p}$$
$$\leqslant \Big(\sum_{j=1}^{n} a_j(u)^q\Big)^{1/q} \Big(\sum_{j=1}^{n} a_j(v)^r\Big)^{1/r}.$$

For the factorisation of $w = \sum_{k=1}^{\infty} \tau_k (e_k|\cdot) f_k$, just take

$$u = \sum_{k=1}^{\infty} \tau_k^{p/q} (g_k|\cdot) f_k, \qquad v = \sum_{k=1}^{\infty} \tau_k^{p/r} (e_k|\cdot) g_k$$

for any orthonormal sequence $(g_k)_{k=1}^{\infty}$. □

Proposition D.2.5. *The functional*

$$\operatorname{tr} : \mathscr{C}^1(H) \to \mathbb{C}, \quad u \mapsto \operatorname{tr}(u) := \sum_{i \in I} (h_i|uh_i),$$

where $(h_i)_{i \in I}$ is any orthonormal basis of H, is well defined and contractive.

Proof. Any $u \in \mathscr{C}^1$ has a factorisation $u = v^* w$ with $v, w \in \mathscr{C}^2$ and $\|u\|_1 = \|v\|_2 \|w\|_2$. Then the expression defining $\operatorname{tr}(u)$ is just $(v|w)_{\mathscr{C}^2}$, which is independent of the chosen orthonormal basis. Since the terms of the series $(vh_i|wh_i) = (h_i|uh_i)$ are independent of the particular factorisation of u as $v^* w$, also the sum of the series is. Finally,

$$|\operatorname{tr}(u)| = |(v|w)_{\mathscr{C}^2}| \leqslant \|v\|_2 \|w\|_2 = \|u\|_1.$$

□

Notice that for $x, y \in H$,

$$\operatorname{tr}\big((x|\cdot)y\big) = \sum_{i \in I} (x|h_i)(h_i|y) = (x|y).$$

Theorem D.2.6. *For any Hilbert space H and $p \in (1, \infty)$, the mapping*

$$\Lambda : \mathscr{C}^{p'}(H) \to (\mathscr{C}^p(H))^*, \quad \langle \Lambda v, u \rangle := \operatorname{tr}(vu)$$

is an isometric isomorphism. This is also true for $p \in \{1, \infty\}$ when interpreted as

$$\Lambda : \mathscr{C}^1(H) \to (\mathscr{K}(H))^*, \qquad \Lambda : \mathscr{L}(H) \to (\mathscr{C}^1(H))^*.$$

Proof. First, Λ is well defined and contractive, since

$$|\operatorname{tr}(vu)| \leqslant \|vu\|_1 \leqslant \|v\|_{p'}\|u\|_p,$$

where $\|v\|_\infty := \|v\|$ is the usual operator norm in either $\mathscr{K}(H)$ or $\mathscr{L}(H)$. Testing $v = \sum_{k=1}^\infty \tau_k (e_k|\cdot) f_k \in \mathscr{C}^{p'}$, $p' \in [1,\infty)$, with $u = \sum_{k=1}^\infty \tau_k^{p'-1}(f_k|\cdot)e_k$ gives

$$\operatorname{tr}(vu) = \sum_{k=1}^\infty (f_k|vuf_k) = \sum_{k=1}^\infty \tau_k^{p'} = \|v\|_{p'}^{p'} = \|u\|_p^p = \|v\|_{p'}\|u\|_p$$

showing that Λ is isometric. For $v \in \mathscr{L}(H)$ we have even more simply with $u = (x|\cdot)y$ that $\operatorname{tr}(vu) = (x|vy)$, so taking the supremum over $\|x\| = \|y\| = 1$ shows that $\|\Lambda v\|_{(\mathscr{C}^1)^*} = \|v\|$. It remains to be shown that Λ is onto.

Let $\phi \in (\mathscr{C}^p)^*$, $p \in [1,\infty)$, or else $\phi \in (\mathscr{K}(H))^*$, and define on $H \times H$ the bilinear form $b(x,y) := \phi\big((x|\cdot)y\big)$. Since $\|(x|\cdot)y\|_p = \|x\|\|y\|$, it follows that $|b(x,y)| \leqslant \|\phi\|_{(\mathscr{C}^p)^*}\|x\|\|y\|$, and hence b is represented by a bounded linear operator, $b(x,y) = (x|vy)$ with $\|v\| \leqslant \|\phi\|_{(\mathscr{C}^p)^*}$. For $p = 1$, this is already all that we claimed, since now

$$\phi((x|\cdot)y) = b(x,y) = (x|vy) = \operatorname{tr}(v(x|\cdot)y) = \langle \Lambda v, (x|\cdot)y\rangle. \tag{D.7}$$

By linearity, the functionals ϕ and Λv coincide on all finite rank operators, and by density and continuity on their full domain.

For $p \in (1,\infty]$ (the end point meaning $\phi \in (\mathscr{K}(H))^*$), we still need to check that $v \in \mathscr{C}^{p'}$, in particular that it is compact. Fix an $\varepsilon > 0$ and continue to choose pairs (f_k, g_k) of unit vectors, $k \in \mathbb{Z}_+$, such that $(g_k|vf_k) \geqslant \varepsilon$ where f_k is orthogonal to $\operatorname{span}(f_j)_{j=1}^{k-1}$ and g_k to $\operatorname{span}(g_j)_{j=1}^{k-1}$, as long as this is possible. This process will stop after finitely many steps, since if $(f_k, g_k)_{k=1}^n$ are chosen in this way, there holds

$$\varepsilon n \leqslant \sum_{k=1}^n (g_k|vf_k) = \phi\Big(\sum_{k=1}^n (g_k|\cdot)f_k\Big) \leqslant \|\phi\|_{(\mathscr{C}^p)^*} n^{1/p},$$

so $n \leqslant (\|\phi\|_{(\mathscr{C}^p)^*}/\varepsilon)^{p'}$.

Let $(f_k)_{k=1}^n$ and $(g_k)_{k=1}^n$ be the maximal orthonormal sequences as chosen above, and let π_1 and π_2 be the orthogonal projections onto their respective linear spans. By the required maximality, any unit vectors $x \in \mathsf{R}(I-\pi_1)$ and $y \in \mathsf{R}(I-\pi_2)$ will satisfy $|(y|vx)| < \varepsilon$. This means that $\|(I-\pi_2)v(I-\pi_1)\| \leqslant \varepsilon$. But

$$(I-\pi_2)v(I-\pi_1) = v - (\pi_2 v + v\pi_1 - \pi_2 v \pi_1)$$

is the difference of v from a certain finite rank operator. We just showed that its norm can be made smaller than a given $\varepsilon > 0$, and hence v is in the closure of the finite rank operators, i.e., it is compact.

Thus v has the singular value decomposition $v = \sum_{k=1}^\infty \tau_k(e_k|\cdot)f_k$. But then

$$\tau_k = (f_k|ve_k) = b(f_k, e_k) = \phi\big((f_k|\cdot)e_k\big),$$

and hence

$$\sum_{k=1}^{n} \tau_k^{p'} = \phi\Big(\sum_{k=1}^{n} \tau_k^{p'-1}(f_k|\cdot)e_k\Big) \leqslant \|\phi\|_{(\mathscr{C}^p)^*}\Big(\sum_{k=1}^{n} \tau_k^{(p'-1)p}\Big)^{1/p},$$

which upon simplification gives

$$\Big(\sum_{k=1}^{n} \tau_k^{p'}\Big)^{1/p'} \leqslant \|\phi\|_{(\mathscr{C}^p)^*}.$$

Passing to the limit $n \to \infty$ shows that $\|v\|_{p'} \leqslant \|\phi\|_{(\mathscr{C}^p)^*}$. Now the proof is completed by repeating the observation (D.7) and the sentence after it to check that indeed $\phi = \Lambda v$ and hence Λ is onto, as we wanted to prove. □

D.3 Interpolation

We have already seen that the spaces $\mathscr{C}^p$ share many of the properties of the familiar ℓ^p spaces, although in many cases this required a new argument. Here is one more analogy, which is obtained as an essentially direct consequence of the corresponding property of the ℓ^p spaces. We write $\mathscr{C}^\infty(H) := \mathscr{K}(H)$.

Proposition D.3.1. *For any Hilbert space H and numbers $1 \leqslant p_0, p_1 \leqslant \infty$, $0 < \theta < 1$ and $1/p = (1-\theta)/p_0 + \theta/p_1$, there holds*

$$[\mathscr{C}^{p_0}(H), \mathscr{C}^{p_1}(H)]_\theta = \mathscr{C}^p(H),$$

isometrically.

Proof. In accordance with the definition of $\mathscr{C}^\infty$ as the space of compact operators, the space ℓ^∞ should be (inside this proof only, in order to simplify writing) interpreted as c_0.

Suppose first that $u \in \mathscr{C}^p$; thus $u = \sum_{k=1}^\infty \tau_k(e_k|\cdot)f_k$ with $\tau := (\tau_k)_{k=1}^\infty \in \ell^p = [\ell^{p_0}, \ell^{p_1}]_\theta$. By the definition of the complex interpolation space and using the notation introduced for this purpose, this means that $\tau = T(\theta)$ for some $T \in \mathscr{H}(\ell^{p_0}, \ell^{p_1})$ with $\|T\|_{\mathscr{H}(\ell^{p_0},\ell^{p_1})} \leqslant (1+\varepsilon)\|\tau\|_{\ell^p}$. Taking $U(s) := \sum_{k=1}^\infty T_k(s)(e_k|\cdot)f_k$, we have $U \in \mathscr{H}(\mathscr{C}^{p_0}, \mathscr{C}^{p_1})$, $U(\theta) = u$, and hence $u \in [\mathscr{C}^{p_0}, \mathscr{C}^{p_1}]_\theta$ with

$$\begin{aligned}\|u\|_{[\mathscr{C}^{p_0},\mathscr{C}^{p_1}]_\theta} &\leqslant \|U\|_{\mathscr{H}(\mathscr{C}^{p_0},\mathscr{C}^{p_1})} \\ &= \|T\|_{\mathscr{H}(\ell^{p_0},\ell^{p_1})} \leqslant (1+\varepsilon)\|\tau\|_{\ell^p} = (1+\varepsilon)\|u\|_{\mathscr{C}^p}.\end{aligned}$$

Let conversely $u \in [\mathscr{C}^{p_0}, \mathscr{C}^{p_1}]_\theta$. Then $u = U(\theta)$ for some $U \in \mathscr{H}(\mathscr{C}^{p_0}, \mathscr{C}^{p_1})$ with the usual norm estimate. At every $s \in S$, there is a singular value decomposition $U(S) = \sum_{k=1}^\infty \tau_k(s)(e_k(s)|\cdot)f_k(s)$. We write simply τ_k, e_k and f_k for $\tau_k(\theta)$, $e_k(\theta)$ and $f_k(\theta)$.

Motivated by the fact that $\tau_k = (f_k|U(\theta)e_k)$, we define

$$T_k(s) := (f_k|U(s)e_k) = \sum_{j=1}^{\infty} \tau_j(s)(f_k|f_j(s))(e_j(s)|e_k).$$

Thus the sequence $T(s) := (T_k(s))_{k=1}^{\infty}$ is obtained from the sequence $\tau(s) := (\tau_j(s))_{j=1}^{\infty}$ by multiplying with the infinite matrix $A(s) := (\alpha_{kj}(s))_{k,j=1}^{\infty}$ with $\alpha_{kj}(s) = (f_k|f_j(s))(e_j(s)|e_k)$. This matrix satisfies

$$\sum_{j=1}^{\infty} |\alpha_{kj}(s)| \leqslant \Big(\sum_{j=1}^{\infty} |(f_k|f_j(s))|^2\Big)^{1/2}\Big(\sum_{j=1}^{\infty} |(e_j(s)|e_k)|^2\Big)^{1/2} \leqslant \|f_k\|\|e_k\| \leqslant 1,$$

and similarly $\sum_{k=1}^{\infty} |\alpha_{kj}(s)| \leqslant \|f_j(s)\|\|e_j(s)\| \leqslant 1$. Multiplication by such a matrix contracts all ℓ^q spaces, $q \in [1,\infty]$; indeed

$$\sum_{k=1}^{\infty}\Big|\sum_{j=1}^{\infty}\alpha_{kj}y_j\Big|^q \leqslant \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}|\alpha_{kj}||y_j|^q = \sum_{j=1}^{\infty}|y_j|^q\sum_{k=1}^{\infty}|\alpha_{kj}| \leqslant \sum_{j=1}^{\infty}|y_j|^q$$

with an even easier estimate when $q = \infty$.

We conclude that $T \in \mathscr{H}(\ell^{p_0}, \ell^{p_1})$; the size estimates for $T(s)$ follow from those for $\tau(s)$, and the analyticity (which could fail for $\tau(s)$) from the analyticity of $U(s)$ by the representation of T in terms of it. Hence $\tau = T(\theta) \in [\ell^{p_0}, \ell^{p_1}]_\theta = \ell^p$ and thus $u = \sum_{k=1}^{\infty} \tau_k(e_k|\cdot)f_k \in \mathscr{C}^p$ with

$$\|u\|_{\mathscr{C}^p} = \|\tau\|_{\ell^p} \leqslant \|T\|_{\mathscr{H}(\ell^{p_0},\ell^{p_1})} \leqslant \|U\|_{\mathscr{H}(\mathscr{C}^{p_0},\mathscr{C}^{p_1})} \leqslant (1+\varepsilon)\|u\|_{[\mathscr{C}^{p_0},\mathscr{C}^{p_1}]_\theta}.$$

□

D.4 Notes

This material is quite standard and can be found in many textbooks. Our presentation is adapted from Diestel, Jarchow, and Tonge [1995] and Gohberg, Goldberg, and Krupnik [2000].

The singular value decomposition is also called the *Schmidt representation*, e.g., in Gohberg, Goldberg, and Krupnik [2000].

References

Adams, M. and V. Guillemin. *Measure theory and probability.* Birkhäuser Boston Inc., Boston, MA, 1996. Corrected reprint of the 1986 original. 515

Adams, R. A. and J. J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics.* Elsevier/Academic Press, Amsterdam, second edition, 2003. 161

Alaoglu, L. Weak topologies of normed linear spaces. *Ann. of Math. (2)*, 41: 252–267, 1940. 533

Albiac, F. and N. J. Kalton. *Topics in Banach space theory*, volume 233 of *Graduate Texts in Mathematics.* Springer, New York, 2006. 354, 363, 533

Aldous, D. J. Unconditional bases and martingales in $L_p(F)$. *Math. Proc. Cambridge Philos. Soc.*, 85(1):117–123, 1979. 356, 363

Aliprantis, C. D. and O. Burkinshaw. *Positive operators.* Springer, Dordrecht, 2006. Reprint of the 1985 original. 156

Almeida, V., J. J. Betancor, and A. J. Castro. UMD Banach spaces and the maximal regularity for the square root of several operators. *Semigroup Forum*, 88(1):21–51, 2014. 491, 492

Alspach, D. and E. Odell. L_p spaces. In *Handbook of the geometry of Banach spaces, Vol. I*, pages 123–159. North-Holland, Amsterdam, 2001. 154

Amann, H. *Linear and quasilinear parabolic problems. Vol. I: Abstract linear theory*, volume 89 of *Monographs in Mathematics.* Birkhäuser Boston Inc., Boston, MA, 1995. 159, 160, 161

Amann, H. *Anisotropic function spaces and maximal regularity for parabolic problems. Part 1: Function spaces.* Jindřich Nečas Center for Mathematical Modeling Lecture Notes, 6. Matfyzpress, Prague, 2009. 489

Arendt, W. and S. Bu. The operator-valued Marcinkiewicz multiplier theorem and maximal regularity. *Math. Z.*, 240(2):311–343, 2002. 490

Arendt, W. and N. Nikolski. Vector-valued holomorphic functions revisited. *Math. Z.*, 234(4):777–805, 2000. 534

Arendt, W. and N. Nikolski. Addendum: "Vector-valued holomorphic functions revisited" [Math. Z. **234** (2000), no. 4, 777–805]. *Math. Z.*, 252(3): 687–689, 2006. 534

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 63, DOI 10.1007/978-3-319-48520-1

Arendt, W., C. J. K. Batty, M. Hieber, and F. Neubrander. *Vector-valued Laplace transforms and Cauchy problems*, volume 96 of *Monographs in Mathematics*. Birkhäuser/Springer Basel AG, Basel, second edition, 2011. 61, 65, 161

Asmar, N. H. and B. P. Kelly. Vector-valued transference and maximal ergodic theory in UMD-valued function spaces. *Collect. Math.*, 47(1):63–75, 1996. 475

Asmar, N. H., E. Berkson, and T. A. Gillespie. Representations of groups with ordered duals and generalized analyticity. *J. Funct. Anal.*, 90(1):206–235, 1990. 475

Asmar, N. H., B. P. Kelly, and S. Montgomery-Smith. A note on UMD spaces and transference in vector-valued function spaces. *Proc. Edinburgh Math. Soc. (2)*, 39(3):485–490, 1996. 474

Astala, K., T. Iwaniec, and G. Martin. *Elliptic partial differential equations and quasiconformal mappings in the plane*, volume 48 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 2009. 371

Astala, K., T. Iwaniec, I. Prause, and E. Saksman. Burkholder integrals, Morrey's problem and quasiconformal mappings. *J. Amer. Math. Soc.*, 25 (2):507–531, 2012. 371

Badrikian, A., G. W. Johnson, and I. Yoo. The composition of operator-valued measurable functions is measurable. *Proc. Amer. Math. Soc.*, 123 (6):1815–1820, 1995. 63

Banach, S. *Théorie des opérations linéaires*. Warszawa, Mathematisches Seminar der Univ. Warschau (Monografje Matematyczne I), 1932. 533, 534

Bañuelos, R. and P. Janakiraman. L^p-bounds for the Beurling-Ahlfors transform. *Trans. Amer. Math. Soc.*, 360(7):3603–3612, 2008. 494

Bates, S., W. B. Johnson, J. Lindenstrauss, D. Preiss, and G. Schechtman. Affine approximation of Lipschitz functions and nonlinear quotients. *Geom. Funct. Anal.*, 9(6):1092–1127, 1999. 489

Bauer, H. *Measure and integration theory*, volume 26 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 2001. 515

Benyamini, Y. and J. Lindenstrauss. *Geometric nonlinear functional analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2000. 65

Bergh, J. and J. Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223. 158, 562

Berkson, E. and T. A. Gillespie. Spectral decompositions and harmonic analysis on UMD spaces. *Studia Math.*, 112(1):13–49, 1994. 475

Berkson, E. and T. A. Gillespie. An $\mathfrak{M}_q(\mathbb{T})$-functional calculus for power-bounded operators on certain UMD spaces. *Studia Math.*, 167(3):245–257, 2005. 494

Berkson, E., T. A. Gillespie, and P. S. Muhly. Abstract spectral decompositions guaranteed by the Hilbert transform. *Proc. London Math. Soc. (3)*, 53(3):489–517, 1986. 356, 475

Berkson, E., T. A. Gillespie, and P. S. Muhly. A generalization of Macaev's theorem to noncommutative L^p-spaces. *Integral Equations Operator Theory*, 10(2):164–186, 1987. 356

Berkson, E., T. A. Gillespie, and P. S. Muhly. Generalized analyticity in UMD spaces. *Ark. Mat.*, 27(1):1–14, 1989. 475

Berkson, E., T. A. Gillespie, and J. L. Torrea. Vector valued transference. In P. Liu, editor, *Functional space theory and its applications (Wuhan, 2003)*, pages 1–27. Research Information Ltd. UK, Hemel Hempstead, 2004. 482

Bessaga, C. and A. Pełczyński. On bases and unconditional convergence of series in Banach spaces. *Studia Math.*, 17:151–164, 1958. 61

Betancor, J. J., A. J. Castro, J. Curbelo, J. C. Fariña, and L. Rodríguez-Mesa. γ-radonifying operators and UMD-valued Littlewood-Paley-Stein functions in the Hermite setting on BMO and Hardy spaces. *J. Funct. Anal.*, 263 (12):3804–3856, 2012. 492

Betancor, J. J., A. J. Castro, J. Curbelo, and L. Rodríguez-Mesa. Characterization of UMD Banach spaces by imaginary powers of Hermite and Laguerre operators. *Complex Anal. Oper. Theory*, 7(4):1019–1048, 2013a. 492

Betancor, J. J., R. Crescimbeni, J. C. Fariña, and L. Rodríguez-Mesa. Multipliers and imaginary powers of the Schrödinger operators characterizing UMD Banach spaces. *Ann. Acad. Sci. Fenn. Math.*, 38(1):209–227, 2013b. 492

Betancor, J. J., A. J. Castro, J. Curbelo, J. C. Fariña, and L. Rodríguez-Mesa. Square functions in the Hermite setting for functions with values in UMD spaces. *Ann. Mat. Pura Appl. (4)*, 193(5):1397–1430, 2014. 492

Betancor, J. J., A. J. Castro, and L. Rodríguez-Mesa. UMD-valued square functions associated with Bessel operators in Hardy and BMO spaces. *Integral Equations Operator Theory*, 81(3):319–374, 2015. 492

Betancor, J. J., A. J. Castro, and L. Rodriguez-Mesa. Square functions and spectral multipliers for Bessel operators in UMD spaces. *Banach J. Math. Analysis*, 10(2):338–384, 2016. 492

Billard, P. Sur la convergence presque partout des séries de Fourier-Walsh des fonctions de l'espace $L^2(0, 1)$. *Studia Math.*, 28:363–388, 1966/1967. 362

Billard, P., S. Kwapień, A. Pelczyński, and Ch. Samuel. Biorthogonal systems of random unconditional convergence in Banach spaces. In *Texas Functional Analysis Seminar 1985–1986 (Austin, TX, 1985–1986)*, Longhorn Notes, pages 13–35. Univ. Texas, Austin, TX, 1986. 355

Billingsley, P. *Convergence of probability measures*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. 515

Birkhoff, G. Integration of functions with values in a Banach space. *Trans. Amer. Math. Soc.*, 38(2):357–378, 1935. 63

Birman, M. Sh. and M. Z. Solomyak. Double Stieltjes operator integrals. II. *Probl. Mat. Fiz.*, 2:26–60, 1967. 482

Blasco, O. Hardy spaces of vector-valued functions: duality. *Trans. Amer. Math. Soc.*, 308(2):495–507, 1988. 65, 478

Blasco, O. and J. M. A. M. van Neerven. Spaces of operator-valued functions measurable with respect to the strong operator topology. In *Vector measures, integration and related topics*, volume 201 of *Oper. Theory Adv. Appl.*, pages 65–78. Birkhäuser Verlag, Basel, 2010. 64

Blasco, O. and Q. H. Xu. Interpolation between vector-valued Hardy spaces. *J. Funct. Anal.*, 102(2):331–359, 1991. 562

Blower, G. A multiplier characterization of analytic UMD spaces. *Studia Math.*, 96(2):117–124, 1990. 490

Bochner, S. *Abstrakte Funktionen und die Besselsche Ungleichung.* 1933a. in: S. Bochner, Collected papers. Part 1, American Mathematical Society, Providence, RI, 1992. 156

Bochner, S. Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind. *Fund. Math.*, 20:262–276, 1933b. 59, 61

Bochner, S. and A. E. Taylor. Linear functionals on certain spaces of abstractly-valued functions. *Ann. of Math. (2)*, 39(4):913–944, 1938. 157

Bogachev, V. I. *Gaussian measures*, volume 62 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 1998. 59

Bogachev, V. I. *Measure theory. Vol. I.* Springer-Verlag, Berlin, 2007a. 515

Bogachev, V. I. *Measure theory. Vol. II.* Springer-Verlag, Berlin, 2007b. 61, 515

Bohnenblust, H. F. and A. Sobczyk. Extensions of functionals on complex linear spaces. *Bull. Amer. Math. Soc.*, 44(2):91–93, 1938. 533

Bourgain, J. Some remarks on Banach spaces in which martingale difference sequences are unconditional. *Ark. Mat.*, 21(2):163–168, 1983. 354, 469, 474, 492

Bourgain, J. Vector valued singular integrals and the H^1-BMO duality. In *Israel seminar on geometrical aspects of functional analysis (1983/84)*, pages XVI, 23. Tel Aviv Univ., Tel Aviv, 1984a. 359, 471, 479, 480, 481, 485

Bourgain, J. Extension of a result of Benedek, Calderón and Panzone. *Ark. Mat.*, 22(1):91–95, 1984b. 261, 357, 481

Bourgain, J. On martingales transforms in finite-dimensional lattices with an appendix on the K-convexity constant. *Math. Nachr.*, 119:41–53, 1984c. 363

Bourgain, J. On square functions on the trigonometric system. *Bull. Soc. Math. Belg. Sér. B*, 37(1):20–26, 1985. 481

Bourgain, J. Real isomorphic complex Banach spaces need not be complex isomorphic. *Proc. Amer. Math. Soc.*, 96(2):221–226, 1986a. 536

Bourgain, J. Vector-valued singular integrals and the H^1-BMO duality. In *Probability theory and harmonic analysis (Cleveland, Ohio, 1983)*, volume 98 of *Monogr. Textbooks Pure Appl. Math.*, pages 1–19. Dekker, New York, 1986b. 356, 359, 471, 479, 480, 481, 483, 485

Bourgain, J. Vector-valued Hausdorff-Young inequalities and applications. In *Geometric aspects of functional analysis (1986/87)*, volume 1317 of *Lecture Notes in Math.*, pages 239–249. Springer, Berlin, 1988. 160

Bourgain, J. and W. J. Davis. Martingale transforms and complex uniform convexity. *Trans. Amer. Math. Soc.*, 294(2):501–515, 1986. 359

Bourgin, R. D. *Geometric aspects of convex sets with the Radon-Nikodým property*, volume 993 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1983. 65

Brezis, H. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011. 161, 533

Brzeźniak, Z. Stochastic partial differential equations in M-type 2 Banach spaces. *Potential Anal.*, 4(1):1–45, 1995. 264, 499

Brzeźniak, Z. On stochastic convolution in Banach spaces and applications. *Stochastics Stochastics Rep.*, 61(3-4):245–295, 1997. 264

Brzeźniak, Z. and J. M. A. M. van Neerven. Stochastic convolution in separable Banach spaces and the stochastic linear Cauchy problem. *Studia Math.*, 143(1):43–74, 2000. 533

Brzeźniak, Z. and S. Peszat. Maximal inequalities and exponential estimates for stochastic convolutions in Banach spaces. In *Stochastic processes, physics and geometry: new interplays, I (Leipzig, 1999)*, volume 28 of *CMS Conf. Proc.*, pages 55–64. Amer. Math. Soc., Providence, RI, 2000. 264

Bu, S. Operator-valued Fourier multipliers on multi-dimensional Hardy spaces. *Chin. Ann. Math. Ser. B*, 32(2):293–302, 2011. 490

Bu, S. and C. Le Merdy. H^p-maximal regularity and operator valued multipliers on Hardy spaces. *Canad. J. Math.*, 59(6):1207–1222, 2007. 490

Bukhvalov, A. V. The duality of functors that are generated by spaces of vector-valued functions. *Izv. Akad. Nauk SSSR Ser. Mat.*, 39(6):1284–1309, 1438, 1975. 154, 155

Bukhvalov, A. V. Continuity of operators in spaces of vector-valued functions; applications to the theory of bases. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 157:5–22, 178, 1987. 356

Burkholder, D. L. Maximal inequalities as necessary conditions for almost everywhere convergence. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 3:75–88 (1964), 1964. 160

Burkholder, D. L. Martingale transforms. *Ann. Math. Statist.*, 37:1494–1504, 1966. 263, 354

Burkholder, D. L. Distribution function inequalities for martingales. *Ann. Probab.*, 1:19–42, 1973. 263

Burkholder, D. L. A geometrical characterization of Banach spaces in which martingale difference sequences are unconditional. *Ann. Probab.*, 9(6):997–1011, 1981a. 263, 354, 355, 359

Burkholder, D. L. Martingale transforms and the geometry of Banach spaces. In *Probability in Banach spaces, III (Medford, Mass., 1980)*, volume 860 of *Lecture Notes in Math.*, pages 35–50. Springer, Berlin-New York, 1981b. 367, 368

Burkholder, D. L. A nonlinear partial differential equation and the unconditional constant of the Haar system in L^p. *Bull. Amer. Math. Soc. (N.S.)*, 7(3):591–595, 1982. 494

Burkholder, D. L. A geometric condition that implies the existence of certain singular integrals of Banach-space-valued functions. In *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)*, Wadsworth Math. Ser., pages 270–286. Wadsworth, Belmont, CA, 1983. 354, 364, 469

Burkholder, D. L. Boundary value problems and sharp inequalities for martingale transforms. *Ann. Probab.*, 12(3):647–702, 1984. 355, 359, 367

Burkholder, D. L. Martingales and Fourier analysis in Banach spaces. In *Probability and analysis (Varenna, 1985)*, volume 1206 of *Lecture Notes in Math.*, pages 61–108. Springer, Berlin, 1986. 355, 366, 367, 368

Burkholder, D. L. Explorations in martingale theory and its applications. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 1–66. Springer, Berlin, 1991. 358, 367

Burkholder, D. L. Martingales and singular integrals in Banach spaces. In *Handbook of the geometry of Banach spaces, Vol. I*, pages 233–269. North-Holland, Amsterdam, 2001. 262, 265, 354, 497

Burkholder, D. L. and R. F. Gundy. Extrapolation and interpolation of quasi-linear operators on martingales. *Acta Math.*, 124(1):249–304, 1970. 263

Burkholder, D. L., R. F. Gundy, and M. L. Silverstein. A maximal function characterization of the class H^p. *Trans. Amer. Math. Soc.*, 157:137–153, 1971. 477

Burkholder, D. L., B. J. Davis, and R. F. Gundy. Integral inequalities for convex functions of operators on martingales. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 223–240. Univ. California Press, Berkeley, Calif., 1972. 259

Calderón, A.-P. Intermediate spaces and interpolation, the complex method. *Studia Math.*, 24:113–190, 1964. 159, 563

Calderón, A.-P. and A. Zygmund. On the existence of certain singular integrals. *Acta Math.*, 88:85–139, 1952. 159, 262, 480

Carleson, L. On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116:135–157, 1966. 362, 476, 496

Carleson, L. On the Littlewood–Paley theorem. *Report (Mittag–Leffler Institute)*, 1967. 481

Caspers, M., S. Montgomery Smith, D. Potapov, and F. Sukochev. The best constants for operator Lipschitz functions on Schatten classes. *J. Funct. Anal.*, 267(10):3557–3579, 2014. 485

Chang, S.-Y. A., J. M. Wilson, and T. H. Wolff. Some weighted norm inequalities concerning the Schrödinger operators. *Comment. Math. Helv.*, 60(2): 217–246, 1985. 260

Chatterji, S. D. *Martingales of Banach-valued random variables.* PhD thesis, Michigan State University, 1960. 259, 261, 262

Chatterji, S. D. A note on the convergence of Banach-space valued martingales. *Math. Ann.*, 153:142–149, 1964. 259, 261, 262

Chatterji, S. D. Martingale convergence and the Radon-Nikodym theorem in Banach spaces. *Math. Scand.*, 22:21–41, 1968. 262

Christ, M. A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral. *Colloq. Math.*, 60/61(2):601–628, 1990. 261

Chung, K. L. and R. J. Williams. *Introduction to stochastic integration.* Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, second edition, 1990. 164

Clarkson, J. A. Uniformly convex spaces. *Trans. Amer. Math. Soc.*, 40:396–414, 1936. 159

Clément, Ph. and B. de Pagter. Some remarks on the Banach space valued Hilbert transform. *Indag. Math. (N.S.)*, 2(4):453–460, 1991. 471

Clément, Ph. and J. Prüss. An operator-valued transference principle and maximal regularity on vector-valued L_p-spaces. In *Evolution equations and their applications in physical and life sciences (Bad Herrenalb, 1998)*, volume 215 of *Lecture Notes in Pure and Appl. Math.*, pages 67–87. Dekker, New York, 2001. 479

Clément, Ph., B. de Pagter, F. A. Sukochev, and H. Witvliet. Schauder decompositions and multiplier theorems. *Studia Math.*, 138(2):135–163, 2000. 355, 359, 362

Cobos, F. Some spaces in which martingale difference sequences are unconditional. *Bull. Polish Acad. Sci. Math.*, 34(11-12):695–703 (1987), 1986. 356

Cobos, F. and D. L. Fernandez. Hardy-Sobolev spaces and Besov spaces with a function parameter. In *Function spaces and applications (Lund, 1986)*, volume 1302 of *Lecture Notes in Math.*, pages 158–170. Springer, Berlin, 1988. 356

Coifman, R. R. A real variable characterization of H^p. *Studia Math.*, 3(51): 269–274, 1974. 478

Coifman, R. R. and G. Weiss. *Analyse harmonique non-commutative sur certains espaces homogènes.* Lecture Notes in Mathematics, Vol. 242. Springer-Verlag, Berlin-New York, 1971. Étude de certaines intégrales singulières. 260

Coifman, R. R. and G. Weiss. *Transference methods in analysis.* American Mathematical Society, Providence, R.I., 1976. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, No. 31. 475

Coifman, R. R., J. L. Rubio de Francia, and S. Semmes. Multiplicateurs de Fourier de $L^p(\mathbf{R})$ et estimations quadratiques. *C. R. Acad. Sci. Paris Sér. I Math.*, 306(8):351–354, 1988. 482

Conde, J. M. A note on dyadic coverings and nondoubling Calderón-Zygmund theory. *J. Math. Anal. Appl.*, 397(2):785–790, 2013. 260

Conway, J. B. *A course in functional analysis*, volume 96 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1990. 533

Córdoba, A. Some remarks on the Littlewood-Paley theory. In *Proceedings of the Seminar on Harmonic Analysis (Pisa, 1980)*, number suppl. 1, pages 75–80, 1981. 481

Cox, S. G. and S. Geiss. On decoupling in Banach spaces. *Preprint*, 2016. 364

Cox, S. G. and M. C. Veraar. Some remarks on tangent martingale difference sequences in L^1-spaces. *Electron. Comm. Probab.*, 12:421–433, 2007. 366

Cox, S. G. and M. C. Veraar. Vector-valued decoupling and the Burkholder-Davis-Gundy inequality. *Illinois J. Math.*, 55(1):343–375 (2012), 2011. 358, 366, 499

Cruz-Uribe, D. V. Two weight norm inequalities for fractional integral operators and commutators. *arXiv:1412.4157*, 2014. 260

Cruz-Uribe, D. V., J. M. Martell, and C. Pérez. *Weights, extrapolation and the theory of Rubio de Francia*, volume 215 of *Operator Theory: Advances and Applications*. Birkhäuser/Springer Basel AG, Basel, 2011. 157, 261

David, G. and J.-L. Journé. A boundedness criterion for generalized Calderón-Zygmund operators. *Ann. of Math. (2)*, 120(2):371–397, 1984. 471

David, G., J.-L. Journé, and S. Semmes. Opérateurs de Calderón-Zygmund, fonctions para-accrétives et interpolation. *Rev. Mat. Iberoam.*, 1(4):1–56, 1985. 472

Davies, E. B. Lipschitz continuity of functions of operators in the Schatten classes. *J. London Math. Soc. (2)*, 37(1):148–157, 1988. 484

Davis, B. On the integrability of the martingale square function. *Israel J. Math.*, 8:187–190, 1970. 263

Davis, W. J., N. Ghoussoub, W. B. Johnson, S. Kwapien, and B. Maurey. Weak convergence of vector valued martingales. In *Probability in Banach spaces 6*, pages 41–50. Springer, 1990. 262

Defant, A. and K. Floret. *Tensor norms and operator ideals*, volume 176 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1993. 62

Defant, M. On the vector-valued Hilbert transform. *Math. Nachr.*, 141:251–265, 1989. 157

Denk, R., M. Hieber, and J. Prüss. $\mathcal{R}$-boundedness, Fourier multipliers and problems of elliptic and parabolic type. *Mem. Amer. Math. Soc.*, 166(788): viii+114, 2003. 485

Dettweiler, E. On the martingale problem for Banach space valued stochastic differential equations. *J. Theoret. Probab.*, 2(2):159–191, 1989. 499

Dettweiler, E. Stochastic integration relative to Brownian motion on a general Banach space. *Doğa Mat.*, 15(2):58–97, 1991. 264, 499

Di Nezza, E., G. Palatucci, and E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573, 2012. 161

Di Plinio, F. and Y. Ou. Banach-valued multilinear singular integrals. *arXiv:1506.05827*, 2015. 266, 500

Diestel, J. and J. J. Uhl, Jr. *Vector measures*. American Mathematical Society, Providence, R.I., 1977. 59, 61, 62, 65, 161, 259, 534

Diestel, J., H. Jarchow, and A. Tonge. *Absolutely summing operators*, volume 43 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. 575

Diestel, J., J. H. Fourie, and J. Swart. *The metric theory of tensor products. Grothendieck's résumé revisited.* American Mathematical Society, Providence, RI, 2008. 62

Dieudonné, J. Complex structures on real Banach spaces. *Proc. Amer. Math. Soc.*, 3:162–164, 1952. 535

Dieudonné, J. *History of functional analysis*, volume 49 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam-New York, 1981. 533

Dilworth, S. J. and M. Girardi. Bochner vs. Pettis norm: examples and results. In *Banach spaces (Mérida, 1992)*, volume 144 of *Contemp. Math.*, pages 69–80. Amer. Math. Soc., Providence, RI, 1993. 61, 63

Dilworth, S. J. and M. Girardi. Nowhere weak differentiability of the Pettis integral. *Quaestiones Math.*, 18(4):365–380, 1995. 61

Dinculeanu, N. *Vector integration and stochastic integration in Banach spaces.* Pure and Applied Mathematics (New York). Wiley-Interscience, New York, 2000. 59

Dodds, P. G., T. K. Dodds, and B. de Pagter. Fully symmetric operator spaces. *Integral Equations Operator Theory*, 15(6):942–972, 1992. 495

Doob, J. L. *Stochastic processes.* John Wiley & Sons Inc., New York, 1953. 259, 262

Doob, J. L. *Classical potential theory and its probabilistic counterpart*, volume 262 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, New York, 1984. 259

Dore, G. and A. Venni. On the closedness of the sum of two closed operators. *Math. Z.*, 196(2):189–201, 1987. 472

Dulst, D. van. *Reflexive and superreflexive Banach spaces*, volume 102 of *Mathematical Centre Tracts*. Mathematisch Centrum, Amsterdam, 1978. 363

Dulst, D. van. *Characterizations of Banach spaces not containing l^1*, volume 59 of *CWI Tract*. Stichting Mathematisch Centrum Centrum voor Wiskunde en Informatica, Amsterdam, 1989. 61

Dunford, N. Integration in general analysis. *Trans. Amer. Math. Soc.*, 37(3): 441–453, 1935. 61

Dunford, N. and J. T. Schwartz. *Linear operators. I: General theory.* Pure and Applied Mathematics, Vol. 7. Interscience Publishers, Inc., New York; Interscience Publishers, Ltd., London, 1958. 59, 61, 533

Edgar, G. A. Measurability in a Banach space. *Indiana Univ. Math. J.*, 26 (4):663–677, 1977. 60

Edgar, G. A. Measurability in a Banach space. II. *Indiana Univ. Math. J.*, 28(4):559–579, 1979. 60

Enflo, P. Banach spaces which can be given an equivalent uniformly convex norm. In *Proceedings of the International Symposium on Partial Differential*

Equations and the Geometry of Normed Linear Spaces (Jerusalem, 1972), volume 13, pages 281–288 (1973), 1972. 264, 363

Evans, L. C. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010. 161

Fefferman, C. Characterizations of bounded mean oscillation. *Bull. Amer. Math. Soc.*, 77(4):587–588, 1971. 65

Fefferman, C. and E. M. Stein. Some maximal inequalities. *Amer. J. Math.*, 93:107–115, 1971. 261

Fernandez, D. L. and J. B. Garcia. Interpolation of Orlicz-valued function spaces and U.M.D. property. *Studia Math.*, 99(1):23–40, 1991. 356

Figiel, T. On equivalence of some bases to the Haar system in spaces of vector-valued functions. *Bull. Polish Acad. Sci. Math.*, 36(3-4):119–131 (1989), 1988. 360, 469, 470

Figiel, T. Singular integral operators: a martingale approach. In *Geometry of Banach spaces (Strobl, 1989)*, volume 158 of *London Math. Soc. Lecture Note Ser.*, pages 95–110. Cambridge Univ. Press, Cambridge, 1990. 469, 470, 471

Figiel, T. and P. Wojtaszczyk. Special bases in function spaces. In *Handbook of the geometry of Banach spaces, Vol. I*, pages 561–597. North-Holland, Amsterdam, 2001. 359

Fourie, J. H. On operator-valued measurable functions. In *Vector measures, integration and related topics*, volume 201 of *Oper. Theory Adv. Appl.*, pages 205–214. Birkhäuser Verlag, Basel, 2010. 64

Fremlin, D. H. Borel sets in non-separable Banach spaces. *Hokkaido Math. J.*, 9:179–188, 1980. 60

Fremlin, D. H. *Measure theory. Vol. 1: The irreducible minimum.* Torres Fremlin, Colchester, 2000. Corrected third printing in 2004. 515

Fremlin, D. H. *Measure theory. Vol. 2: Broad foundations.* Torres Fremlin, Colchester, 2001. Corrected second printing in 2003. 515

Fremlin, D. H. *Measure theory. Vol. 3: Measure algebras.* Torres Fremlin, Colchester, 2002. Corrected second printing in 2004. 515

Fremlin, D. H. *Measure theory. Vol. 4: Topological measure spaces. Part I, II.* Torres Fremlin, Colchester, 2003. 515

García-Cuerva, J. and J. L. Rubio de Francia. *Weighted norm inequalities and related topics*, volume 116 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1985. 156, 157

García-Cuerva, J., R. Macías, and J. L. Torrea. The Hardy-Littlewood property of Banach lattices. *Israel J. Math.*, 83(1-2):177–201, 1993. 261, 357

García-Cuerva, J., K. S. Kazaryan, V. I. Kolyada, and J. L. Torrea. The Hausdorff-Young inequality with vector-valued coefficients and applications. *Uspekhi Mat. Nauk*, 53(3(321)):3–84, 1998. translation in Russian Math. Surveys **53** (1998), no. 3, 435–513. 161, 563

Garling, D. J. H. Brownian motion and UMD-spaces. In *Probability and Banach spaces (Zaragoza, 1985)*, volume 1221 of *Lecture Notes in Math.*, pages 36–49. Springer, Berlin, 1986. 358, 364, 497, 499

Garling, D. J. H. On martingales with values in a complex Banach space. *Math. Proc. Cambridge Philos. Soc.*, 104(2):399–406, 1988. 372, 490

Garling, D. J. H. Random martingale transform inequalities. In *Probability in Banach spaces 6 (Sandbjerg, 1986)*, volume 20 of *Progr. Probab.*, pages 101–119. Birkhäuser Boston, Boston, MA, 1990. 358, 363, 499

Garling, D. J. H. *Inequalities: a journey into linear analysis.* Cambridge University Press, Cambridge, 2007. 156, 260

Garsia, A. M. *Martingale inequalities: Seminar notes on recent progress.* W. A. Benjamin, Inc., Reading, Mass.-London-Amsterdam, 1973. Mathematics Lecture Notes Series. 260

Geiss, S. BMO_ψ-spaces and applications to extrapolation theory. *Studia Math.*, 122(3):235–274, 1997. 263

Geiss, S. A counterexample concerning the relation between decoupling constants and UMD-constants. *Trans. Amer. Math. Soc.*, 351(4):1355–1375, 1999. 358, 498

Geiss, S. and P. F. X. Müller. Extrapolation of vector-valued rerrangement operators. *J. Lond. Math. Soc. (2)*, 80(3):798–814, 2009. 360

Geiss, S., S. Montgomery-Smith, and E. Saksman. On singular integral and martingale transforms. *Trans. Amer. Math. Soc.*, 362(2):553–575, 2010. 474, 475, 489, 494, 497

Ghoussoub, N., G. Godefroy, B. Maurey, and W. Schachermayer. Some topological and geometrical structures in Banach spaces. *Mem. Amer. Math. Soc.*, 70(378):iv+116, 1987. 65

Gijswijt, D. and J. M. A. M. van Neerven. On the constant in a transference inequality for the vector-valued Fourier transform. *Austr. J. Math. Anal. Appl.*, 13:1–11, 2016. Article 2. 160

Girardi, M. and L. W. Weis. Operator-valued Fourier multiplier theorems on $L_p(X)$ and geometry of Banach spaces. *J. Funct. Anal.*, 204(2):320–354, 2003. 485

Girardi, M. and L. W. Weis. Operator-valued martingale transforms and R-boundedness. *Illinois J. Math.*, 49(2):487–516 (electronic), 2005. 263, 360

Gohberg, I. C. and M. G. Kreĭn. *Theory and applications of Volterra operators in Hilbert space.* Translations of Mathematical Monographs, Vol. 24. American Mathematical Society, Providence, R.I., 1970. 484

Gohberg, I. C., S. Goldberg, and N. Ya. Krupnik. *Traces and determinants of linear operators*, volume 116 of *Operator Theory: Advances and Applications.* Birkhäuser Verlag, Basel, 2000. 575

Gohberg, I.C. and N. Ya. Krupnik. Norm of the Hilbert transformation in the L_p-space. *Funkts. Anal. Prilozh.*, 2(2):91–92, 1968. 484

Grafakos, L. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics.* Springer, New York, second edition, 2008. 479, 489

Grosse-Erdmann, K.-G. On the Borel-Okada theorem and the Hadamard multiplication theorem. *Complex Variables Theory Appl.*, 22(1-2):101–112, 1993. 534

Guerre-Delabrière, S. Some remarks on complex powers of $(-\Delta)$ and UMD spaces. *Illinois J. Math.*, 35(3):401–407, 1991. 474, 492

Gundy, R. F. A decomposition for L^1-bounded martingales. *Ann. Math. Statist.*, 39:134–138, 1968. 262

Haase, M. H. A. *The functional calculus for sectorial operators*, volume 169 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 2006. 491

Haase, M. H. A. Convexity inequalities for positive operators. *Positivity*, 11 (1):57–68, 2007. 163

Haase, M. H. A. The complex inversion formula revisited. *J. Aust. Math. Soc.*, 84(1):73–83, 2008. 476

Haase, M. H. A. A transference principle for general groups and functional calculus on UMD spaces. *Math. Ann.*, 345(2):245–265, 2009. 475

Hahn, H. Über lineare Gleichungssysteme in linearen Räumen. *J. Reine Angew. Math.*, 157:214–229, 1927. 533

Haller, R., H. Heck, and A. Noll. Mikhlin's theorem for operator-valued Fourier multipliers in n variables. *Math. Nachr.*, 244:110–130, 2002. 480, 485

Hänninen, T. S. and T. P. Hytönen. The A_2 theorem and the local oscillation decomposition for Banach space valued functions. *J. Operator Theory*, 72 (1):193–218, 2014. 474

Hänninen, T. S. and T. P. Hytönen. Operator-valued dyadic shifts and the $T(1)$ theorem. *Monatsh. Math.*, pages 1–41, 2016. 366, 472

Harboure, E., J. L. Torrea, and B. Viviani. Vector-valued extensions of operators related to the Ornstein-Uhlenbeck semigroup. *J. Anal. Math.*, 91: 1–29, 2003. 492

Hardy, G. H. and J. E. Littlewood. A maximal theorem with function-theoretic applications. *Acta Math.*, 54(1):81–116, 1930. 159

Helly, E. Über lineare Functionaloperationen. *Sitzungsber. Kaiserl. (Öster.) Akad. Wissen.*, 121:265–297, 1912. 533

Helly, E. Über Systeme linearer Gleichungen mit unendlich vielen Unbekannten. *Monatsh. Math. Phys.*, 31(1):60–91, 1921. 363

Hernandez, R. Espaces L^p, factorisations et produits tensoriels. In *Seminar on the geometry of Banach spaces, Vol. I, II (Paris, 1983)*, volume 18 of *Publ. Math. Univ. Paris VII*, pages 63–79. Univ. Paris VII, Paris, 1984. 154

Herz, C. The theory of p-spaces with an application to convolution operators. *Trans. Amer. Math. Soc.*, 154:69–82, 1971. 154

Herz, C. Bounded mean oscillation and regulated martingales. *Trans. Amer. Math. Soc.*, 193:199–215, 1974. 259, 260

Hewitt, E. and K. Stromberg. *Real and abstract analysis*. Springer-Verlag, New York, 1975. Third printing, Graduate Texts in Mathematics, No. 25. 61, 515

Hieber, M. and J. Prüss. Functional calculi for linear operators in vector-valued L^p-spaces via the transference principle. *Adv. Differential Equations*, 3(6):847–872, 1998. 475

Hille, E. and R. S. Phillips. *Functional analysis and semi-groups.* American Mathematical Society Colloquium Publications, vol. 31. American Mathematical Society, Providence, R.I., 1957. rev. ed. 61, 64

Hinrichs, A. On the type constants with respect to systems of characters of a compact abelian group. *Studia Math.*, 118(3):231–243, 1996. 161

Hitczenko, P. On tangent sequences of UMD-space valued random vectors. Unpublished manuscript, 1988a. 364, 365

Hitczenko, P. Comparison of moments for tangent sequences of random variables. *Probab. Theory Relat. Fields*, 78(2):223–230, 1988b. 364

Hitczenko, P. Domination inequality for martingale transforms of a Rademacher sequence. *Israel J. Math.*, 84(1-2):161–178, 1993. 499

Hitczenko, P. On a domination of sums of random variables by sums of conditionally independent ones. *Ann. Probab.*, 22(1):453–468, 1994. 358, 366, 499

Hoffmann-Jørgensen, J. and G. Pisier. The law of large numbers and the central limit theorem in Banach spaces. *Ann. Probab.*, 4(4):587–599, 1976. 157

Holmstedt, T. Interpolation of quasi-normed spaces. *Math. Scand.*, 26:177–199, 1970. 563

Hörmander, L. Estimates for translation invariant operators in L^p spaces. *Acta Math.*, 104:93–140, 1960. 480, 486

Hunt, R. A. An extension of the Marcinkiewicz interpolation theorem to Lorentz spaces. *Bull. Amer. Math. Soc.*, 70:803–807, 1964. 159

Hytönen, T. P. Fourier embeddings and Mihlin-type multiplier theorems. *Math. Nachr.*, 274/275:74–103, 2004. 485

Hytönen, T. P. An operator-valued Tb theorem. *J. Funct. Anal.*, 234(2): 420–463, 2006. 472

Hytönen, T. P. Littlewood-Paley-Stein theory for semigroups in UMD spaces. *Rev. Mat. Iberoam.*, 23(3):973–1009, 2007. 475, 483, 495

Hytönen, T. P. On Petermichl's dyadic shift and the Hilbert transform. *C. R. Math. Acad. Sci. Paris*, 346(21-22):1133–1136, 2008. 471

Hytönen, T. P. New thoughts on the vector-valued Mihlin-Hörmander multiplier theorem. *Proc. Amer. Math. Soc.*, 138(7):2553–2560, 2010a. 485

Hytönen, T. P. The real-variable Hardy space and BMO. Lecture notes, University of Helsinki, http://www.helsinki.fi/~tpehyton/Hardy/hardy.pdf, 2010b. 260

Hytönen, T. P. The sharp weighted bound for general Calderón-Zygmund operators. *Ann. of Math. (2)*, 175(3):1473–1506, 2012. 474

Hytönen, T. P. The vector-valued nonhomogeneous Tb theorem. *Int. Math. Res. Not. IMRN*, (2):451–511, 2014. 366, 472

Hytönen, T. P. and A. Kairema. Systems of dyadic cubes in a doubling metric space. *Colloq. Math.*, 126(1):1–33, 2012. 261

Hytönen, T. P. and M. Kemppainen. On the relation of Carleson's embedding and the maximal theorem in the context of Banach space geometry. *Math. Scand.*, 109(2):269–284, 2011. 266

Hytönen, T. P. and M. T. Lacey. Pointwise convergence of Walsh–Fourier series of vector-valued functions. *arXiv:1202.0209*, 2012. 362

Hytönen, T. P. and M. T. Lacey. Pointwise convergence of vector-valued Fourier series. *Math. Ann.*, 357(4):1329–1361, 2013. 476, 495, 496

Hytönen, T. P. and A. Naor. Pisier's inequality revisited. *Studia Math.*, 215 (3):221–235, 2013. 371

Hytönen, T. P. and P. Portal. Vector-valued multiparameter singular integrals and pseudodifferential operators. *Adv. Math.*, 217(2):519–536, 2008. 481, 485

Hytönen, T. P. and D. Potapov. Vector-valued multiplier theorems of Coifman-Rubio de Francia-Semmes type. *Arch. Math. (Basel)*, 87(3):245–254, 2006. 482

Hytönen, T. P. and A. V. Vähäkangas. The local non-homogeneous Tb theorem for vector-valued functions. *Glasg. Math. J.*, 57(1):17–82, 2015. 472

Hytönen, T. P. and L. W. Weis. A $T1$ theorem for integral transformations with operator-valued kernel. *J. Reine Angew. Math.*, 599:155–200, 2006. 471

Hytönen, T. P. and L. W. Weis. Singular convolution integrals with operator-valued kernel. *Math. Z.*, 255(2):393–425, 2007. 481

Hytönen, T. P., A. McIntosh, and P. Portal. Kato's square root problem in Banach spaces. *J. Funct. Anal.*, 254(3):675–726, 2008. 266, 500

Hytönen, T. P., J. L. Torrea, and D. V. Yakubovich. The Littlewood-Paley-Rubio de Francia property of a Banach space for the case of equal intervals. *Proc. Roy. Soc. Edinburgh Sect. A*, 139(4):819–832, 2009. 482

Hytönen, T. P., S. Li, and A. Naor. Quantitative affine approximation for UMD targets. *Discrete Analysis*, 2016. 489

Ionescu Tulcea, A. and C. Ionescu Tulcea. Abstract ergodic theorems. *Trans. Amer. Math. Soc.*, 107:107–124, 1963. 262

Itô, K. and M. Nisio. On the convergence of sums of independent Banach space valued random variables. *Osaka J. Math.*, 5:35–48, 1968. 262

Iwaniec, T. Extremal inequalities in Sobolev spaces and quasiconformal mappings. *Z. Anal. Anwendungen*, 1(6):1–16, 1982. 370, 493, 494

Jacod, J. Une généralisation des semimartingales: les processus admettant un processus à accroissements indépendants tangent. In *Seminar on probability, XVIII*, volume 1059 of *Lecture Notes in Math.*, pages 91–118. Springer, Berlin, 1984. 364

James, R. C. Bases and reflexivity of Banach spaces. *Ann. of Math. (2)*, 52: 518–527, 1950. 535

James, R. C. Characterizations of reflexivity. *Studia Math.*, 23:205–216, 1963/1964. 363

James, R. C. Some self-dual properties of normed linear spaces. In *Symposium on Infinite-Dimensional Topology (Louisiana State Univ., Baton Rouge,* *La., 1967)*, pages 159–175. Ann. of Math. Studies, No. 69. Princeton Univ. Press, Princeton, N.J., 1972a. 264, 363

James, R. C. Super-reflexive Banach spaces. *Canad. J. Math.*, 24:896–904, 1972b. 363

James, R. C. Nonreflexive spaces of type 2. *Israel J. Math.*, 30(1-2):1–13, 1978. 500

Janicka, L. and N. J. Kalton. Vector measures of infinite variation. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 25(3):239–241, 1977. 63

Jensen, J. L. W. V. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Math.*, 30(1):175–193, 1906. 61, 158

Jessen, B. On the approximation of Lebesgue integrals by Riemann sums. *Ann. of Math. (2)*, 35(2):248–251, 1934. 262

Jiao, Y., L. Wu, and M. Popa. Operator-valued martingale transforms in rearrangement invariant spaces and applications. *Sci. China Math.*, 56(4): 831–844, 2013. 360

John, F. and L. Nirenberg. On functions of bounded mean oscillation. *Comm. Pure Appl. Math.*, 14:415–426, 1961. 259

Johnson, G. W. The product of strong operator measurable functions is strong operator measurable. *Proc. Amer. Math. Soc.*, 117(4):1097–1104, 1993. 63

Junge, M. Doob's inequality for non-commutative martingales. *J. Reine Angew. Math.*, 549:149–190, 2002. 259

Kahane, J.-P. Sur les sommes vectorielles $\sum \pm u_n$. *C. R. Acad. Sci. Paris*, 259:2577–2580, 1964. 260

Kahane, J.-P. *Some random series of functions*, volume 5 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 1985. 259, 260

Kallenberg, O. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002. 163, 259, 515

Kalton, N. J. Differentials of complex interpolation processes for Köthe function spaces. *Trans. Amer. Math. Soc.*, 333(2):479–529, 1992. 357

Kalton, N. J. An elementary example of a Banach space not isomorphic to its complex conjugate. *Canad. Math. Bull.*, 38(2):218–222, 1995. 536

Kalton, N. J. and L. W. Weis. The H^∞-calculus and sums of closed operators. *Math. Ann.*, 321(2):319–345, 2001. 473

Kalton, N. J. and L. W. Weis. The H^∞-calculus and square function estimates. In *Nigel J. Kalton Selecta, Vol. 1*, pages 715–764. Springer-Verlag, 2016. 482

Kalton, N. J., P. Kunstmann, and L. W. Weis. Perturbation and interpolation theorems for the H^∞-calculus with applications to differential operators. *Math. Ann.*, 336(4):747–801, 2006. 563

Kato, M. and Y. Takahashi. Type, cotype constants and Clarkson's inequalities for Banach spaces. *Math. Nachr.*, 186:187–196, 1997. 159

Kato, T. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition. 534

Kelley, J. L. and I. Namioka. *Linear topological spaces.* Springer-Verlag, New York-Heidelberg, 1976. Second corrected printing, Graduate Texts in Mathematics, No. 36. 534

Kemppainen, M. On the Rademacher maximal function. *Studia Math.*, 203 (1):1–31, 2011. 266

Kemppainen, M. Some remarks on the dyadic Rademacher maximal function. *Colloq. Math.*, 131(1):113–128, 2013. 266

Khintchine, A. Über dyadische Brüche. *Math. Z.*, 18(1):109–116, 1923. 260

Kolmogorov, A. Sur les fonctions harmoniques conjuguées et les séries de Fourier. *Fund. Math.*, 7:24–29, 1925. 159, 469

König, H. *Eigenvalue distribution of compact operators*, volume 16 of *Operator Theory: Advances and Applications.* Birkhäuser Verlag, Basel, 1986. 159, 562

König, H. On the Fourier-coefficients of vector-valued functions. *Math. Nachr.*, 152:215–227, 1991. 161

König, H. Vector-valued Lagrange interpolation and mean convergence of Hermite series. In *Functional analysis (Essen, 1991)*, volume 150 of *Lecture Notes in Pure and Appl. Math.*, pages 227–247. Dekker, New York, 1994. 477

König, H. and N. J. Nielsen. Vector-valued L_p-convergence of orthogonal series and Lagrange interpolation. *Forum Math.*, 6(2):183–207, 1994. 477

Kreĭn, M. G. Some new studies in the theory of perturbations of self-adjoint operators. In *First Math. Summer School, Part I (Russian)*, pages 103–187. Izdat. “Naukova Dumka”, Kiev, 1964. 484

Kreĭn, M. G. and V. Šmulian. On regulary convex sets in the space conjugate to a Banach space. *Ann. of Math. (2)*, 41:556–583, 1940. 533

Kreĭn, S. G., Yu. Ī. Petunīn, and E. M. Semënov. *Interpolation of linear operators*, volume 54 of *Translations of Mathematical Monographs.* American Mathematical Society, Providence, R.I., 1982. 158, 159, 562

Kunstmann, P. C. and L. W. Weis. Maximal L_p-regularity for parabolic equations, Fourier multiplier theorems and H^∞-functional calculus. In *Functional analytic methods for evolution equations*, volume 1855 of *Lecture Notes in Math.*, pages 65–311. Springer, Berlin, 2004. 485

Kwapień, S. Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients. *Studia Math.*, 44:583–595, 1972a. 160, 265

Kwapień, S. On operators factorizable through L_p space. In *Actes du Colloque d'Analyse Fonctionnelle de Bordeaux (Univ. de Bordeaux, 1971)*, pages 215–225. Bull. Soc. Math. France, Mém. No. 31–32. Soc. Math. France, Paris, 1972b. 154

Kwapień, S. and A. Pełczyński. The main triangle projection in matrix spaces and its applications. *Studia Math.*, 34:43–68, 1970. 354

Kwapień, S. and W. A. Woyczyński. Decoupling of martingale transforms and stochastic integrals for processes with independent increments. In *Probability theory and harmonic analysis (Cleveland, Ohio, 1983)*, volume 98 of

Monogr. Textbooks Pure Appl. Math., pages 139–148. Dekker, New York, 1986. 364

Kwapień, S. and W. A. Woyczyński. Tangent sequences of random variables: basic inequalities and their applications. In *Almost everywhere convergence (Columbus, OH, 1988)*, pages 237–265. Academic Press, Boston, MA, 1989. 364

Kwapień, S. and W. A. Woyczyński. Semimartingale integrals via decoupling inequalities and tangent processes. *Probab. Math. Statist.*, 12(2):165–200 (1992), 1991. 364, 365

Kwapień, S. and W. A. Woyczyński. *Random series and stochastic integrals: single and multiple, 2nd ed.* Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 2002. 364

Lacey, M. and C. Thiele. A proof of boundedness of the Carleson operator. *Math. Res. Lett.*, 7(4):361–370, 2000. 476

Lechner, R. and M. Passenbrunner. Adaptive deterministic dyadic grids on spaces of homogeneous type. *Bull. Pol. Acad. Sci. Math.*, 62(2):139–160, 2014. 472

Ledoux, M. and M. Talagrand. *Probability in Banach spaces: isoperimetry and processes*, volume 23 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3).* Springer-Verlag, Berlin, 1991. 262

Lehto, O. Remarks on the integrability of the derivatives of quasiconformal mappings. *Ann. Acad. Sci. Fenn. Ser. A I No.*, 371:8, 1965. 493

Lerner, A. K. A simple proof of the A_2 conjecture. *Int. Math. Res. Not. IMRN*, (14):3159–3170, 2013. 474

Li, D. and H. Queffélec. *Introduction à l'étude des espaces de Banach*, volume 12 of *Cours Spécialisés.* Société Mathématique de France, Paris, 2004. 533

Lindemulder, N. Difference norms for vector-valued Bessel potential spaces with an application to pointwise multipliers. *arXiv:1505.08039*, 2015. 489

Lindemulder, N. Vector-valued extensions of linear operators on L^∞. In *Ordered Structures and Applications (Positivity VII, Zaanen Centennial Conference).* Birkhäuser Verlag, 2016. 156

Lindenstrauss, J. and L. Tzafriri. *Classical Banach spaces. Vol. I: Sequence spaces*, volume 92 of *Ergebnisse der Mathematik und ihrer Grenzgebiete.* Springer-Verlag, Berlin-New York, 1977. 354

Lions, J.-L. and J. Peetre. Sur une classe d'espaces d'interpolation. *Inst. Hautes Études Sci. Publ. Math.*, (19):5–68, 1964. 61, 563

Liptser, R. Sh. and A. N. Shiryayev. *Theory of martingales*, volume 49 of *Mathematics and its Applications (Soviet Series).* Kluwer Academic Publishers Group, Dordrecht, 1989. 259

Littlewood, J. E. and R. E. A. C. Paley. Theorems on Fourier series and power series. *J. London Math. Soc.*, 6:230–233, 1931. 479

Littlewood, J. E. and R. E. A. C. Paley. Theorems on Fourier series and power series. II. *Proc. London Math. Soc.*, 42(1):52–89, 1936. 480

Liu, P. D. Spaces in which martingale difference sequences are unconditional. *J. Systems Sci. Math. Sci.*, 9(3):251–259, 1989. 356

Lizorkin, P. I. (L_p, L_q)-multipliers of Fourier integrals. *Dokl. Akad. Nauk SSSR*, 152:808–811, 1963. 485, 486

Long, R. L. *Martingale spaces and inequalities.* Peking University Press, Beijing; Friedr. Vieweg & Sohn, Braunschweig, 1993. 260

López-Sánchez, L. D., Martell. J. M., and J. J. Parcet. Dyadic harmonic analysis beyond doubling measures. *Adv. Math.*, 267:44–93, 2014. 262

Lotz, H. P. Über das Spektrum positiver Operatoren. *Math. Z.*, 108:15–32, 1968. 535

Lü, Q. and J. M. A. M. van Neerven. Conditonal expectations in $L^p(\mu; L^q(\nu))$. *arXiv:1606.02780*, 2016. 163, 164

Lü, Q., J. Yong, and X. Zhang. Representation of Itô integrals by Lebesgue/Bochner integrals. *J. Eur. Math. Soc. (JEMS)*, 14(6):1795–1823, 2012. 164

Lunardi, A. *Interpolation theory.* Appunti. Scuola Normale Superiore di Pisa (Nuova Serie). Edizioni della Normale, Pisa, second edition, 2009. 158, 161, 562

Luxemburg, W. A. J. and A. C. Zaanen. The linear modulus of an order bounded linear transformation. II. *Nederl. Akad. Wetensch. Proc. Ser. A*, 74:435–447, 1971. 535

Macaev, V. I. A class of completely continuous operators. *Dokl. Akad. Nauk SSSR*, 139:548–551, 1961. 484

Marcinkiewicz, J. Sur l'interpolation d'opérations. *C. R. Acad. Sci., Paris*, 208:1272–1273, 1939a. 158, 159

Marcinkiewicz, J. Sur les multiplicateurs des séries de Fourier. *Studia Math.*, 8:78–91, 1939b. 480, 485, 489

Marcinkiewicz, J. *Collected papers.* Edited by Antoni Zygmund. Instytut Matematyczny Polskiej Akademii Nauk. Państwowe Wydawnictwo Naukowe, Warsaw, 1964. 159

Marcinkiewicz, J. and A. Zygmund. Quelques inégalités pour les opérations linéaires. *Fund. Math.*, 32(1):115–121, 1939. 156

Martikainen, H. Vector-valued non-homogeneous Tb theorem on metric measure spaces. *Rev. Mat. Iberoam.*, 28(4):961–998, 2012. 472

Martínez, T. and J. L. Torrea. Operator-valued martingale transforms. *Tohoku Math. J. (2)*, 52(3):449–474, 2000. 261, 263

Martínez, T., J. L. Torrea, and Q. Xu. Vector-valued Littlewood-Paley-Stein theory for semigroups. *Adv. Math.*, 203(2):430–475, 2006. 483

Maurey, B. Système de Haar. In *Séminaire Maurey-Schwartz 1974–1975: Espaces L^p, applications radonifiantes et géométrie des espaces de Banach, Exp. Nos. 1 et 2.* Centre Math., École Polytech., Paris, 1975. 266, 354, 355, 363

McConnell, T. R. On Fourier multiplier transformations of Banach-valued functions. *Trans. Amer. Math. Soc.*, 285(2):739–757, 1984. 479, 485, 486, 490

McConnell, T. R. Decoupling and stochastic integration in UMD Banach spaces. *Probab. Math. Statist.*, 10(2):283–295, 1989. 364, 365, 499

Mei, T. BMO is the intersection of two translates of dyadic BMO. *C. R. Math. Acad. Sci. Paris*, 336(12):1003–1006, 2003. 260

Mei, T. Operator valued Hardy spaces. *Mem. Amer. Math. Soc.*, 188(881): vi+64, 2007. 260

Métivier, M. Martingales à valeurs vectorielles. Applications à la dérivation des mesures vectorielles. *Ann. Inst. Fourier (Grenoble)*, 17(fasc. 2):175–208 (1968), 1967. 262

Meyer-Nieberg, P. *Banach lattices.* Universitext. Springer-Verlag, Berlin, 1991. 156

Meyries, M. and M. C. Veraar. Pointwise multiplication on vector-valued function spaces with power weights. *J. Fourier Anal. Appl.*, 21(1):95–136, 2015. 489

Mihlin, S. G. On the multipliers of Fourier integrals. *Dokl. Akad. Nauk SSSR (N.S.)*, 109:701–703, 1956. 480

Mihlin, S. G. Fourier integrals and multiple singular integrals. *Vestn. Leningr. Univ. (Ser. Mat. Mekh. Astron.)*, 12(7):143–155, 1957. 480, 485, 489

Milman, M. Complex interpolation and geometry of Banach spaces. *Ann. Mat. Pura Appl. (4)*, 136:317–328, 1984. 159

Mittelmeyer, G. and M. Wolff. Über den Absolutbetrag auf komplexen Vektorverbänden. *Math. Z.*, 137:87–92, 1974. 535

Monna, A. F. *Functional analysis in historical perspective.* John Wiley & Sons, New York-Toronto, Ont., 1973. 533

Monniaux, S. and J. Prüss. A theorem of the Dore-Venni type for noncommuting operators. *Trans. Amer. Math. Soc.*, 349(12):4787–4814, 1997. 473

Montgomery-Smith, S. J. Concrete representation of martingales. *Electron. J. Probab.*, 3:No. 15, 15 pp. (electronic), 1998. 364

Morrey, Jr., C. B. Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific J. Math.*, 2:25–53, 1952. 371

Morrey, Jr., C. B. *Multiple integrals in the calculus of variations.* Die Grundlehren der mathematischen Wissenschaften, Band 130. Springer-Verlag New York, Inc., New York, 1966. 371

Muckenhoupt, B. Poisson integrals for Hermite and Laguerre expansions. *Trans. Amer. Math. Soc.*, 139:231–242, 1969. 492

Müller, P. F. X. *Isomorphisms between H^1 spaces*, volume 66 of *Mathematics Institute of the Polish Academy of Sciences. Mathematical Monographs (New Series).* Birkhäuser Verlag, Basel, 2005. 260

Müller, P. F. X. and M. Passenbrunner. A decomposition theorem for singular integral operators on spaces of homogeneous type. *J. Funct. Anal.*, 262(4): 1427–1465, 2012. 472

Müller, P. F. X. and G. Schechtman. Several results concerning unconditionality in vector valued L^p and $H^1(\mathscr{F}_n)$ spaces. *Illinois J. Math.*, 35(2): 220–233, 1991. 356

Muñoz, G. A., Y. Sarantopoulos, and A. Tonge. Complexifications of real Banach spaces, polynomials and multilinear maps. *Studia Math.*, 134(1): 1–33, 1999. 534

Murray, F. J. Linear transformations in L_p, $p > 1$. *Trans. Amer. Math. Soc.*, 39(1):83–100, 1936. 533

Musiał, K. Pettis integral. In *Handbook of measure theory, Vol. I, II*, pages 531–586. North-Holland, Amsterdam, 2002. 61

Naor, A. and G. Schechtman. Remarks on non linear type and Pisier's inequality. *J. Reine Angew. Math.*, 552:213–236, 2002. 371

Nazarov, F., S. Treil, and A. Volberg. Accretive system Tb-theorems on nonhomogeneous spaces. *Duke Math. J.*, 113(2):259–312, 2002. 472

Nazarov, F., S. Treil, and A. Volberg. The Tb-theorem on non-homogeneous spaces. *Acta Math.*, 190(2):151–239, 2003. 471, 472

Neerven, J. M. A. M. van. The norm of a complex Banach lattice. *Positivity*, 1(4):381–390, 1997. 535

Neerven, J. M. A. M. van and M. C. Veraar. On the stochastic Fubini theorem in infinite dimensions. In *Stochastic partial differential equations and applications—VII*, volume 245 of *Lect. Notes Pure Appl. Math.*, pages 323–336. Chapman & Hall/CRC, Boca Raton, FL, 2006. 358

Neerven, J. M. A. M. van and J. Zhu. A maximal inequality for stochastic convolutions in 2-smooth Banach spaces. *Electron. Commun. Probab.*, 16: 689–705, 2011. 264

Neerven, J. M. A. M. van, M. C. Veraar, and L. W. Weis. Stochastic integration in UMD Banach spaces. *Annals Probab.*, 35:1438–1478, 2007. 499

Neerven, J. M. A. M. van, M. C. Veraar, and L. W. Weis. Stochastic integration in Banach spaces – a survey. In *Stochastic Analysis: A Series of Lectures*, volume 68 of *Progress in Probability*. Birkhäuser Verlag, 2015. 499

Neidhardt, A. L. *Stochastic integrals in 2-uniformly smooth Banach spaces.* PhD thesis, University of Wisconsin, 1978. 264, 499

Neveu, J. Relations entre la théorie des martingales et la théorie ergodique. *Ann. Inst. Fourier (Grenoble)*, 15(fasc. 1):31–42, 1965. 261

Neveu, J. *Discrete-parameter martingales.* North-Holland Publishing Co., Amsterdam, revised edition, 1975. 162, 259

Okikiolu, K. Characterization of subsets of rectifiable curves in $\mathbf{R}^n$. *J. London Math. Soc. (2)*, 46(2):336–348, 1992. 260

Osękowski, A. Weak norm inequalities for martingales and geometry of Banach spaces. *Statist. Probab. Lett.*, 82(3):411–418, 2012a. 360

Osękowski, A. *Sharp martingale and semimartingale inequalities*, volume 72 of *Instytut Matematyczny Polskiej Akademii Nauk. Monografie Matematyczne (New Series).* Birkhäuser/Springer Basel AG, Basel, 2012b. 367, 369

Osękowski, A. On the UMD constant of the space ℓ_1^N. *Colloq. Math.*, 142(1): 135–147, 2016. 368

Pagter, B. de and H. Witvliet. Unconditional decompositions and UMD-spaces. In *Semi-groupes d'opérateurs et calcul fonctionnel (Besançon,* *1998)*, volume 16 of *Publ. Math. UFR Sci. Tech. Besançon*, pages 79–111. Univ. Franche-Comté, Besançon, 1999. 355

Pagter, B. de, H. Witvliet, and F. A. Sukochev. Double operator integrals. *J. Funct. Anal.*, 192(1):52–111, 2002. 485

Paley, R. E. A. C. A remarkable series of orthogonal functions (I). *Proc. London Math. Soc.*, 2(1):241–264, 1932. 156

Parcet, J., F. Soria, and Q. Xu. On the growth of vector-valued Fourier series. *Indiana Univ. Math. J.*, 62(6):1765–1784, 2013. 476, 496

Parthasarathy, K. R. *Probability measures on metric spaces.* Academic Press, New York-London, 1967. 515

Pedregal, P. and V. Šverák. A note on quasiconvexity and rank-one convexity for 2×2 matrices. *J. Convex Anal.*, 5(1):107–117, 1998. 371

Peetre, J. Sur le nombre de paramétres dans la définition de certains espaces d'interpolation. *Ricerche Mat.*, 12:248–261, 1963. 563

Peetre, J. Sur la transformation de Fourier des fonctions à valeurs vectorielles. *Rend. Sem. Mat. Univ. Padova*, 42:15–26, 1969. 157, 160, 563

Peña, V. H. de la and E. Giné. *Decoupling.* Probability and its Applications (New York). Springer-Verlag, New York, 1999. 364

Petermichl, S. Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol. *C. R. Acad. Sci. Paris Sér. I Math.*, 330(6):455–460, 2000. 470, 471, 473

Petermichl, S. The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic. *Amer. J. Math.*, 129(5):1355–1375, 2007. 473

Petermichl, S. and S. Pott. A version of Burkholder's theorem for operator-weighted spaces. *Proc. Amer. Math. Soc.*, 131(11):3457–3461 (electronic), 2003. 470

Pettis, B. J. On integration in vector spaces. *Trans. Amer. Math. Soc.*, 44 (2):277–304, 1938. 59, 61

Piasecki, M. A geometrical characterization of AUMD Banach spaces via subharmonic functions. *Demonstratio Math.*, 30(3):641–654, 1997a. 372

Piasecki, M. A characterization of complex AUMD Banach spaces via tangent martingales. *Demonstratio Math.*, 30(4):715–728, 1997b. 372

Pichorides, S. K. On the best values of the constants in the theorems of M. Riesz, Zygmund and Kolmogorov. *Studia Math.*, 44:165–179. (errata insert), 1972. Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, II. 484

Pietsch, A. *History of Banach spaces and linear operators.* Birkhäuser Boston Inc., Boston, MA, 2007. 63, 533

Pietsch, A. and J. Wenzel. *Orthonormal systems and Banach space geometry*, volume 70 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1998. 161, 259, 354, 358, 359, 497, 498

Pinelis, I. Optimum bounds for the distributions of martingales in Banach spaces. *Ann. Probab.*, 22(4):1679–1706, 1994. 264

Pisier, G. Martingales with values in uniformly convex spaces. *Israel J. Math.*, 20(3-4):326–350, 1975a. 263, 363

Pisier, G. Un exemple concernant la super-réflexivité. In *Séminaire Maurey-Schwartz 1974–1975: Espaces L^p, applications radonifiantes et géométrie des espaces de Banach, Annexe No. 2*, page 12. Centre Math. École Polytech., Paris, 1975b. 354

Pisier, G. Some results on Banach spaces without local unconditional structure. *Compositio Math.*, 37(1):3–19, 1978. 355

Pisier, G. *Factorization of linear operators and geometry of Banach spaces*, volume 60 of *CBMS Regional Conference Series in Mathematics*. Washington, DC, 1986a. 154

Pisier, G. Probabilistic methods in the geometry of Banach spaces. In *Probability and analysis (Varenna, 1985)*, volume 1206 of *Lecture Notes in Math.*, pages 167–241. Springer, Berlin, 1986b. 371

Pisier, G. Complex interpolation between Hilbert, Banach and operator spaces. *Mem. Amer. Math. Soc.*, 208(978):vi+78, 2010. 154

Pisier, G. *Martingales in Banach spaces*, volume 155 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2016. 66, 259, 264, 354, 478, 562

Pisier, G. and Q. Xu. Non-commutative martingale inequalities. *Comm. Math. Phys.*, 189(3):667–698, 1997. 356, 358, 484

Pisier, G. and Q. Xu. Non-commutative L^p-spaces. In *Handbook of the geometry of Banach spaces, Vol. 2*, pages 1459–1517. North-Holland, Amsterdam, 2003. 356

Portal, P. and Ž. Štrkalj. Pseudodifferential operators on Bochner spaces and an application. *Math. Z.*, 253(4):805–819, 2006. 485

Potapov, D. and F. Sukochev. A vector-valued Littlewood-Paley theorem for an arbitrary set of intervals and its applications. *Dokl. Akad. Nauk*, 404(5): 598–600, 2005. 482

Potapov, D. and F. Sukochev. Operator-Lipschitz functions in Schatten-von Neumann classes. *Acta Math.*, 207(2):375–389, 2011. 484, 485

Potapov, D., F. Sukochev, and Q. Xu. On the vector-valued Littlewood-Paley-Rubio de Francia inequality. *Rev. Mat. Iberoam.*, 28(3):839–856, 2012. 482

Pott, S. and A. Stoica. Linear bounds for Calderón-Zygmund operators with even kernel on UMD spaces. *J. Funct. Anal.*, 266(5):3303–3319, 2014. 472

Prüss, J. and G. Simonett. *Moving interfaces and quasilinear parabolic evolution equations*, volume 105 of *Monographs in Mathematics*. Birkhäuser, Basel, 2016. 161, 473, 485

Qiu, Y. On the UMD constants for a class of iterated $L_p(L_q)$ spaces. *J. Funct. Anal.*, 263(8):2409–2429, 2012. 164, 363

Randrianantoanina, N. Non-commutative martingale transforms. *J. Funct. Anal.*, 194(1):181–212, 2002. 484

Revuz, D. and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, third edition, 1999. 259

Riesz, M. Les fonctions conjuguées et les series de Fourier. *C. R. Acad. Sci., Paris*, 178:1464–1467, 1924. 469, 476

Riesz, M. Sur les maxima des formes bilinéaires et sur les fonctionnelles linéaires. *Acta Math.*, 49(3-4):465–497, 1927. 158

Riesz, M. Sur les fonctions conjuguées. *Math. Z.*, 27(1):218–244, 1928. 469, 480

Rogers, L. C. G. and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 1: Foundations.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Ltd., Chichester, second edition, 1994. 259

Rosenthal, H. P. On the subspaces of L^p ($p > 2$) spanned by sequences of independent random variables. *Israel J. Math.*, 8:273–303, 1970. 154

Rosenthal, H. P. On subspaces of L^p. *Ann. of Math. (2)*, 97:344–373, 1973. 154

Rosiński, J. and Z. Suchanecki. On the space of vector-valued functions integrable with respect to the white noise. *Colloq. Math.*, 43(1):183–201 (1981), 1980. 157

Ross, K. A. and K. Stromberg. Jessen's theorem on Riemann sums for locally compact groups. *Pacific J. Math.*, 20:135–147, 1967. 262

Rozendaal, J. and M. C. Veraar. Fourier multiplier theorems involving type and cotype. *Banach Journal of Mathematical Analysis, arXiv:1605.09340*, 2016a. 486

Rozendaal, J. and M. C. Veraar. Fourier multiplier theorems on Besov spaces under type and cotype conditions. *arXiv:1605.09340*, 2016b. 486

Rubio de Francia, J. L. Estimates for some square functions of Littlewood-Paley type. *Publ. Sec. Mat. Univ. Autònoma Barcelona*, 27(2):81–108, 1983. 481

Rubio de Francia, J. L. Factorization theory and A_p weights. *Amer. J. Math.*, 106(3):533–547, 1984. 157, 261

Rubio de Francia, J. L. A Littlewood-Paley inequality for arbitrary intervals. *Rev. Mat. Iberoam.*, 1(2):1–14, 1985. 481, 482, 496

Rubio de Francia, J. L. Martingale and integral transforms of Banach space valued functions. In *Probability and Banach spaces (Zaragoza, 1985)*, volume 1221 of *Lecture Notes in Math.*, pages 195–222. Springer, Berlin, 1986. 158, 261, 354, 356, 357, 362, 476, 494, 495, 496

Rudin, W. *Functional analysis.* International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991. 533

Runst, Th. and W. Sickel. *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, volume 3 of *de Gruyter Series in Nonlinear Analysis and Applications.* Walter de Gruyter & Co., Berlin, 1996. 161

Ruston, A. F. *Fredholm theory in Banach spaces*, volume 86 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 1986. 534

Scalora, F. S. Abstract martingale convergence theorems. *Pacific J. Math.*, 11:347–374, 1961. 262

Schaefer, H. H. *Banach lattices and positive operators.* Springer-Verlag, New York, 1974. Die Grundlehren der mathematischen Wissenschaften, Band 215. 535

Schipper, W. J. de. A note on the modulus of an order bounded linear operator between complex vector lattices. *Nederl. Akad. Wetensch. Proc. Ser. A*, 76: 355–367, 1973. 535

Schlüchtermann, G. A note on multiplication of strong operator measurable functions. *Proc. Amer. Math. Soc.*, 123(9):2815–2816, 1995. 63

Schmeisser, H.-J. Vector-valued Sobolev and Besov spaces. In *Seminar analysis of the Karl-Weierstrass-Institute of Mathematics 1985/86 (Berlin, 1985/86)*, volume 96 of *Teubner-Texte Math.*, pages 4–44. Teubner, Leipzig, 1987. 159, 161

Schwartz, L. *Radon measures on arbitrary topological spaces and cylindrical measures.* Oxford University Press, London, 1973. Tata Institute of Fundamental Research Studies in Mathematics, No. 6. 61

Seeley, R. Norms and domains of the complex powers A_B^z. *Amer. J. Math.*, 93:299–309, 1971. 489

Seidler, J. Exponential estimates for stochastic convolutions in 2-smooth Banach spaces. *Electron. J. Probab.*, 15:no. 50, 1556–1573, 2010. 264

Sierpinski, W. Sur les fonctions d'ensemble additives et continues. *Fund. Math.*, 3:240–246, 1922. 515

Soukhomlinoff, G. Über Fortsetzung von linearen Funktionalen in linearen komplexen Räumen und linearen Quaternionräumen. *Mat. Sbornik*, 45(2): 353–358, 1938. 533

Stein, E. M. *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N. J., 1970a. 471, 478

Stein, E. M. *Topics in harmonic analysis related to the Littlewood-Paley theory.* Annals of Mathematics Studies, No. 63. Princeton University Press, Princeton, N. J., 1970b. 262, 359, 483, 495

Stein, E. M. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series.* Princeton University Press, Princeton, NJ, 1993. 65, 478

Stein, E. M. and R. Shakarchi. *Real analysis: measure theory, integration, and Hilbert spaces.* Princeton Lectures in Analysis, III. Princeton University Press, Princeton, NJ, 2005. 161

Stein, E. M. and G. Weiss. *Introduction to Fourier analysis on Euclidean spaces.* Princeton University Press, Princeton, N. J., 1971. Princeton Mathematical Series, No. 32. 160, 479

Štrkalj, Ž. and L. W. Weis. On operator-valued Fourier multiplier theorems. *Trans. Amer. Math. Soc.*, 359(8):3529–3547 (electronic), 2007. 480, 485, 490

Stromberg, K. R. *Probability for analysts.* Chapman & Hall Probability Series. Chapman & Hall, New York, 1994. 162

Suárez, J. and L. W. Weis. Interpolation of Banach spaces by the γ-method. In *Methods in Banach space theory*, volume 337 of *London Math. Soc. Lecture Note Ser.*, pages 293–306. Cambridge Univ. Press, Cambridge, 2006. 563

Sukochev, F. A. and S. V. Ferleger. Harmonic analysis in UMD-spaces: applications to basis theory. *Mat. Zametki*, 58(6):890–905, 960, 1995. 356

Šverák, V. Rank-one convexity does not imply quasiconvexity. *Proc. Roy. Soc. Edinburgh Sect. A*, 120(1-2):185–189, 1992. 371

Taggart, R. J. Pointwise convergence for semigroups in vector-valued L^p spaces. *Math. Z.*, 261(4):933–949, 2009. 495

Talagrand, M. Comparaison des boreliens d'un espace de Banach pour les topologies fortes et faibles. *Indiana Univ. Math. J.*, 27(6):1001–1004, 1978. 60

Talagrand, M. Pettis integral and measure theory. *Mem. Amer. Math. Soc.*, 51(307):ix+224, 1984. 60, 61

Talagrand, M. Isoperimetry, logarithmic Sobolev inequalities on the discrete cube, and Margulis' graph connectivity theorem. *Geom. Funct. Anal.*, 3(3): 295–314, 1993. 371

Tanaka, H. and Y. Terasawa. Positive operators and maximal operators in a filtered measure space. *J. Funct. Anal.*, 264(4):920–946, 2013. 162

Taylor, A. E. Additions to the theory of polynomials in normed linear spaces. *Tôhoku Math. J*, 44:302–318, 1938. 534

Thomas, G. E. F. Totally summable functions with values in locally convex spaces. In *Measure theory (Proc. Conf., Oberwolfach, 1975)*, pages 117–131. Lecture Notes in Math., Vol. 541. Springer, Berlin, 1976. 63

Thorin, G. O. An extension of a convexity theorem due to M. Riesz. Fysiograf. Sällsk. Lund Förhdl. 8, 166-170. Meddel. Lunds. Univ. Mat. Sem. 4 (1939), 5 p (1938, 1939), 1938. 158

Tolsa, X. BMO, H^1, and Calderón-Zygmund operators for non doubling measures. *Math. Ann.*, 319(1):89–149, 2001. 262

Tozoni, S. A. Weighted inequalities for vector operators on martingales. *J. Math. Anal. Appl.*, 191(2):229–249, 1995. 263, 360, 362

Tozoni, S. A. Vector-valued extensions of operators on martingales. *J. Math. Anal. Appl.*, 201(1):128–151, 1996. 261, 360

Trautman, D. A. A note on MT operators. *Proc. Amer. Math. Soc.*, 97(3): 445–448, 1986. 359

Triebel, H. *Interpolation theory, function spaces, differential operators*, volume 18 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1978. 158, 159, 161, 562, 563

Triebel, H. *Fractals and spectra*, volume 91 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1997. 161

Uhl, Jr., J. J. Applications of Radon-Nikodým theorems to martingales of vector valued functions. *Bull. Amer. Math. Soc.*, 75:840–842, 1969a. 262

Uhl, Jr., J. J. The Radon-Nikodym theorem and the mean convergence of Banach space valued martingales. *Proc. Amer. Math. Soc.*, 21:139–144, 1969b. 262

Vakhania, N. N., V. I. Tarieladze, and S. A. Chobanyan. *Probability distributions on Banach spaces*, volume 14 of *Mathematics and its Applications (Soviet Series).* D. Reidel Publishing Co., Dordrecht, 1987. 59, 60

Veraar, M. C. Randomized UMD Banach spaces and decoupling inequalities for stochastic integrals. *Proc. Amer. Math. Soc.*, 135:1477–1486, 2007. 358

Veraar, M. C. and L. W. Weis. A note on maximal estimates for stochastic convolutions. *Czechoslovak Math. J.*, 61(136)(3):743–758, 2011. 264

Virot, B. Extensions vectorielles d'opérateurs linéaires bornés sur L^p. In *Harmonic analysis*, volume 8 of *Publ. Math. Orsay 81*, pages Exp. No. 7, 31. Univ. Paris XI, Orsay, 1981. 154

Watson, D. K. Weighted estimates for singular integrals via Fourier transform estimates. *Duke Math. J.*, 60(2):389–399, 1990. 495

Weis, L. W. Operator-valued Fourier multiplier theorems and maximal L_p-regularity. *Math. Ann.*, 319(4):735–758, 2001. 479

Weisz, F. Almost everywhere convergence of Banach space-valued Vilenkin-Fourier series. *Acta Math. Hungar.*, 116(1-2):47–59, 2007a. 362

Weisz, F. Hardy spaces and convergence of vector-valued Vilenkin-Fourier series. *Publ. Math. Debrecen*, 71(3-4):413–424, 2007b. 362

Wenzel, J. The UMD constants of the summation operators. *Quaest. Math.*, 27(2):111–136, 2004. 359, 497, 498

Wenzel, J. Two examples concerning martingales in Banach spaces. *J. London Math. Soc. (2)*, 72(2):442–456, 2005a. 358

Wenzel, J. Strong martingale type and uniform smoothness. *J. Convex Anal.*, 12(1):159–171, 2005b. 264

Werner, D. *Funktionalanalysis.* Springer-Verlag, Berlin, extended edition, 2000. 533

Wiener, N. The ergodic theorem. *Duke Math. J.*, 5(1):1–18, 1939. 160

Williams, D. *Probability with martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991. 259

Witvliet, H. *Unconditional Schauder decompositions and multiplier theorems.* PhD thesis, Delft University of Technology, 2000. 355

Wojtaszczyk, P. *Banach spaces for analysts*, volume 25 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1991. 533

Xu, Q. Littlewood-Paley theory for functions with values in uniformly convex spaces. *J. Reine Angew. Math.*, 504:195–226, 1998. 483

Yor, M. Sur les intégrales stochastiques à valeurs dans un espace de Banach. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 10:31–36, 1974. 157

Yosida, K. *Functional analysis*, volume 123 of *Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, Berlin, sixth edition, 1980. 533

Zhang, X. and C. Zhang. Atomic decompositions of Banach lattice-valued martingales. *Statist. Probab. Lett.*, 82(3):664–671, 2012. 358

Zimmermann, F. On vector-valued Fourier multiplier theorems. *Studia Math.*, 93(3):201–222, 1989. 485, 486

Zinn, J. Comparison of martingale difference sequences. In *Probability in Banach spaces, V (Medford, Mass., 1984)*, volume 1153 of *Lecture Notes in Math.*, pages 453–457. Springer, Berlin, 1985. 364

Zygmund, A. On a theorem of Marcinkiewicz concerning interpolation of operations. *J. Math. Pures Appl. (9)*, 35:223–248, 1956. 158, 159

Zyl, G. van. Complexification of the projective and injective tensor products. *Studia Math.*, 189(2):105–112, 2008. 535

Index

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 63, DOI 10.1007/978-3-319-48520-1