Martingales in Banach Spaces
(in connection with Type and Cotype)

Course IHP

Feb 2-8, 2011

Gilles Pisier

February 9, 2011

Contents

Introduction

Martingales (with discrete time) lie at the centre of these notes (which might become a book). They are known to have major applications to virtually every corner of Probability Theory. Our central theme is their applications to the Geometry of Banach spaces.

We should emphasize that we do *not* assume any knowledge about scalar valued martingales. Actually, the beginning gives a self-contained introduction to the basic martingale convergence theorems for which the use of the norm of a vector valued random variable instead of the modulus of a scalar one makes little difference. Only when we consider the "boundedness implies convergence" phenomenon does it start to matter. Indeed, this requires the Banach space B to have the Radon-Nikodym property (RNP in short).

While the RNP is infinite dimensional and we will concentrate on finite dimensional (also called "local") properties, it is a convenient way to introduce the stronger properties of uniform convexity and smoothness and superreflexivity. Indeed, the martingale inequalities satisfied by super-reflexive spaces can be interpreted as "quantitative versions" of the RNP: roughly RNP means martingales converge and superreflexivity produces a uniform speed for their convergence.

Our main theme in the first part is super-reflexivity and its connections with uniform convexity and smoothness. Roughly we relate the geometric properties of a Banach space B with the study of the p-variation

$$S_p(f) = \left(\sum\nolimits_1^\infty \|f_n - f_{n-1}\|_B^p\right)^{1/p}$$

of B-valued martingales (f_n). Depending whether $S_p(f) \in L_p$ is necessary or sufficient for the convergence of (f_n) in $L_p(B)$, we can find an equivalent norm on B with modulus of uniform convexity (resp. smoothness) "at least as good as" the function $t \to t^p$.

We also consider the strong p-variation

$$V_p(f) = \sup_{0=n(0)<n(1)<n(2)<\cdots} \left(\sum\nolimits_1^\infty \|f_{n(k)} - f_{n(k-1)}\|_B^p\right)^{1/p}$$

of a martingale. For that topic (exceptionally) we devote an entire chapter only to the scalar case. Our crucial tool here is the "real interpolation method".

The first part of the notes with the first 7 chapters are all related to super-reflexivity, or more precisely, to the martingale versions of type and cotype. We will see by an example (see chapter 7) that the latter are strictly stronger than type and cotype.

However, if martingale difference sequences are unconditional, then the martingale versions of type and cotype reduce to the usual ones. This could be one way to motivate the introduction of the UMD property in these notes, but UMD is important in its own right: it is the key to harmonic analysis for Banach space valued functions.

The chapter 8 is devoted to UMD Banach spaces and forms a second part of the notes.

A major feature of the UMD property is its equivalence to the boundedness of the Hilbert transform (HT in short) but we keave this for the final version of these notes.

We also describe in chapter 9 some exciting recent work on non-linear properties of metric spaces analogous to uniform convexity/smoothness and type for metric spaces.

We will now review the contents of these notes chapter by chapter.

Chapter 1 begins with preliminary background: We introduce Banach space valued L_p-spaces, conditional expectations and the central notion in this book, namely Banach space valued martingales associated to a filtration $(\mathcal{A}_n)_{n\geq 0}$ on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We describe the classical examples of filtrations (the dyadic one and the Haar one) in §1.3. If B is an arbitrary Banach space and the martingale (f_n) is associated to some f in $L_p(B)$ by $f_n = \mathbb{E}^{\mathcal{A}_n}(f)$ $(1 \leq p < \infty)$ then, assuming $\mathcal{A} = \mathcal{A}_\infty$ for simplicity, the fundamental convergence theorems say that

$$f_n \to f$$

both in $L_p(B)$ and almost surely (a.s. in short).

The convergence in $L_p(B)$ is Theorem 1.5, while the a.s. convergence is Theorem 1.14. The latter is based on Doob's classical *maximal inequalities* (Theorem 1.9) that are proved using the crucial notion of *stopping time.* We also describe the dual form of Doob's inequality due to Burkholder–Davis–Gundy (see Theorem 1.10). Doob's maximal inequality shows that the convergence of f_n to f in $L_p(B)$ "automatically" implies a.s. convergence. This of course is special to martingales but in general it requires $p \geq 1$. However, for martingales that are sums of independent, symmetric random variables (Y_n) (i.e. we have $f_n = \sum_1^n Y_k$), this result holds for $0 < p < 1$ (see Theorem 1.22). It also holds, roughly, for $p = 0$.

In §1.5, we prove the strong law of large numbers using the a.s. convergence of reverse B-valued martingales.

To get to a.s. convergence, all the preceding results need to assume in the first place some form of convergence, e.g. in $L_p(B)$. In classical (i.e. real valued) martingale theory, it suffices to assume *boundedness* of the martingale $\{f_n\}$ in L_p $(p \geq 1)$ to obtain its a.s. convergence (as well as norm convergence if $1 < p < \infty$). However, this "boundedness $\Rightarrow$ convergence" phenomenon no longer holds in the B-valued case unless B has a specific property called the Radon–Nikodym property (RNP in short) that we introduce and study in Chapter 2. The RNP of a Banach space B expresses the validity of a certain form of the Radon–Nikodym theorem for B-valued measures, but it turns out to be equivalent to the assertion that all martingales bounded in $L_p(B)$ converge a.s. (and in $L_p(B)$ if $p > 1$) for some (or equivalently all) $1 \leq p < \infty$. Moreover, the RNP is equivalent to a certain "geometric" property called "dentability". All this is included in Theorem 2.5. The basic examples of Banach spaces with the RNP are the reflexive ones and separable duals (see Corollary 2.11).

Moreover, a dual space B^* has the RNP iff the classical duality $L_p(B)^* = L_{p'}(B^*)$ is valid for some (or all) $1 < p < \infty$ with $\frac{1}{p} + \frac{1}{p'} = 1$, see Theorem 2.16. Actually, for a general B one can also describe $L_p(B)^*$ as a space of martingales bounded in $L_{p'}(B^*)$, but in general the latter is larger than the (Bochner sense) space $L_{p'}(B^*)$ itself, see Proposition 2.14.

In §2.2, we discuss the Krein–Milman property (KMP): this says that any bounded closed convex set $C \subset B$ is the closed convex hull of its extreme points. This is closely related to dentability, but although it is known that RNP $\Rightarrow$ KMP (see Theorem 2.21) the converse implication is still open.

Chapter 3 is devoted to *super-reflexivity.* A Banach space B is super-reflexive if every space that is *finitely representable* in B is reflexive. In §3.1 we introduce finite representability and general super-properties in connection with ultraproducts. We include some background about the latter in an appendix to Chapter 3.

In §3, we concentrate on super-P when P is either "reflexivity" or the RNP. We prove that super-reflexivity is equivalent to the super-RNP (see Theorem 3.11). We give (see Theorem 3.10) a fundamental characterization of reflexivity, from which one can also derive easily (see Theorem 3.22) one of super-reflexivity.

As in the preceding chapter, we replace B by $L_2(B)$ and view martingale difference sequences as monotone basic sequences in $L_2(B)$. Then we deduce the martingale inequalities from those satisfied by general basic sequences in super-reflexive spaces.

In §3.3, we show that uniformly non-square Banach spaces are reflexive, and hence automatically super-reflexive (see Theorem 3.24 and Corollary 3.26). More generally, we go on to prove that B is super-reflexive iff it is J-convex, or equivalently iff it is J-(n, ε) convex for some $n \geq 2$ and some $\varepsilon > 0$. We say that B is J-(n, ε) convex if for any n-tuple $(x_1, \ldots, x_n)$ in the unit ball of B there is an integer $j = 1, \ldots, n$ such that

$$\left\| \sum_{i<j} x_i - \sum_{i \geq j} x_i \right\| \leq n(1 - \varepsilon).$$

When $n = 2$, we recover the notion of "uniformly non-square". The implication super-reflexive $\Rightarrow$ J-convex is rather easy to derive (as we do in Corollary 3.34) from the fundamental reflexivity criterion stated as Theorem 3.10. The converse implication (due to James) is much more delicate. We prove it following essentially the Brunel–Sucheston approach ([77]), that in our opinion is much easier to grasp. This construction shows that a non-super-reflexive (or merely non-reflexive) space B contains very *extreme finite dimensional* structures that constitute obstructions to either reflexivity or the RNP. For instance any such B admits a space $\widetilde{B}$ finitely representable in B for which there is a dyadic martingale (f_n) with values in the unit ball of $\widetilde{B}$ such that

$$\forall n \geq 1 \qquad \|f_n - f_{n-1}\|_B \equiv 1.$$

Thus the unit ball of $\widetilde{B}$ contains an extremely sparsely separated infinite dyadic tree. (See Remark 1.25 for concrete examples of such trees.)

In §3.4, we finally connect super-reflexivity and uniform convexity. We prove that B is super-reflexive iff it is isomorphic to either a uniformly convex space, or a uniformly smooth one, or a uniformly non-square one. By the preceding Chapter 4, we already know that the renormings can be achieved with moduli of convexity and smoothness of "power type". Using interpolation (see Proposition 3.42) we can even obtain a renorming that is both p-uniformly smooth and q-uniformly convex for some $1 < p, q < \infty$, but it is still open whether this holds with the optimal choice of $p > 1$ and $q < \infty$. To end Chapter 3, we give a characterization of super-reflexivity by the validity of a version of the strong law of large numbers for B-valued martingales.

In Chapter 4, we turn to uniform convexity and uniform smoothness of Banach spaces. We show that certain martingale inequalities characterize Banach spaces B that admit an equivalent norm for which there is a constant C and $2 \leq q < \infty$ (resp. $1 < p \leq 2$) such that for any x, y in B

$$\|x\|^q + C\|y\|^q \leq \frac{\|x+y\|^q + \|x-y\|^q}{2} \tag{1}$$

$\Big($resp.

$$\frac{\|x+y\|^p + \|x-y\|^p}{2} \leq \|x\|^p + C\|y\|^p\Big). \tag{2}$$

This is the content of Corollary 4.7 (resp. Corollary 4.22). We use this in Theorem 4.1 (resp. Th. 4.24) to show that actually *any* uniformly convex (resp. smooth) Banach space admits for some $2 \leq q < \infty$ (resp. $1 < p \leq 2$) such an equivalent renorming. The inequality (1) (resp. (2)) holds iff the modulus of uniform convexity (resp. smoothness) $\delta(\varepsilon)$ (resp. $\rho(t)$) satisfies $\inf_{\varepsilon>0} \delta(\varepsilon)\varepsilon^{-q} > 0$ (resp. $\sup_{t>0} \rho(t)t^{-p} < \infty$). In that case we say that the space is q-uniformly convex (resp. p-uniformly smooth). The proof also uses inequalities going back to Gurarii, James and Lindenstrauss on monotone basic sequences. We apply the latter to martingale difference sequences viewed as monotone basic sequences in $L_p(B)$. Our treatment of uniform smoothness in §4.2 runs parallel to that of uniform convexity in §4.1.

In §4.3, we estimate the moduli of uniform convexity and smoothness of L_p for $1 < p < \infty$. In particular, L_p is p-uniformly convex if $2 \leq p < \infty$ and p-uniformly smooth if $1 < p \leq 2$.

In §4.5, we prove analogues of Burkholder's inequalities but with the square function now replaced by

$$S_p(f) = \left(\|f_0\|_B^p + \sum\nolimits_1^\infty \|f_n - f_{n-1}\|_B^p\right)^{1/p}.$$

Unfortunately the results are now only one-sided: if B satisfies (1) (resp. (2)) then $\|S_q(f)\|_r$ is dominated by (resp. $\|S_p(f)\|_r$ dominates) $\|f\|_{L_r(B)}$ for all $1 < r < \infty$, but here $p \le 2 \le q$ and the case $p = q$ is reduced to the Hilbert space case.

In §4.6, we return to the strong p-variation and prove analogous results to the preceding ones but this time with $W_q(f)$ and $W_p(f)$ in place of $S_q(f)$ and $S_p(f)$ and $1 < p < 2 < q < \infty$. The technique here is similar to that used for the scalar case in Chapter 6.

In Chapter 5, although we mention the complex method, we concentrate on the real method of interpolation for pairs of Banach spaces (B_0, B_1) assumed compatible for interpolation purposes. The complex interpolation space is denoted by $(B_0, B_1)_\theta$. It depends on the single parameter $0 < \theta < 1$, and requires B_0, B_1 to be both *complex* Banach spaces. Complex interpolation is a sort of "abstract" generalization of the classical Riesz–Thorin theorem, asserting that if an operator T has norm 1 simultaneously on both spaces $B_0 = L_{p_0}$ and $B_1 = L_{p_1}$, with $1 \le p_0 < p_1 \le \infty$, then it also has norm 1 on the space L_p for any p such that $p_0 < p < p_1$.

The real interpolation space is denoted by $(B_0, B_1)_{\theta,q}$. It depends on two parameters $0 < \theta < 1$, $1 \le q \le \infty$, and now (B_0, B_1) can be a pair of *real* Banach spaces. Real interpolation is a sort of abstract generalization of the Marcinkiewicz classical theorem already proved in an appendix to Chapter 8. The real interpolation space is introduced using the "K-functional" defined, for any $B_0 + B_1$, by

$$\forall t > 0 \quad K_t(x) = \inf\{\|x_0\|_{B_0} + t\|x_1\|_{B_1} \mid x_0 \in B_0, x_1 \in B_1, x = x_0 + x_1\}.$$

When $B_0 = L_1(\Omega, \mu),\ B_1 = L_\infty(\Omega, \mu)$ we find

$$K_t(x) = \int_0^t x^*(s)ds$$

where x^* is the non-increasing rearrangement of $|x|$ and (Ω, μ) is an arbitrary measure space. We prove this in Theorem 5.3 together with the identification of $(L_1, L_\infty)_{\theta,q}$ with the Lorentz space $L_{p,q}$ for $p = (1-\theta)^{-1}$.

Real interpolation will be crucially used in the later Chapters 6 and 7 in connection with our study of the "strong p-variation" of martingales. The two interpolation methods satisfy distinct properties but are somewhat parallel to each other. For instance, duality, reiteration and interpolation between vector valued L_p-spaces are given parallel treatments in Chapter 5. The classical reference on interpolation is [5] (see also [35]).

In Chapter 6 we study the strong p-variation $W_p(f)$ of a scalar martingale (f_n). This is defined as the supremum of

$$\Big(|f_{n(0)}|^p + \sum\nolimits_{k=1}^{\infty} |f_{n(k)} - f_{n(k-1)}|^p\Big)^{1/p}$$

over all possible increasing sequences

$$0 = n(0) < n(1) < n(2) < \cdots .$$

The main results are Theorem 6.2 and Proposition 6.6. Roughly this says that, if $1 \le p < 2$, $W_p(f)$ is essentially "controlled" by $(\sum |f_n - f_{n-1}|^p)^{1/p}$, i.e. by the finest partition corresponding to consecutive $n(k)$'s; while, in sharp contrast, if $2 < p < \infty$, it is "controlled" by $|f_\infty| = \lim |f_n|$, or equivalently by the coarsest partition corresponding to the choice $n(0) = 0$, $n(1) = \infty$.

The proofs combine a simple stopping time argument with the reiteration theorem of the real interpolation method.

In Chapter 7, we study the real interpolation spaces $(v_1, \ell_\infty)_{\theta,q}$. As usual, ℓ_∞ (resp. v_1) is the space of scalar sequences (x_n) such that $\sup |x_n| < \infty$ (resp. $\sum_1^\infty |x_n - x_{n-1}| < \infty$) equipped with its natural norm. The inclusion $v_1 \to \ell_\infty$ plays a major part (perhaps behind the scene) in our treatment of (super) reflexivity in Chapter 3. Indeed, by the fundamental Theorem 3.10, B is non-reflexive iff the inclusion $\mathcal{J}\colon\ v_1 \to \ell_\infty$ factors through B, i.e. it admits a factorization

$$v_1 \xrightarrow{a} B \xrightarrow{b} \ell_\infty,$$

with bounded linear maps a, b such that $\mathcal{J} = ba$.

The work of James on J-convexity (described in Chapter 3) left open an important point: whether any Banach space B such that ℓ_1^n is not finitely representable in B (i.e. is not almost isometrically embeddable in B) must be reflexive. James proved that the answer is yes if $n = 2$, but for $n > 2$ this remained open until James himself settled it in [166] by a counterexample for $n = 3$ (see also [168] for simplifications). In the theory of type (and cotype), it is the same to say that, for some $n \ge 2, B$ does not contain ℓ_1^n almost isometrically or to say that B has type p for some $p > 1$ (see the survey [206]). Moreover, type p can be equivalently defined by an inequality analogous to that of p-uniformly smoothness but only for martingales with independent increments. Thus it is natural to wonder whether the strongest notion of "type p", namely type 2, implies reflexivity. In another tour de force, James [167] proved that *it is not so.* His example is rather complicated. However, it turns out that the real interpolation spaces $\mathcal{W}_{p,q} = (v_1, \ell_\infty)_{\theta,q}$ $(1 < p, q < \infty, 1 - \theta = 1/p)$ provide very nice examples of the same kind. Thus, following [235] we prove in Corollary 7.19, that $\mathcal{W}_{p,q}$ has exactly the same type and cotype exponents as the Lorentz space $\ell_{p,q} = (\ell_1, \ell_\infty)_{\theta,q}$ as long as $p \ne 2$, although as already explained $\mathcal{W}_{p,q}$ is *not reflexive* since it lies between v_1 and ℓ_∞. The singularity at $p = 2$ is necessary since (unlike $\ell_2 = \ell_{2,2}$) the space $\mathcal{W}_{2,2}$, being non-reflexive, cannot have both type 2 and cotype 2 since that would force it to be isomorphic to Hilbert space.

In Chapter 7, we include a discussion of the classical James space (usually denoted by J) that we denote by v_2^0. The spaces $\mathcal{W}_{p,q}$ are in many ways similar to the James space, in particular if $1 < p, q < \infty$ they are of codimension 1 in their bidual (see Remark 7.8). We can derive the type and cotype of $\mathcal{W}_{p,q}$ in two ways. The first one proves that the vector valued spaces $\mathcal{W}_{p,q}(L_r)$ satisfy the same kind of "Hölder–Minkowski" inequality than the Lorentz spaces $\ell_{p,q}$ with the only exception of $p = r$. This is the substance of Corollary 7.18. Another way (see Remark 7.25) goes through estimates of the K-functional for the pairs (v_1, ℓ_∞) and also (v_r, ℓ_∞) for $1 < r < \infty$, see Lemma 7.22. Indeed, by the reiteration theorem, we may identify $(v_1, \ell_\infty)_{\theta,q}$ and $(v_r, \ell_\infty)_{\theta,q}$ if $\theta > \theta(r) = 1 - \frac{1}{r}$, and similarly in the vector valued case, see Theorem 7.23. We also use reiteration in Theorem 7.14 to describe the space $(v_r, \ell_\infty)_{\theta,q}$ for $0 < r < 1$. In the final Theorem 7.26, we give an alternate description of $\mathcal{W}_p = \mathcal{W}_{p,p}$ that should convince the reader that it is a very natural space (this is closely connected to "splines" in approximation theory). The description is as follows: a sequence $x = (x_n)_n$ belongs to $\mathcal{W}_p$ iff $\sum_N S_N(x)^p < \infty$ where $S_N(x)$ is the distance in ℓ_∞ of x from the subspace of all sequences (y_n) such that $\text{card}\{n \mid |y_n - y_{n-1}| \neq 0\} \leq N$.

Chapter 8 is devoted to the UMD property. After a brief presentation of Burkholder's inequalities in the scalar case, we concentrate on their analogue for Banach space valued martingales (f_n). In the scalar case, when $1 < p < \infty$, we have

$$\sup_n \|f_n\|_p \simeq \|\sup |f_n|\|_p \simeq \|S(f)\|_p$$

where $S(f) = (|f_0|^2 + \sum (f_n - f_{n-1})^2)^{1/2}$, and where $A_p \simeq B_p$ means that there are positive constants C'_p and C''_p such that $C'_p A_p \leq B_p \leq C''_p A_p$. In the Banach space valued case, we replace $S(f)$ by:

$$R(f)(\omega) = \sup\nolimits_N \left(\int \left\| f_0(\omega) + \sum\nolimits_1^N \varepsilon_n (f_n - f_{n-1})(\omega) \right\|^2 d\mu \right)^{1/2} \tag{3}$$

where μ is the uniform probability measure on the set D of all choices of signs $(\varepsilon_n)_n$ with $\varepsilon_n = \pm 1$.

In §8.2 we prove Kahane's inequality, i.e. the equivalence of all the L_p-norms for series of the form $\sum_1^\infty \varepsilon_n x_n$ with x_n in an arbitrary Banach space when $0 < p < \infty$, see (8.13); in particular, up to equivalence, we can substitute to the L_2-norm in (3) any other L_p-norm for $p < \infty$.
Let $\{x_n\}$ be a sequence in a Banach space, such that the series $\sum \varepsilon_n x_n$ converges almost surely. We set

$$R(\{x_n\}) = \left(\int_D \left\| \sum \varepsilon_n x_n \right\|^2 d\mu \right)^{1/2}.$$

With this notation we have

$$R(f)(\omega) = R(\{f_0(\omega), f_1(\omega) - f_0(\omega), f_2(\omega) - f_1(\omega), \cdots\}). \tag{4}$$

The UMD_p and UMD properties are introduced in §8.3. Consider the series

$$\tilde{f}_\varepsilon = f_0 + \sum\nolimits_1^\infty \varepsilon_n (f_n - f_{n-1}). \tag{5}$$

By definition, when B is UMD_p, (f_n) converges in $L_p(B)$ iff (5) converges in $L_p(B)$ for all choices of signs $\varepsilon_n = \pm 1$ or equivalently iff it converges for almost all (ε_n). Moreover, we have then for $1 < p < \infty$ and all choices of signs $\varepsilon = (\varepsilon_n)$

$$\|\tilde{f}_\varepsilon\|_{L_p(B)} \simeq \|f\|_{L_p(B)} \tag*{$(3)_p$}$$

$$\sup_{n \geq 0} \|f_n\|_{L_p(B)} \simeq \|R(f)\|_p. \tag*{$(4)_p$}$$

See Proposition 8.9. The case $p = 1$ (due to Burgess Davis) is treated in §8.4. The main result of §8.3 is the equivalence of UMD_p and UMD_q for any $1 < p, q < \infty$. We give two proofs of this, the first one is based on distributional (also called "good λ") inequalities. This is an extrapolation principle that allows to show that, for a given Banach space B, $(3)_q \Rightarrow (3)_p$ for any $1 < p < q$. In the scalar case one starts from the case $q = 2$, that is obvious by orthogonality, and uses the preceding implication to deduce from it the case $1 < p < 2$ and then $2 < p < \infty$ by duality.

The second proof is based on Gundy's decomposition, that is a martingale version of the Calderón–Zygmund decomposition in classical harmonic analysis. There one proves a weak type (1,1) estimate and then invokes the Marcinkiewicz theorem to obtain the case $1 < p < 2$. We describe the latter in an appendix to Chapter 8.

In §8.6 we show that to check that a space B is UMD_p we may restrict ourselves to martingales adapted to the *dyadic* filtration and the associated UMD-constant remains the same.

In §8.7, we prove the Burkholder–Rosenthal inequalities. In the scalar case this boils down to the equivalence

$$\sup_n \|f_n\|_p \simeq \|\sigma(f)\|_p + \|\sup_n |f_n - f_{n-1}|\|_p$$

valid for $2 < p < \infty$.
Rosenthal originally proved this when f_n is a sum of independent variables and Burkholder extended it to martingales. We describe a remarkable example of complemented subspace of L_p (the Rosenthal space X_p) that motivated Rosenthal's work.

In §8.8, we describe Stein's inequality and its B-valued analogue when B is a UMD Banach space. Let $(\mathcal{A}_n)_{n \geq 0}$ be a filtration as usual, and let $(x_n)_{n \geq 0}$ be now an *arbitrary* sequence in L_p. Let $y_n = \mathbb{E}^{\mathcal{A}_n} x_n$. Stein's inequality asserts that for any $1 < p < \infty$ there is a constant C_p such that

$$\left\| \left(\sum |y_n|^2 \right)^{1/2} \right\|_p \leq C_p \left\| \left(\sum |x_n|^2 \right)^{1/2} \right\|_p, \tag{8}$$

for any (x_n) in L_p.

For x_n in $L_p(B)$, with B UMD the same result remains valid if we replace on both sides of (8) the expression $(\sum |x_n|^2)^{1/2}$ by

$$\left(\int \left\|\sum \varepsilon_n x_n\right\|_B^2 \, d\mu\right)^{1/2}.$$

See (8.53).

In §8.9, we discuss the space BMO and the B-valued version of H^1 in the martingale context. This leads naturally to the atomic version of B-valued H^1, denoted by $H^1_{at}(B)$. Its dual can be identified with a BMO-space for B^*-valued martingales, at least for a "regular" filtration $(\mathcal{A}_n)$. Equivalently, the space $H^1_{at}(B)$ can be identified with $H^1_{\max}(B)$ that is defined as the completion of $L_1(B)$ with respect to the norm $f \mapsto \mathbb{E}\sup_n \|f_n\|_B$ (here $f_n = \mathbb{E}^{\mathcal{A}_n} f$).

In §8.10, we describe Burkholder's geometric characterization of UMD spaces in terms of ζ-convexity (Theorem 8.47) but we prefer to give the full details of a more recent result (Theorem 8.48). The latter asserts that a real Banach space of the form $B = X \oplus X^*$ is UMD iff the function

$$x \oplus \xi \to \xi(x)$$

is the difference of two real valued convex continuous functions on B. After an already mentioned first appendix devoted to the Marcinkiewicz theorem, the second one collects several facts (to be used later on) on reverse Hölder inequalities. A typical result is that, when $0 < r < p < \infty$, if (Z_n) are i.i.d. copies of a random variable, then the sequence $\{n^{-1/p} \sup\limits_{1\le k\le n} |z_k| \mid n \ge 1\}$ is bounded in L_r iff Z is in weak-L_p, in other words iff $\sup\limits_{t>0} t^p \mathbb{P}\{|z| > t\} < \infty$. We call it reverse Hölder because the assumption is boundedness in L_r with $r < p$ and the conclusion is in weak-L_p (or $L_{p,\infty}$) and a fortiori in L_q for all $r < q < p$.

In Chapter 9, we give two characterizations of super-reflexive Banach spaces by properties of the underlying *metric spaces*. The relevant properties involve finite metric spaces. Given a sequence $\mathcal{T} = (T_n, d_n)$ of finite metric spaces, we say that the sequence $\mathcal{T}$ embeds Lipschitz uniformly in a metric space (T, d) if for some constant C there are subsets $\widetilde{T}_n \subset T$, and bijective mappings $f_n\colon\ T_n \to \widetilde{T}_n$ with Lipschitz norms satisfying

$$\sup_n \|f_n\|_{\mathrm{Lip}} \|f_n^{-1}\|_{\mathrm{Lip}} < \infty.$$

Consider for instance the case when T_n is a finite dyadic tree restricted to its first $1 + 2 + \cdots + 2^n = 2^{n+1} - 1$ points viewed as a graph and equipped with the usual geodesic distance. In Theorem 9.1, we prove following [86] that a Banach space B is super-reflexive iff it does not contain the sequence of these dyadic trees Lipschitz uniformly. More recently (cf. [173]), it was proved that the trees can be replaced in this result by the "*diamond graphs*". We describe the analogous characterization with diamond graphs in §9.2.

In §9.3, we discuss several non-linear notions of "type p" for metric spaces, notably the notion of Markov type p and we prove the recent result from [212] that p-uniformly smooth implies Markov type p. The proof uses martingale inequalities for martingales naturally associated to Markov chains on finite state spaces.

Acknowledgement I am very grateful to all those who helped me to correct mistakes, misprints and suggested improvements of all kind, in particular Julien Giol, Rostyslav Kravchenko, Javier Parcet, Yanqi Qiu, ...

Chapter 1

Banach space valued martingales

1.1 Banach space valued L_p-spaces. Conditional expectations

Let $(\Omega, \mathcal{A}, m)$ be a measure space. Let B be a Banach space. We will denote by $F(B)$ the space of all measurable step functions, i.e. the functions $f\colon \Omega \to B$ for which there is a partition of Ω, say $\Omega = A_1 \cup \ldots \cup A_N$ with $A_k \in \mathcal{A}$, and elements $b_k \in B$ such that

$$\forall \omega \in \Omega \qquad f(\omega) = \sum\nolimits_1^N 1_{A_k}(\omega) b_k. \tag{1.1}$$

Equivalently, $F(B)$ is the space of all measurable functions $f\colon \Omega \to B$ taking only finitely many values.

Definition. We will say that a function $f\colon \Omega \to B$ is Bochner measurable if there is a sequence (f_n) in $F(B)$ tending to f pointwise.

We will denote by $L_0(\Omega, \mathcal{A}, m; B)$ the set of equivalence classes (modulo equality almost everywhere) of Bochner measurable functions.

Let $1 \le p \le \infty$. We will denote by $L_p(\Omega, \mathcal{A}, m; B)$ the subspace of $L_0(\Omega, \mathcal{A}, m)$ formed of all the functions f such that $\int \|f\|_B^p \, dm < \infty$ for $p < \infty$, and $\text{ess sup}\|f(\cdot)\|_B < \infty$ for $p = \infty$. We equip this space with the norm

$$\|f\|_{L_p(B)} = \left(\int \|f\|_B^p \, dm \right)^{1/p} \quad \text{for} \quad p < \infty$$
$$\|f\|_{L_\infty(B)} = \text{ess sup}\|f(\cdot)\|_B \quad \text{for} \quad p = \infty,$$

with which it becomes a Banach space.

Of course, this definition coincides with the usual one in the scalar valued case i.e. if $B = \mathbb{R}$ (or $\mathbb{C}$). In that case, we often denote simply by $L_p(\Omega, \mathcal{A}, m)$ (or sometimes $L_p(m)$, or even L_p) the resulting space of scalar valued functions.

For brievity, we will often write simply $L_p(\mathbb{P}; B)$ or, if there is no risk of confusion, simply $L_p(B)$ instead of $L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$.

Given $\varphi_1, \dots, \varphi_N \in L_p$ and $b_1, \dots, b_N \in B$ we can define a function $f\colon\ \Omega \to B$ in $L_p(B)$ by setting $f(\omega) = \sum_1^N \varphi_k(\omega) b_k$. We will denote this function by $\sum_1^N \varphi_k \otimes b_k$ and by $L_p \otimes B$ the subspace of $L_p(B)$ formed of all such functions.

Proposition 1.1. *Let $1 \le p < \infty$.*

(i) *$F(B) \cap L_p(B)$ is dense in $L_p(B)$.*

(ii) *The subspace $L_p \otimes B \subset L_p(B)$ is dense in $L_p(B)$.*

Proof. Consider $f \in L_p(B)$. Let $f_n \in F(B)$ be such that $f_n \to f$ pointwise. Then $\|f_n(\cdot)\|_B \to \|f(\cdot)\|_B$ pointwise, so that if we set $g_n(\omega) = f_n(\omega) 1_{\{\|f_n\| < 2\|f\|\}}$ we still have $g_n \to f$ pointwise and in addition $\sup_n \|g_n - f\| \le \sup_n \|g_n\| + \|f\| \le 3\|f\|$. Therefore, by dominated convergence, we must have $\int \|g_n - f\|_B^p \, dm \to 0$ and of course $g_n \in F(B) \cap L_p(B)$. This proves (i). The second point is then obvious since $F(B) \cap L_p(B) \subset L_p \otimes B$ (indeed we can take $\varphi_k = 1_{A_k}$ with $m(A_k) < \infty$, as in (1.1)). $\square$

Remark 1.2. If B is finite dimensional, then $F(B)$ is dense in $L_\infty(B)$ but this is no longer true in the infinite dimensional case, because the unit ball of B is not compact.

We now turn to the definition of the integral of a function in $L_1(B)$. Consider a function f of the form (1.1) in $L_1(B) \cap F(B)$. We define

$$\int f \, dm = \sum\nolimits_1^N m(A_k) b_k.$$

This defines a continuous linear map from $L_1(B) \cap F(B)$ to B, since we have obviously by the triangle inequality

$$\left\| \int f \, dm \right\| \le \sum m(A_k) \|b_k\| = \|f\|_{L_1(B)}.$$

By density, this linear map admits an extension defined on the whole of $L_1(B)$, that we still denote by $\int f \, dm$ when $f \in L_1(B)$. The extension clearly satisfies the following fundamental inequality called Jensen's inequality

$$\forall f \in L_1(B) \qquad \left\| \int f \, dm \right\|_B \le \int \|f\|_B \, dm = \|f\|_{L_1(B)}. \tag{1.2}$$

This extends the linear map $f \to \int f \, dm$ from the scalar valued case to the B-valued one. More generally: Let $(\Omega', \mathcal{A}', m')$ be another measure space and let $T\colon\ L_1(\Omega, \mathcal{A}, m) \to L_1(\Omega', \mathcal{A}', m')$ be a bounded operator. We may clearly define unambiguously a linear operator $T_0\colon\ F(B)\cap L_1(m;B)\to L_1(m',B)$ by setting for any f of the form (1.1)

$$T_0(f)=\sum\nolimits_1^N T(1_{A_k})b_k.$$

We have clearly by the triangle inequality

$$\|T_0(f)\|_{L_1(m';B)}\le\sum\nolimits_1^N\|T(1_{A_k})\|\|b_k\|\le\|T\|\sum m(A_k)\|b_k\|=\|T\|\|f\|_{L_1(B)}.$$

Thus, we can state

Proposition 1.3. *Given a bounded operator $T\colon\ L_1(\Omega,\mathcal{A},m)\to L_1(\Omega',\mathcal{A}',m')$, there is a unique bounded linear map $\widetilde{T}\colon\ L_1(\Omega,\mathcal{A},m;B)\to L_1(\Omega',\mathcal{A}',m';B)$ such that*

$$\forall\varphi\in L_1(\Omega,\mathcal{A},m)\ \forall b\in B\qquad \widetilde{T}(\varphi\otimes b)=T(\varphi)b. \tag{1.3}$$

Moreover, we have $\|\widetilde{T}\|=\|T\|$.

Proof. By the density of $F(B)\cap L_1(B)$ in $L_1(B)$, the (continuous) map T_0 admits a unique continuous linear extension $\widetilde{T}$ from $L_1(m;B)$ to $L_1(m';B)$, with $\|\widetilde{T}\|\le\|T_0\|\le\|T\|$. If φ is a step function in L_1, then (1.3) is clear by definition of T_0. Approximating φ in L_1 by a step function, we see that (1.3) is true in general. The unicity of $\widetilde{T}$ is clear since (1.3) implies that $\widetilde{T}$ coincides with T_0 on $F(B)\cap L_1(B)$. Finally, considering a fixed b with $\|b\|=1$, we easily derive from (1.3) that $\|T\|\le\|\widetilde{T}\|$, so we obtain $\|T\|=\|\widetilde{T}\|$. □

We start by recalling some well known properties of conditional expectations. Let $(\Omega,\mathcal{A},\mathbb{P})$ be a probability space and let $\mathcal{B}\subset\mathcal{A}$ be a σ-subalgebra. The conditional expectation $f\to\mathbb{E}^{\mathcal{B}}f$ is a positive contraction on $L_p(\Omega,\mathcal{A},\mathbb{P})$ for all $1\le p\le\infty$. It is characterized by the property

$$\begin{gathered}\forall h\in L_\infty(\Omega,\mathcal{B},\mathbb{P})\qquad\forall f\in L_p(\Omega,\mathcal{A},\mathbb{P})\\ \mathbb{E}^{\mathcal{B}}(hf)=h\mathbb{E}^{\mathcal{B}}(f).\end{gathered}$$

On $L_2(\Omega,\mathcal{A},\mathbb{P})$, the conditional expectation $\mathbb{E}^{\mathcal{B}}$ coincides with the orthogonal projection onto the subspace $L_2(\Omega,\mathcal{B},\mathbb{P})$.

It is not true in general that a bounded operator on L_p extends boundedly to $L_p(B)$ as in the preceding Proposition for $p=1$. Nervertheless, it is true for *positive* operators. The conditional expectation of a vector valued function can be defined using that fact, as follows.

Proposition 1.4. *let $1\le p,q\le\infty$. Let $(\Omega,\mathcal{A},\mathbb{P})$ be an arbitrary measure space and let $T\colon\ L_p(\Omega)\to L_q(\Omega)$ be a bounded linear operator. Clearly, there is a unique linear operator*

$$T\otimes I_B\colon\ L_p(\Omega,\mathbb{P})\otimes B\to L_q(\Omega,\mathbb{P})\otimes B$$

such that

$$\forall \varphi \in L_p(\Omega, \mathbb{P}) \quad \forall x \in B \qquad (T \otimes I_B)(\varphi \otimes x) = T(\varphi) \otimes x.$$

Now, if T is positive (i.e. if T preserves nonnegative functions) then $T \otimes I_B$ extends to a bounded operator $\widetilde{T \otimes I_B}$ from $L_p(\Omega, \mathbb{P}; B)$ to $L_q(\Omega, \mathbb{P}; B)$ which has the same norm as T, i.e.

$$\|\widetilde{T \otimes I_B}\|_{L_p(B) \to L_q(B)} = \|T\|_{L_p \to L_q}.$$

Proof. It clearly suffices to show that

$$\text{(1.4)} \qquad \forall f \in L_p(\Omega, \mathbb{P}) \otimes B \qquad \|(T \otimes I_B) f(\cdot)\|_B \overset{\text{a.s.}}{\leq} T(\|f(\cdot)\|_B).$$

For that purpose, we can assume B separable (or even finite dimensional) so that there is a countable subset $D \subset B^*$ verifying

$$\forall x \in B \qquad \|x\| = \sup_{\xi \in D} |\xi(x)|.$$

Clearly for any ξ in B^* we have

$$\langle \xi, (T \otimes I_B) f(\cdot) \rangle = T(\langle \xi, f(\cdot) \rangle)$$

and hence by the positivity of T for any finite subset $D' \subset D$

$$\sup_{\xi \in D'} |\langle \xi, (T \otimes I_B) f(\cdot) \rangle| \overset{\text{a.s.}}{\leq} T(\sup_{\xi \in D} |\langle \xi, f(\cdot) \rangle|)$$

therefore we obtain (1.4) and the proposition follows. □

Remark. Let B_1 be another Banach space and let $u \colon B \to B_1$ be a bounded operator. Then for any f in $L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$ we have

$$\widetilde{T \otimes I_{B_1}}(u(f)) = u[\widetilde{T \otimes I_B}(f)].$$

In particular, for any ξ in B^* we have

$$\text{(1.5)} \qquad T(\xi(f)) = \xi(T(f)).$$

Indeed, this is immediately checked for f in $L_p(\Omega, \mathbb{P}) \otimes B$, and the general case is obtained after completion.

Note that now that $\widetilde{T \otimes I_B}$ makes sense, the preceding argument can be repeated to show that

$$\text{(1.6)} \qquad \forall f \in L_p(\Omega, \mathbb{P}; B) \qquad \|(\widetilde{T \otimes I_B}) f\|_B \overset{\text{a.s.}}{\leq} T(\|f(\cdot)\|_B).$$

A priori, in the above (1.6) we implicitly assume that B is a real Banach space, but actually if B is a complex space (and T is $\mathbb{C}$-linear on complex valued L_1), we may consider B a fortiori as a real space and then (1.6) remains valid.

In particular, the preceding proposition applies for $T = \mathbb{E}^{\mathcal{B}}$. For any f in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$ we will denote again simply by $\mathbb{E}^{\mathcal{B}}(f)$ the function $\widetilde{T \otimes I_B}(f)$ for $T = \mathbb{E}^{\mathcal{B}}$. Note that $g = \mathbb{E}^{\mathcal{B}}(f)$ is characterized by the following properties

(i) $g \in L_1(\Omega, \mathcal{B}, \mathbb{P}; B)$

(ii) $\forall E \in \mathcal{B} \int_E g d\mathbb{P} = \int_E f d\mathbb{P}$.

Indeed, this is easy to check by 'scalarization', since it holds in the scalar case. More precisely, a B-valued function g has these properties iff for any ξ in B^* the scalar valued function $\langle \xi, g(\cdot) \rangle$ has similar properties, or equivalently $\langle \xi, g(\cdot) \rangle = \mathbb{E}^{\mathcal{B}} \langle \xi, f \rangle$, and hence by (1.5) $\langle \xi, g \rangle = \langle \xi, \mathbb{E}^{\mathcal{B}} f \rangle$ which means $g = \mathbb{E}^{\mathcal{B}} f$ as announced.

1.2 Martingales: basic properties

Let B be a Banach space. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. A sequence $(M_n)_{n \geq 0}$ in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$ is called a martingale if there exists an increasing sequence of σ-subalgebras $\mathcal{A}_0 \subset \mathcal{A}_1 \subset \cdots \subset \mathcal{A}_n \subset \cdots \subset \mathcal{A}$ (this is called "a filtration") such that for each $n \geq 0$ M_n is $\mathcal{A}_n$-measurable and satisfies

$$M_n = \mathbb{E}^{\mathcal{A}_n}(M_{n+1}).$$

For the precise definition of the conditional expectation in the Banach space valued case, see the above Proposition 1.4. This implies of course that

$$\forall n < m \qquad M_n = \mathbb{E}^{\mathcal{A}_n} M_m.$$

In particular if (M_n) is a B-valued martingale, the above property (ii) (in the preceding section) yields in the case $\mathcal{B} = \mathcal{A}_n$ and $n \leq m$

$$\forall n \leq m \quad \forall A \in \mathcal{A}_n \qquad \int_A M_n d\mathbb{P} = \int_A M_m d\mathbb{P}. \tag{1.7}$$

A sequence of random variables (M_n) is called adapted to the filtration $(\mathcal{A}_n)_{n \geq 0}$ if M_n is $\mathcal{A}_n$-measurable for each $n \geq 0$. Note that the martingale property $M_n = \mathbb{E}^{\mathcal{A}_n}(M_{n+1})$ automatically implies that (M_n) is adapted. Of course, the minimal choice of $\mathcal{A}_n$ is simply $\mathcal{A}_n = \sigma(M_0, M_1, \ldots, M_n)$.

We will also need the definition of a submartingale. A sequence $(M_n)_{n \geq 0}$ of scalar valued random variables in L_1 is called a submartingale if there are σ-subalgebras $\mathcal{A}_n$ as above such that M_n is $\mathcal{A}_n$-measurable and

$$\forall n \geq 0 \qquad M_n \leq \mathbb{E}^{\mathcal{A}_n} M_{n+1}.$$

This implies of course that

$$\forall n < m \qquad M_n \leq \mathbb{E}^{\mathcal{A}_n} M_m.$$

More generally, if I is any partially ordered set, then a collection $(M_i)_{i\in I}$ in $L_1(\Omega,\mathbb{P};B)$ is called a martingale (indexed by I) if there are σ-subalgebras $\mathcal{A}_i \subset \mathcal{A}$ such that $\mathcal{A}_i \subset \mathcal{A}_j$ whenever $i<j$ and $M_i = \mathbb{E}^{\mathcal{A}_i} M_j$.

In particular, when

$$I = \{0,-1,-2,\ldots\}$$

is the set of all negative integers, the corresponding sequence is usually called a reverse martingale.

The following convergence theorem is fundamental.

Theorem 1.5. *Let $(\mathcal{A}_n)$ be a fixed increasing sequence of σ-subalgebras of $\mathcal{A}$. Let $\mathcal{A}_\infty$ be the σ-algebra generated by $\bigcup_{n\geq 0}\mathcal{A}_n$. Let $1\leq p<\infty$ and consider M in $L_p(\Omega,\mathbb{P};B)$. Let us define $M_n = \mathbb{E}^{\mathcal{A}_n}(M)$. Then $(M_n)_{n\geq 0}$ is a martingale such that $M_n \to \mathbb{E}^{\mathcal{A}_\infty}(M)$ in $L_p(\Omega,\mathbb{P};B)$ when $n\to\infty$.*

Proof. Note that since $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ we have $\mathbb{E}^{\mathcal{A}_n}\mathbb{E}^{\mathcal{A}_{n+1}} = \mathbb{E}^{\mathcal{A}_n}$, and similarly $\mathbb{E}^{\mathcal{A}_n}\mathbb{E}^{\mathcal{A}_\infty} = \mathbb{E}^{\mathcal{A}_n}$. Replacing M by $\mathbb{E}^{\mathcal{A}_\infty}M$ we can assume w.l.o.g. that M is $\mathcal{A}_\infty$-measurable. We will use the following fact: the union $\bigcup_n L_p(\Omega,\mathcal{A}_n,\mathbb{P};B)$ is dense in $L_p(\Omega,\mathcal{A}_\infty,\mathbb{P};B)$. Indeed, let $\mathcal{C}$ be the class of all sets A such that $1_A \in \overline{\bigcup_n L_\infty(\Omega,\mathcal{A}_n,\mathbb{P})}$, where the closure is meant in $L_p(\Omega,\mathbb{P})$ (recall $p<\infty$). Clearly $\mathcal{C}\supset \bigcup_{n\geq 0}\mathcal{A}_n$ and $\mathcal{C}$ is a σ-algebra hence $\mathcal{C}\supset\mathcal{A}_\infty$. This gives the scalar case version of the above fact. Now, any f in $L_p(\Omega,\mathcal{A}_\infty,\mathbb{P};B)$ can be approximated (by definition of the spaces $L_p(B)$) by functions of the form $\sum_1^n 1_{A_i}x_i$ with $x_i\in B$ and $A_i\in\mathcal{A}_\infty$. But since $1_{A_i}\in\overline{\bigcup_n L_\infty(\Omega,\mathcal{A}_n,\mathbb{P})}$ we clearly have $f\in\overline{\bigcup_n L_p(\Omega,\mathcal{A}_n,\mathbb{P};B)}$ as announced.

We can now prove Theorem 1.5. Let $\varepsilon>0$. By the above fact there is an integer k and g in $L_p(\Omega,\mathcal{A}_k,\mathbb{P};B)$ such that $\|M-g\|_p<\varepsilon$. We have then $g=\mathbb{E}^{\mathcal{A}_n}g$ for all $n\geq k$, hence

$$\forall n\geq k \qquad M_n - M = \mathbb{E}^{\mathcal{A}_n}(M-g)+g-M$$

and finally

$$\begin{aligned}\|M_n-M\|_p &\leq \|\mathbb{E}^{\mathcal{A}_n}(M-g)\|_p + \|g-M\|_p\\ &\leq 2\varepsilon.\end{aligned}$$

This completes the proof. □

Corollary 1.6. *In the scalar case (or the f.d. case) every martingale which is bounded in L_p for some $1\leq p<\infty$ and which is uniformly integrable if $p=1$ is actually convergent in L_p to a limit M_∞ such that $M_n=\mathbb{E}^{\mathcal{A}_n}M_\infty \quad \forall n\geq 0$.*

Proof. Let (M_{n_k}) be a subsequence converging weakly to a limit which we denote by M_∞. Clearly $M_\infty \in L_p(\Omega, \mathcal{A}_\infty, \mathbb{P})$ and we have $\forall A \in \mathcal{A}_n$

$$\int_A M_\infty d\mathbb{P} = \lim \int_A M_{n_k} d\mathbb{P},$$

but whenever $n_k \geq n$, we have $\int_A M_{n_k} d\mathbb{P} = \int_A M_n d\mathbb{P}$ by the martingale property. Hence

$$\forall A \in \mathcal{A}_n \qquad \int_A M_\infty d\mathbb{P} = \int_A M_n d\mathbb{P}$$

which forces $M_n = \mathbb{E}^{\mathcal{A}_n} M_\infty$. We then conclude by Theorem 1.5 that $M_n \to M_\infty$ in L_p-norm. $\square$

Note that conversely any martingale which converges in L_1 is clearly uniformly integrable.

Remark 1.7. Fix $1 \leq p < \infty$. Let I be a directed set, with order denoted simply by $\leq$. This means that for any pair i, j in I there is $k \in I$ such that $i \leq k$ and $j \leq k$. Let $(\mathcal{A}_i)$ be a family of σ-algebras directed by inclusion (i.e. we have $\mathcal{A}_i \subset \mathcal{A}_j$ whenever $i \leq j$). The extension of the notion of martingale is obvious: A collection of random variables $(f_i)_{i \in I}$ in $L_p(B)$ will be called a martingale if $f_i = \mathbb{E}^{\mathcal{A}_i}(f_j)$ holds whenever $i \leq j$. The resulting net converges in $L_p(B)$ iff for any increasing sequence $i_1 \leq \cdots \leq i_n \leq i_{n+1} \leq \cdots$, the (usual sense) martingale (f_{i_n}) converges in $L_p(B)$. Indeed, this merely follows from the metrizability of $L_p(B)$! More precisely, if we assume that $\sigma\Big(\bigcup_{i \in I} \mathcal{A}_i\Big) = \mathcal{A}$, then for any f in $L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$, the directed net $(\mathbb{E}^{\mathcal{A}_i} f)_{i \in I}$ converges to f in $L_p(B)$. Indeed, this net must satisfy the Cauchy criterion, because otherwise we would be able for some $\delta > 0$ to construct (by induction) an increasing sequence $i(1) \leq i(2) \leq \ldots$ in I such that $\|\mathbb{E}^{\mathcal{A}_{i(k)}} f - \mathbb{E}^{\mathcal{A}_{i(k-1)}} f\|_{L_p(B)} > \delta$ for all $k > 1$, and this would then contradict Theorem 1.5. Thus, $\mathbb{E}^{\mathcal{A}_i} f$ converges to a limit F in $L_p(B)$, and hence for any set $A \subset \Omega$ in $\bigcup_{j \in I} \mathcal{A}_j$ we must have

$$\int_A f = \lim_i \int_A \mathbb{E}^{\mathcal{A}_i} f = \int_A F.$$

Since the equality $\int_A f = \int_A F$ must remain true on the σ-algebra generated by $\bigcup \mathcal{A}_j$, we conclude that $f = F$, thus completing the proof that $\mathbb{E}^{\mathcal{A}_i} f \to f$ in $L_p(B)$.

1.3 Examples of filtrations

The most classical example of filtration is the one associated to a sequence of independent (real valued) random variables $(Y_n)_{n \geq 1}$ on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let $\mathcal{A}_n = \sigma(Y_1, \ldots, Y_n)$ for all $n \geq 1$ and $\mathcal{A}_0 = \{\phi, \Omega\}$. In that case, a sequence of random variables $(f_n)_{n \geq 0}$ is adapted to the filtration $(\mathcal{A}_n)_{n \geq 0}$ iff

f_0 is constant and, for each $n \geq 1$, f_n depends only on $Y_1, \ldots, Y_n$, i.e. there is a (Borel measurable) function F_n on $\mathbb{R}^n$ such that

$$f_n = F_n(Y_1, \ldots, Y_n).$$

The martingale condition can then be written as

$$\forall n \geq 0 \qquad F_n(Y_1, \ldots, Y_n) = \int F_{n+1}(Y_1, \ldots, Y_n, y) \quad d\mathbb{P}_n(y)$$

where $\mathbb{P}_n$ is the probability distribution (or "the law") of Y_{n+1}.

An equivalent but more "intrinsic" model arises when one considers $\Omega = \mathbb{R}^{\mathbb{N}_*}$ equipped with the product probability measure $\mathbb{P} = \bigotimes_{n \geq 1} \mathbb{P}_n$. If one denotes by $Y = (Y_n)_{n \geq 1}$ a generic point in Ω, the random variable $Y \to Y_n$ appears as the n-th coordinate, and $Y \to F_n(Y)$ is $\mathcal{A}_n$-measurable iff $F_n(Y)$ depends only on the n first coordinates of Y.

The *dyadic filtration* $(\mathcal{D}_n)_{n \geq 0}$ on $D = \{-1, 1\}^{\mathbb{N}_*}$ is the fundamental example of this kind: Here we denote by

$$\varepsilon_n \colon\ D \to \{-1, 1\} \qquad (n = 1, 2, \ldots)$$

the n-th coordinate, we equip D with the probability measure

$$\mu = \otimes(\delta_1 + \delta_{-1})/2,$$

and we set $\mathcal{D}_n = \sigma(\varepsilon_1, \ldots, \varepsilon_n)$, $\mathcal{D}_0 = \{\phi, D\}$.
Clearly, the variables (ε_n) are independent on $(D, \mathcal{D}, \mu)$ and take the values ± 1 with equal probability $1/2$.
Note that $\mathcal{D}_n$ admits exactly 2^n atoms and moreover $\dim L_2(D, \mathcal{D}_n, \mu) = 2^n$. For any finite subset $A \subset [1, 2, \ldots]$, let $w_A = \prod_{n \in A} \varepsilon_n$ with the convention $w_\phi \equiv 1$. It is easy to check that $\{w_A \mid A \subset [1, \ldots, n]\}$ (resp. $\{w_A \mid |A| < \infty\}$) is an orthonormal basis of $L_2(D, \mathcal{D}_n, \mu)$ (resp. $L_2(D, \mathcal{D}, \mu)$).
Given a Banach space B, a B-valued martingale $f_n \colon\ D \to B$ adapted to the dyadic filtration $(\mathcal{D}_n)$ is caracterized by the property that

$$\forall n \geq 1 \qquad (f_n - f_{n-1})(\varepsilon_1, \ldots, \varepsilon_n) = \varepsilon_n \varphi_{n-1}(\varepsilon_1, \ldots, \varepsilon_{n-1}),$$

where φ_{n-1} depends only on $\varepsilon_1, \ldots, \varepsilon_{n-1}$. We leave the easy verification of this to the reader.
Of course the preceding remarks remain valid if one works with any sequence of ± 1-valued independent random variables (ε_n) such that $\mathbb{P}(\varepsilon_n = \pm 1) = 1/2$ on an "abstract" probability space $(\Omega, \mathbb{P})$.

In classical analysis, it is customary to use the Rademacher functions $(r_n)_{n \geq 1}$ on the Lebesgue interval $([0, 1], dt)$ instead of (ε_n). We need some notation to introduce these. Given an interval $I \subset \mathbb{R}$ we divide I into parts of equal length and we denote by I^+ and I^- respectively the left and right half of I. Note that we do not specify whether the end points belong to I since the latter are negligible for the Lebesgue measure on $[0, 1]$ (or $[0, 1[$ or $\mathbb{R}$). Let

$$h_I = 1_{I^+} - 1_{I^-}.$$

We denote $I_1(1) = [0,1[$, $I_2(1) = [0, \frac{1}{2}[$, $I_2(2) = \left[\frac{1}{2}, 1\right[$ and more generally

$$I_n(k) = \left[\frac{k-1}{2^{n-1}}, \frac{k}{2^{n-1}}\right[$$

for $k = 1, 2, \ldots, 2^{n-1}$ $(n \geq 1)$.
We then set $h_1 \equiv 1, h_2 = h_{I_1(1)},\ h_3 = \sqrt{2}\ h_{I_2(1)},\ h_4 = \sqrt{2}\ h_{I_2(2)}$ and more generally

$$\forall n \geq 1\ \forall k = 1, \ldots, 2^{n-1} \qquad h_{2^{n-1}+k} = |I_n(k)|^{-1/2} h_{I_n(k)}.$$

Note that $\|h_n\|_2 = 1$ for all $n \geq 1$.

The Rademacher function r_n can be defined, for each $n \geq 1$, by

$$r_n = \sum\nolimits_{k=1}^{2^{n-1}} h_{I_n(k)}.$$

Then the sequence $(r_n)_{n\geq 1}$ has the same distribution on $([0,1], dt)$ as the sequence $(\varepsilon_n)_{n\geq 1}$ on (D, μ). Let $\mathcal{A}_n = \sigma(r_1, \ldots, r_n)$. Then $\mathcal{A}_n$ is generated by the 2^n-atoms $\{I_{n+1}(k) \mid 1 \leq k \leq 2^n\}$, each having length 2^{-n}. The dimension of $L_2([0,1], \mathcal{A}_n)$ is 2^n and the functions $\{h_1, \ldots, h_{2^n}\}$ (resp. $\{h_n \mid n \geq 1\}$) form an orthonormal basis of $L_2([0,1], \mathcal{A}_n)$ (resp. $L_2([0,1])$).

The *"Haar filtration"* $(\mathcal{B}_n)_{n\geq 1}$ on $[0,1]$ is defined by

$$\mathcal{B}_n = \sigma(h_1, \ldots, h_n),$$

so that we have $\sigma(h_1, \ldots, h_{2^n}) = \sigma(r_1, \ldots, r_n)$ or equivalently $\mathcal{B}_{2^n} = \mathcal{A}_n$ for all $n \geq 1$ (note that here $\mathcal{B}_1$ is trivial). It is easy to check that $\mathcal{B}_n$ is an atomic σ-algebra, with exactly n atoms. Since the conditional expectation $\mathbb{E}^{\mathcal{B}_n}$ is the orthogonal projection from L_2 to $L_2(\mathcal{B}_n)$, we have for any f in $L_2([0,1])$

$$\forall n \geq 1 \qquad \mathbb{E}^{\mathcal{B}_n} f = \sum\nolimits_1^n \langle f, h_k \rangle h_k$$

and hence for all $n \geq 2$

$$\mathbb{E}^{\mathcal{B}_n} f - \mathbb{E}^{\mathcal{B}_{n-1}} f = \langle f, h_n \rangle h_n. \tag{1.8}$$

More generally for any B-valued martingale $(f_n)_{n\geq\emptyset}$ adapted to $(\mathcal{B}_n)_{n\geq 1}$ we have

$$\forall n \geq 2 \qquad f_n - f_{n-1} = h_n x_n$$

for some sequence (x_n) in B. The Haar functions are in some sense the first example of wavelets (see e.g. [64]). Indeed, if we set

$$h = 1_{[0,\frac{1}{2}[} - 1_{[\frac{1}{2},1[}$$

(this is the same as the function previously denoted by h_2), then the system of functions

$$\left\{2^{\frac{m}{2}} h((t+k)2^m) \mid k, m \in \mathbb{Z}\right\} \tag{1.9}$$

is an orthonormal basis of $L_2(\mathbb{R})$. Note that the constant function 1 is not in $L_2(\mathbb{R})$.

In the system (1.9), the sequence $\{h_n \mid n \geq 1\}$ coincides with the subsystem formed of all functions in (1.9) with support included in $[0,1]$.

1.4 Almost sure convergence and maximal inequalities

To handle the a.s. convergence of martingales, we will need (as usual) the appropriate maximal inequalities. In the martingale case, these are Doob's inequalities. Their proof uses *stopping times* which are a basic tool in martingale theory. Given an increasing sequence $(\mathcal{A}_n)_{n\geq 0}$ of σ-subalgebras on Ω, a random variable $T\colon\ \Omega \to \mathbb{N} \cup \{\infty\}$ is called a stopping time if

$$\forall n \geq 0 \qquad \{T \leq n\} \in \mathcal{A}_n,$$

or equivalently if

$$\forall n \geq 0 \qquad \{T = n\} \in \mathcal{A}_n.$$

If $T < \infty$ a.s., then T is called a finite stopping time.

Proposition 1.8. *For any martingale $(M_n)_{n\geq 0}$ relative to $(\mathcal{A}_n)_{n\geq 0}$ and for every stopping time T, let us denote by $M_{n\wedge T}$ the variable $M_{n\wedge T(\omega)}(\omega)$. Then $(M_{n\wedge T})_{n\geq 0}$ is a martingale relative to $(\mathcal{A}_n)_{n\geq 0}$.*

Proof. Observe that $M_{n\wedge T}$ clearly is in L_1 (since M_n is always assumed in L_1). Moreover, we have

$$M_{n\wedge T} - M_{(n-1)\wedge T} = 1_{\{n\leq T\}}(M_n - M_{n-1}),$$

but $\{n \leq T\}^c = \{T < n\} \in \mathcal{A}_{n-1}$ hence

$$\mathbb{E}^{\mathcal{A}_{n-1}}(M_{n\wedge T} - M_{(n-1)\wedge T}) = 1_{\{n\leq T\}}\mathbb{E}^{\mathcal{A}_{n-1}}(M_n - M_{n-1}) = 0.$$

□

Given a stopping time T, we can define the associated σ-algebra $\mathcal{A}_T$ as follows: we say that a set A in $\mathcal{A}$ belongs to $\mathcal{A}_T$ if $A \cap \{T \leq n\}$ belongs to $\mathcal{A}_n$ for each $n \geq 0$. Then $\mathcal{A}_T$ is a σ-algebra.

Exercises. (i) Consider M_∞ in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$ and let $M_n = \mathbb{E}^{\mathcal{A}_n} M_\infty$ be the associated martingale. Then if T is a stopping time, we have

$$M_T = \mathbb{E}^{\mathcal{A}_T}(M_\infty).$$

Moreover,

$$\mathbb{E}^{\mathcal{A}_n}(M_T) = M_{T\wedge n} = \mathbb{E}^{\mathcal{A}_T}(M_n). \tag{1.10}$$

More generally, if S is any other stopping time, $T \wedge S$ and $T \vee S$ are stopping times and we have

$$\mathbb{E}^{\mathcal{A}_S}(M_T) = M_{T\wedge S} = \mathbb{E}^{\mathcal{A}_T}(M_S).$$

(ii) If $(M_n)_{n\geq 0}$ is a martingale in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$ and if $T_0 \leq T_1 \leq \ldots$ is a sequence of *bounded* stopping times then $(M_{T_k})_{k\geq 0}$ is a martingale relative to the sequence of σ-algebras $\mathcal{A}_{T_0} \subset \mathcal{A}_{T_1} \subset \cdots$. This also holds for unbounded times if we assume as in (i) that $(M_n)_{n\geq 0}$ converges in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$.

Theorem 1.9 (Doob's maximal inequalities)**.** *Let $(M_0, M_1, \dots, M_n)$ be a submartingale in L_1, and let $M_n^* = \sup_{k \le n} M_k$. Then*

$$\forall t > 0 \qquad t\mathbb{P}(\{M_n^* > t\}) \le \int_{\{M_n^* > t\}} M_n d\mathbb{P}, \tag{1.11}$$

and if $M_n^* \ge 0$ then for all $1 < p < \infty$ we have

$$\|M_n^*\|_p \le p' \|M_n\|_p \tag{1.12}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$.

Proof. We can rewrite the submartingale property as saying that for any A in $\mathcal{A}_k$ with $k \le n$ we have

$$\int_A M_k d\mathbb{P} \le \int_A (E^{\mathcal{A}_k} M_n) d\mathbb{P} = \int E^{\mathcal{A}_k}(1_A M_n) d\mathbb{P} = \int_A M_n d\mathbb{P}. \tag{1.13}$$

Fix $t > 0$. Let

$$T = \begin{cases} \inf\{k \le n \mid M_k > t\} & \text{if } M_n^* > t, \\ \infty & \text{otherwise.} \end{cases}$$

Then T is a stopping time relative to the sequence of σ-algebras $(\mathcal{A}'_k)$ defined by $\mathcal{A}'_k = \mathcal{A}_{k \wedge n}$. We have since $M_k > t$ on the set $\{T = k\}$

$$t\mathbb{P}\{M_n^* > t\} = t\mathbb{P}\{T \le n\} = t \sum_{k \le n} \mathbb{P}\{T = k\} \le \sum_{k \le n} \int_{\{T=k\}} M_k$$

hence by (1.13)

$$\le \sum_{k \le n} \int_{\{T=k\}} M_n = \int_{\{T \le n\}} M_n.$$

This proves (1.11). To prove (1.12) we use an extrapolation trick. We have if $M_n^* \ge 0$

$$\begin{aligned} \mathbb{E} M_n^{*p} &= \int_0^\infty p t^{p-1} \mathbb{P}\{M_n^* > t\} dt \\ &\le \int_0^\infty p t^{p-2} \int_{\{M_n^* > t\}} M_n d\mathbb{P} \quad dt \\ &= \int M_n \left(\int_0^{M_n^*} p t^{p-2} dt \right) d\mathbb{P} = \int \frac{p}{p-1} M_n (M_n^*)^{p-1} \quad d\mathbb{P} \end{aligned}$$

hence by Hölder's inequality

$$\begin{aligned} &\le p' \|M_n\|_p \|(M_n^*)^{p-1}\|_{p'} \\ &= p' \|M_n\|_p (\mathbb{E} M_n^{*p})^{\frac{p-1}{p}}, \end{aligned}$$

so that after division by $(\mathbb{E} M_n^{*p})^{\frac{p-1}{p}}$ we obtain (1.12). □

The following inequality is known as the Burkholder–Davis–Gundy inequality. It is dual to Doob's maximal inequality. Indeed, by (1.12) we have for any x in L_p

$$\|(\mathbb{E}_n x)\|_{L_p(\ell_\infty)} = \|\sup_n |\mathbb{E}_n x|\|_p \leq p'\|x\|_p \tag{1.14}$$

therefore it is natural to expect a dual inequality involving an "adjoint mapping" from $L_{p'}(\ell_1)$ to $L_{p'}$, as follows.

Theorem 1.10. *Let $(\theta_n)_{n\geq 0}$ be an arbitrary family of random variables. Then for any $1 \leq p < \infty$*

$$\left\|\sum |\mathbb{E}^{\mathcal{A}_n}\theta_n|\right\|_p \leq p \left\|\sum |\theta_n|\right\|_p. \tag{1.15}$$

In particular if $\theta_n \geq 0$

$$\left\|\sum \mathbb{E}^{\mathcal{A}_n}\theta_n\right\|_p \leq p \left\|\sum \theta_n\right\|_p.$$

Proof. Since $|\mathbb{E}^{\mathcal{A}_n}\theta_n| \leq \mathbb{E}^{\mathcal{A}_n}|\theta_n|$ it suffices to prove this assuming $\theta_n \geq 0$. In that case, consider $f \geq 0$ in $L_{p'}$ with $\|f\|_{p'} = 1$ such that $\left\|\sum \mathbb{E}^{\mathcal{A}_n}\theta_n\right\|_p = \left\langle \sum \mathbb{E}^{\mathcal{A}_n}\theta_n, f\right\rangle$. Then

$$\begin{aligned}\left\langle \sum \mathbb{E}^{\mathcal{A}_n}\theta_n, f\right\rangle &= \sum \langle \theta_n, \mathbb{E}^{\mathcal{A}_n} f\rangle \\ &\leq \left\|\sum \theta_n\right\|_p \|\sup_n \mathbb{E}^{\mathcal{A}_n} f\|_{p'}\end{aligned}$$

hence by Doob's inequality

$$\leq p \left\|\sum \theta_n\right\|_p. \qquad \square$$

Remark 1.11. Note that it is crucial for the validity of Theorems 1.9 and 1.10 that the conditional expectations be *totally ordered*, as in a filtration. However, as we will now see, in some cases we can go beyond that. Let $(\mathcal{A}_n^1)_{n\geq 0}$, $(\mathcal{A}_n^2)_{n\geq 0}, \ldots, (\mathcal{A}_n^d)_{n\geq 0}$ be a d-tuple of (a priori mutually unrelated) filtrations on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let $I_d = \mathbb{N}^d$ and for all $i = (n(1), \ldots, n(d))$ let

$$\mathbb{E}_i = \mathbb{E}^{\mathcal{A}^1_{n(1)}}\mathbb{E}^{\mathcal{A}^2_{n(2)}} \cdots \mathbb{E}^{\mathcal{A}^d_{n(d)}}. \tag{1.16}$$

Then by a simple iteration argument, we find that for any $1 < p \leq \infty$ and any x in L_p we have

$$\|\sup_{i\in I_d} |\mathbb{E}_i x|\ \|_p \leq (p')^d \|x\|_p.$$

A similar iteration holds for the dual to Doob's inequality: for any family $(x_i)_{i\in I_d}$ in $L_{p'}$ we have

$$\left\|\sum |\mathbb{E}_i x_i|\right\|_{p'} \leq (p')^d \left\|\sum |x_i|\right\|_{p'}.$$

To illustrate this (following [91]), consider a dyadic rooted tree T, i.e. the points of T are finite sequences $\xi = (\xi_1, \ldots, \xi_k)$ with $\xi_j \in \{0, 1\}$ and there is also a root (or origin) denoted by ξ_ϕ. We introduce a partial order on T in the natural way, i.e. ξ_ϕ is $\le$ any element and then we set $(\xi_1, \ldots, \xi_k) \le (\xi'_1, \ldots, \xi'_j)$ if $k \le j$ and $(\xi_1, \ldots, \xi_k) = (\xi'_1, \ldots, \xi'_k)$. In other words, $\xi \le \xi'$ if ξ' is on the same "branch" as ξ but "after" ξ.

This is clearly *not* totally ordered since two points situated on disjoint branches are incomparable. Nevertheless, as observed in [91], we have the following: Consider a family $\{\varepsilon_\xi \mid \xi \in T\}$ of independent random variables and for any ξ in T let $\mathcal{A}_\xi = \sigma(\{\varepsilon_\eta \mid \eta \le \xi\}$, and let

$$\mathbb{E}_\xi = \mathbb{E}^{\mathcal{A}_\xi}.$$

We have then for any $1 < p \le \infty$ and any x in L_p

$$\| \sup_{\xi \in T} |\mathbb{E}_\xi x| \, \|_p \le (p')^3 \|x\|_p.$$

The idea is that $\mathbb{E}_\xi$ is actually of the form (1.16) with $d = 3$, see [91] for full details.

Remark 1.12. Let B be a Banach space and let $(M_n)_{n \ge 0}$ be a B-valued martingale. Then the random variables Z_n defined by $Z_n(\omega) = \|M_n(\omega)\|_B$ form a submartingale. Indeed, by (1.6) we have for every k and every f in $L_1(\Omega, \mathbb{P}; B)$

$$\|\mathbb{E}^{\mathcal{A}_k}(f)\| \overset{\text{a.s.}}{\le} \mathbb{E}^{\mathcal{A}_k}(\|f\|_B) \tag{1.17}$$

hence taking $f = M_n$ with $k \le n$ we obtain

$$\|M_k\| \le \mathbb{E}^{\mathcal{A}_k}(\|M_n\|)$$

which shows that (Z_n) is a submartingale. In particular, by (1.13) we have for any A in $\mathcal{A}_k$

$$\mathbb{E}(1_A \|M_k\|) \le \mathbb{E}(1_A \|M_n\|). \tag{1.18}$$

As a consequence, we can apply Doob's inequality to the submartingale (Z_n) and we obtain

Corollary 1.13. *Let (M_n) be a martingale with values in an arbitrary Banach space B. Then*

$$\sup_{t>0} t \mathbb{P}\{\sup_{n \ge 0} \|M_n\| > t\} \le \sup_{n \ge 0} \|M_n\|_{L_1(B)} \tag{1.19}$$

and for all $1 < p < \infty$

$$\| \sup_{n \ge 0} \|M_n\| \|_p \le p' \sup_{n \ge 0} \|M_n\|_{L_p(B)}. \tag{1.20}$$

We can now prove the martingale convergence theorem.

Theorem 1.14. *Let $1 \le p < \infty$. Let B be an arbitrary Banach space. Consider f in $L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$ and let $M_n = \mathbb{E}^{\mathcal{A}_n}(f)$ be the associated martingale. Then M_n converges a.s. to $\mathbb{E}^{\mathcal{A}_\infty}(f)$. Therefore, if a martingale (M_n) is convergent in $L_p(\Omega, \mathbb{P}; B)$ to a limit M_∞, then it necessarily converges a.s. to this limit, and we have $M_n = \mathbb{E}^{\mathcal{A}_n} M_\infty$ for all $n \ge 0$.*

Proof. The proof is based on a general principle, going back to Banach, that allows us to deduce almost sure convergence results from suitable maximal inequalities. By Theorem 1.5, we know that $\mathbb{E}^{\mathcal{A}_n}(f)$ converges in $L_p(B)$ to $M_\infty = \mathbb{E}^{\mathcal{A}_\infty}(f)$. Fix $\varepsilon > 0$ and choose k so that $\sup_{n \ge k} \|M_n - M_k\|_{L_p(B)} < \varepsilon$. We will apply (1.19) and (1.20) to the martingale $(M'_n)_{n \ge 0}$ defined by

$$M'_n = M_n - M_k \text{ if } n \ge k \quad \text{and} \quad M'_n = 0 \text{ if } n \le k.$$

We have in the case $1 < p < \infty$

$$\| \sup_{n \ge k} \|M_n - M_k\| \|_p \le p' \varepsilon$$

and in the case $p = 1$

$$\sup_{t > 0} t \mathbb{P}\{\sup_{n \ge k} \|M_n - M_k\| > t\} \le \varepsilon.$$

Therefore if we define pointwise

$$\ell = \lim_{k \to \infty} \sup_{n, m \ge k} \|M_n - M_m\|$$

we have

$$\ell = \inf_{k \ge 0} \sup_{n, m \ge k} \|M_n - M_m\| \le 2 \sup_{n \ge k} \|M_n - M_k\|.$$

Hence we find $\|\ell\|_p \le 2p'\varepsilon$ and

$$\sup_{t > 0} t \mathbb{P}\{\ell > 2t\} \le \varepsilon,$$

which implies (since $\varepsilon > 0$ is arbitrary) that $\ell = 0$ a.s., and hence by the Cauchy criterion that (M_n) converges a.s. Since $M_n \to M_\infty$ in $L_p(B)$ we have necessarily $M_n \to M_\infty$ a.s. Note that if a martingale M_n tends to a limit M_∞ in $L_p(B)$ then necessarily $M_n = \mathbb{E}^{\mathcal{A}_n}(M_\infty)$. Indeed, $M_n = \mathbb{E}^{\mathcal{A}_n} M_m$ for all $m \ge n$ and by continuity of $\mathbb{E}^{\mathcal{A}_n}$ we have $\mathbb{E}^{\mathcal{A}_n} M_m \to \mathbb{E}^{\mathcal{A}_n} M_\infty$ in $L_p(\mathcal{B})$ so that $M_n = \mathbb{E}^{\mathcal{A}_n} M_\infty$ as announced. This settles the last assertion. □

Corollary 1.15. *Every scalar valued martingale $(M_n)_{n \ge 0}$ which is bounded in L_p for some $p > 1$ (resp. uniformly integrable) must converge a.s. and in L_p (resp. L_1).*

Proof. By Corollary 1.6, if $(M_n)_{n \ge 0}$ is bounded in L_p for some $p > 1$ (resp. uniformly integrable) then M_n converges in L_p (resp. L_1) and by Theorem 1.14 the a.s. convergence is then automatic. □

Let B be a Banach space and let $(M_n)_{n\geq 0}$ be a sequence in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$. We will say that (M_n) is uniformly integrable if the sequence of positive r.v.'s $(\|M_n(\cdot)\|)_{n\geq 0}$ is uniformly integrable. More precisely, this means that $(\|M_n\|)$ is bounded in L_1 and that for any $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\forall A \in \mathcal{A} \qquad \mathbb{P}(A) < \delta \Rightarrow \sup_{n\geq 0} \int_A \|M_n\| < \varepsilon.$$

The following useful lemma illustrates the use of stopping times as a way to properly "truncate" a martingale.

Lemma 1.16. *Let $(M_n)_{n\geq 0}$ be a martingale bounded in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$ where B is an arbitrary Banach space. Fix $t > 0$ and let*

$$T = \begin{cases} \inf\{n \geq 0 \mid \|M_n\| > t\} & \textit{if } \sup \|M_n\| > t, \\ \infty & \textit{otherwise.} \end{cases}$$

Then

$$\mathbb{E}(\|M_T\| 1_{\{T<\infty\}}) \leq \sup_{n\geq 0} \mathbb{E}\|M_n\| \tag{1.21}$$

and moreover the martingale $(M_{n\wedge T})_{n\geq 0}$ is uniformly integrable.

Proof. First we claim that for any $0 \leq k \leq n$ we have

$$\mathbb{E}(1_{\{T=k\}}\|M_k\|) \leq \mathbb{E}(1_{\{T=k\}}\|M_n\|).$$

Indeed $\{T = k\} \in \mathcal{A}_k$ so this is a particular case of (1.18). Summing this with respect to $k \leq n$ we obtain

$$\mathbb{E}(1_{\{T\leq n\}}\|M_T\|) \leq \mathbb{E}(1_{\{T\leq n\}}\|M_n\|),$$

and taking the supremum over $n \geq 0$ we obtain (1.21).
Now recall that by definition $\sup \|M_n\| \leq t$ on $\{T = \infty\}$. More generally, we have $\sup_{n<T} \|M_n\| \leq t$, so that

$$\sup_n \|M_{n\wedge T}\| \leq \max\{1_{\{T<\infty\}}\|M_T\|, t\} \leq 1_{\{T<\infty\}}\|M_T\| + t.$$

Then we can write for any A in $\mathcal{A}$

$$\sup_n \mathbb{E}(1_A\|M_{n\wedge T}\|) \leq \mathbb{E}(1_A Z)$$

where $Z = 1_{\{T<\infty\}}\|M_T\| + t$. Thus we conclude that $(\|M_{n\wedge T}\|)_{n\geq 0}$ is uniformly integrable (since the single variable Z is itself uniformly integrable). □

To obtain what remains of Corollary 1.15 in the case $p = 1$, we will use the following simple fact.

Proposition 1.17. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let B be a Banach space and let $(\mathcal{A}_n)_{n\geq 0}$ be an increasing sequence of σ-subalgebras of $\mathcal{A}$. The following are equivalent:*

(i) Every B-valued martingale adapted to $(\mathcal{A}_n)_{n\geq 0}$ and bounded in $L_1(\Omega, \mathbb{P}; B)$ is a.s. convergent.

(ii) Every B-valued uniformly integrable martingale adapted to $(\mathcal{A}_n)_{n\geq 0}$ is a.s. convergent.

Proof. Assume (ii). Let (M_n) be a martingale bounded in $L_1(B)$. Fix $t > 0$ and consider $(M_{n\wedge T})$ as in Lemma 1.16. Since $(M_{n\wedge T})$ is uniformly integrable, it converges a.s. by (ii). This implies that if $\{T(\omega) = \infty\}$ then $(M_n(\omega))_{n\geq 0}$ is a.s. convergent. But by Doob's inequalities

$$\mathbb{P}\{T < \infty\} = \mathbb{P}\{\sup \|M_n\| > t\} \leq \frac{C}{t}$$

where $C = \sup \mathbb{E}\|M_n\|$. Therefore this probability can be made arbitrarily small by choosing t large, so that we conclude that the martingale $(M_n)_{n\geq 0}$ itself converges a.s. This shows that (ii) $\Rightarrow$ (i). The converse is trivial. □

Finally, we can state what is usually referred to as the "martingale convergence theorem".

Theorem 1.18. *Every L_1-bounded scalar valued martingale converges a.s.*

Proof. By Corollary 1.6, every scalar valued uniformly integrable martingale converges in L_1, and hence by Theorem 1.14 it converges a.s. Thus the present statement follows from the implication (ii) $\Rightarrow$ (i) from Proposition 1.17. □

We will also need the following

Theorem 1.19. *Every submartingale (M_n) bounded in L_1 (resp. and uniformly integrable) converges a.s. (resp. and in L_1.)*

Proof. We use the so-called Doob decomposition: we will write our submartingale as the sum of a martingale $(\widetilde{M}_n)_{n\geq 0}$ and a predictable increasing sequence (A_n) (recall that this means that A_n is $\mathcal{A}_{n-1}$ measurable for each $n \geq 1$). Let us write $\Delta_0 = M_0$ and $\Delta_n = M_n - M_{n-1}$ if $n \geq 1$. Let $d_n = \Delta_n - \mathbb{E}^{\mathcal{A}_{n-1}}(\Delta_n)$ if $n \geq 1$ and $d_0 = \Delta_0$, and let $\widetilde{M}_n = \sum_{k\leq n} d_k$. Then $(\widetilde{M}_n)_{n\geq 0}$ is a martingale. Indeed, by construction we have $\mathbb{E}^{\mathcal{A}_{n-1}}(d_n) = 0$ or equivalently $\mathbb{E}^{\mathcal{A}_{n-1}}\widetilde{M}_n = \widetilde{M}_{n-1}$. To relate $(\widetilde{M}_n)$ to (M_n), we note that

$$M_n = \sum_{0\leq k\leq n} \Delta_k = \sum_{0\leq k\leq n} d_k + \sum_{1\leq k\leq n} \mathbb{E}^{\mathcal{A}_{k-1}}(\Delta_k)$$

hence

$$M_n = \widetilde{M}_n + A_n$$

where

$$A_n = \sum_{1 \le k \le n} \mathbb{E}^{\mathcal{A}_{k-1}}(\Delta_k).$$

Moreover, by the submartingale property $\mathbb{E}^{\mathcal{A}_{n-1}}(\Delta_n) \ge 0$ for all $n \ge 1$ so that

$$0 \le A_1 \le A_2 \le \cdots \le A_{n-1} \le A_n \le \cdots .$$

On one hand, $\mathbb{E}A_n = \sum_{1 \le k \le n} \mathbb{E}\Delta_k = \mathbb{E}M_n - \mathbb{E}M_0$, and since (M_n) is assumed bounded in L_1 we have $\sup_{n \ge 1} \mathbb{E}A_n < \infty$. Therefore by monotonicity A_n converges a.s. and in L_1 when $n \to \infty$ (in particular it is a uniformly integrable sequence). On the other hand, we have

$$\mathbb{E}|\widetilde{M}_n| = \mathbb{E}|M_n - A_n| \le \mathbb{E}|M_n| + \mathbb{E}A_n$$

therefore $(\widetilde{M}_n)$ also is bounded in L_1 and is uniformly integrable if (M_n) is. By the martingale convergence theorem (Theorem 1.18) $(\widetilde{M}_n)$ converges a.s. hence $M_n = \widetilde{M}_n + A_n$ also converges a.s. and, in the uniformly integrable case, it also converges in L_1.

If we impose the initial condition $A_0 = 0$, the above proof also shows uniqueness: Indeed, $M_n = \widetilde{M}_n + A_n$ implies $A_n - A_{n-1} = \Delta_n - d\widetilde{M}_n$ and (assuming A_n $n-1$-measurable) this imposes $A_n - A_{n-1} = \mathbb{E}^{\mathcal{A}_{n-1}}(\Delta_n - d\widetilde{M}_n) = \mathbb{E}^{\mathcal{A}_{n-1}}(\Delta_n)$ which uniquely determines A_n if set $A_0 = 0$.

□

Corollary 1.20. *Let B be an arbitrary Banach space and let $(M_n)_{n \ge 0}$ be a B-valued martingale bounded in $L_1(B)$. Then $\|M_n\|_B$ converges a.s. Moreover, $(M_n)_{n \ge 0}$ converges a.s. in norm iff $\{M_n(\omega) \mid n \ge 0\}$ is relatively compact for almost all ω.*

Proof. The first assertion follows from Theorem 1.19 and Remark 1.12. It suffices to prove the second one for a separable B. Assume that $\{M_n(\omega) \mid n \ge 0\}$ is ω-a.s. relatively compact. Let $f(\omega)$ be a cluster point in B of $\{M_n(\omega) \mid n \ge 0\}$. Note that by Theorem 1.18 for any ξ in B^*, $\xi(M_n(\omega))$ converges ω-a.s., and hence it must converge to $\xi(f(\omega))$. (Incidentally: this shows that f is scalarly measurable, and hence by Appendix 2 is Bochner measurable). Let $D \subset B^*$ be a countable weak-$*$ dense subset. Clearly, $M_n(\omega)$ tends ω-a.s. to $f(\omega)$ in the $\sigma(B, D)$-topology, but if $\{M_n(\omega) \mid n \ge 0\}$ is relatively compact, the latter topology coincides on it with the norm topology, and hence $M_n(\omega) \to f(\omega)$ in norm. Conversely, if $\{M_n(\omega) \mid n \ge 0\}$ is convergent it is obviously relatively compact. □

Remark 1.21. The maximal inequalities for B-valued martingales can be considerably strengthened when $B = \ell_r$ for some $1 < r < \infty$: Consider a filtration $(\mathcal{A}_n)$ as usual, $f \in L_p(\ell_r)$ and let (f_n) be the martingale associated to f. Let (e_k) be the canonical basis of ℓ_r. We may develop f and f_n as

$f = \sum_k f(k)e_k$ and $f_n = \sum_k f_n(k)e_k$. In accordance with previous notation, we set $f(k)^* = \sup_n |f_n(k)|$. Let then

$$f^{**} = \| \sum f(k)^* e_k \|_{\ell_r} = (\sum f(k)^{*r})^{1/r}.$$

Then, for any $1 < p < \infty$, there is a constant $c(p,r)$ such that

$$\|f^{**}\|_p \le c(p,r)\|f\|_{L_p(\ell_r)} = c(p,r)\|(\sum |f(k)|^r)^{1/r}\|_p.$$

Note that $p = r$ is an easy consequence of Doob's inequality. See [56] for $p \le r$ and [187] for the general case and for a weak type-$(1,1)$ inequality that can be proved using the Gundy decomposition described in the next chapter. Finally, the extension to the case $B = L_r$ requires only minor modifications.

There are cases where the maximal inequalities can be extended to L_p with $0 < p < 1$. For instance, let $(Y_n)_{n\ge 0}$ be a sequence of independent B-valued random variables, let $f_n = \sum_0^n Y_n$. If $Y_n \in L_1(B)$ is symmetric for all n (this implies $\mathbb{E}Y_n = 0$), then $(f_n)_{n\ge 0}$ is a martingale satisfying $\mathbb{P}(\sup\limits_n \|f_n\| > t) \le 2\sup\limits_n \mathbb{P}(\|f_n\| > t)$. More generally, we quote without proof the following:

Theorem 1.22. *Let (Y_n) be a sequence of B-valued random variables, such that, for any choice of signs $\xi_n = \pm 1$, the sequence $(\xi_n Y_n)$ has the same distribution as (Y_n). Let $f_n = \sum_0^n Y_k$. We have then:*

$$\forall t > 0 \qquad \mathbb{P}(\sup \|f_n\| > t) \le 2\limsup \mathbb{P}(\|f_n\| > t) \tag{1.22}$$

$$\forall p > 0 \qquad \mathbb{E}\sup \|f_n\|^p \le 2\limsup \mathbb{E}\|f_n\|^p. \tag{1.23}$$

If f_n converges to a limit f_∞ in probability (i.e. $\|f_n - f_\infty\| \to 0$ in probability), then it actually converges a.s. In particular, if f_n converges in L_p $(p > 0)$, then it automatically converges a.s. Finally, if f_n converges a.s. to a limit f_∞, we have

$$\forall t > 0 \qquad \mathbb{P}(\sup \|f_n\| > t) \le 2\mathbb{P}(\|f_\infty\| > t). \tag{1.24}$$

More generally for any Borel convex subset $K \subset B$, we have

$$\mathbb{P}(\cup_n \{f_n \notin K\}) \le 2\mathbb{P}(\{f_\infty \notin K\}).$$

Corollary 1.23. *Let (Y_n) be independent variables in $L_1(B)$ with mean zero (i.e. $\mathbb{E}Y_n = 0$) and let $f_n = \sum_0^n Y_k$ as before. Then, for any $p \ge 1$, we have*

$$\|\sup \|f_n\|\|_p \le 2^{1+1/p} \sup \|f_n\|_{L_p(B)}.$$

Proof. Let $(Y'_n)_n$ be an independent copy of the sequence (Y_n), let $\widetilde{Y}_n = Y_n - Y'_n$ and $f'_n = \sum_0^n Y'_n$. Note that $(\widetilde{Y}_n)$ are independent and symmetric. By (1.23) we have

$$\mathbb{E}\sup \|\widetilde{f}_n\|^p \le 2\sup \mathbb{E}\|\widetilde{f}_n\|^p$$

but now if $p \geq 1$ we have by convexity

$$\begin{aligned}\mathbb{E}\sup\|f_n\|^p &= \mathbb{E}\sup\|f_n - \mathbb{E}f_n'\|^p \leq \mathbb{E}\sup\|f_n - f_n'\|^p \\ &\leq 2\sup\mathbb{E}\|f_n - f_n'\|^p \\ &\leq 2\sup\mathbb{E}(\|f_n\| + \|f_n'\|)^p \\ &\leq 2^p(\mathbb{E}\sup\|f_n\|^p + \mathbb{E}\sup\|f_n'\|^p) = 2^{p+1}\mathbb{E}\sup\|f_n\|^p.\end{aligned}$$

□

Corollary 1.24. *For a series of independent B-valued random variables, convergence in probability implies almost sure convergence.*

Proof. Let $f_n = \sum_0^n Y_k$, with (Y_k) independent. Let (Y_k') be an independent copy of the sequence (Y_k) and let $f_n' = \sum_0^n Y_k'$. Then the variables $(Y_k - Y_k')$ are independent and symmetric. If f_n converges in probability (when $n \to \infty$), then obviously f_n' and hence $f_n - f_n'$ also does. By the preceding Theorem, $f_n - f_n'$ converges a.s., therefore we can choose fixed values $x_n = f_n'(\omega_0)$ such that $f_n - x_n$ converges a.s.. A fortiori, $f_n - x_n$ converges in probability, and since f_n also does, the difference $f_n - (f_n - x_n) = x_n$ also does, which means that (x_n) is convergent in B. Thus the a.s. convergence of $f_n - x_n$ implies that of f_n. □

Remark 1.25. There are well known counterexamples showing that Theorem 1.18 does not extend to the Banach space valued case. For instance, let $\Omega = \{-1,1\}^{\mathbb{N}}$ equipped with the usual probability measure $\mathbb{P}$ and let $\mathcal{A}_n$ be the σ-algebra generated by the $(n+1)$ first coordinates denoted by $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n$. A classical example of a real valued martingale is $M_n = \prod\limits_{k\leq n}(1+\varepsilon_k)$, which is positive and of integral 1. Note however that it does not converge in L_1. Another example is $M_n = \sum\limits_{k\leq n}\alpha_k\varepsilon_k$ where (α_k) are real coefficients. This particular martingale is bounded in L_1 iff $\Sigma|\alpha_n|^2$ is finite. By the martingale convergence theorem, these two martingales must converge a.s. However, we can give very similar Banach space valued examples which do not converge. Take for instance $B = c_0$ and let (e_n) be the canonical basis of c_0. Let $M_n^1 = \sum\limits_{k\leq n}\varepsilon_k e_k$. Then $\|M_n^1(\omega)\|_{c_0} = \sup\limits_{k\leq n}|\varepsilon_k(\omega)| \equiv 1$ but clearly there is no point ω in $\{-1,1\}^{\mathbb{N}}$ such that the sequence $(M_n^1(\omega))_{n\geq 0}$ is convergent in c_0, since we have

$$\forall\omega\in\Omega\quad\forall k<n\qquad\|M_n^1(\omega) - M_k^1(\omega)\|_B = 1.$$

We can give a similar example in L_1. Let $B = L^1(\Omega,\mathbb{P})$ itself and let

$$M_n^2(\omega) = \prod_{k\leq n}(1+\varepsilon_k(\omega)\varepsilon_k). \tag{1.25}$$

Then again $\|M_n^2(\omega)\|_B = 1$ for all ω, but also it is easy to check that

$$\forall\omega\in\Omega\quad\forall k<n\qquad\|M_n^2(\omega) - M_k^2(\omega)\|_B \geq 1, \text{ and } \|M_n^2(\omega) - M_{n-1}^2(\omega)\|_B = 1,$$

so that $(M_n^2)_{n\geq 0}$ is nowhere convergent.

In the next chapter, we will show that the preceding examples cannot occur in a Banach space with the RNP.

1.5 Reverse martingales

We will prove here the following

Theorem. *Let B be an arbitrary Banach space. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $\mathcal{A}_0 \supset \mathcal{A}_{-1} \supset \mathcal{A}_{-2} \supset \cdots$ be a (this time decreasing) sequence of σ-subalgebras of $\mathcal{A}$. Let $\mathcal{A}_{-\infty} = \bigcap\limits_{n\geq 0} \mathcal{A}_{-n}$. Then for any f in $L_p(\Omega, \mathcal{A}, P; B)$, with $1 \leq p < \infty$, the reverse martingale $(\mathbb{E}^{\mathcal{A}_{-n}}(f))_{n\geq 0}$ converges to $\mathbb{E}^{\mathcal{A}_{-\infty}}(f)$ a.s. and in $L_p(B)$.*

We first check the convergence in $L_p(B)$. Since the operators $(\mathbb{E}^{\mathcal{A}_{-n}})_{n\geq 0}$ are equicontinuous on $L_p(B)$ it suffices to check this for f in a dense subset of $L_p(B)$. In particular, it suffices to consider f of the form $f = \sum\limits_1^n \varphi_i x_i$ with φ_i an indicator function and x_i in B. Since $\varphi_i \in L_2(\Omega, \mathbb{P})$, we have (by classical Hilbert space theory) $\mathbb{E}^{\mathcal{A}_{-n}}\varphi_i \to \mathbb{E}^{\mathcal{A}_{-\infty}}\varphi_i$ in $L_2(\Omega, \mathbb{P})$ when $n \to \infty$. (Note that $L_2(\Omega, \mathcal{A}_{-\infty}, \mathbb{P})$ is the intersection of the family $(L_2(\Omega, \mathcal{A}_{-n}, \mathbb{P}))_{n\geq 0}$.)
Observe that $\|f - g\|_p \leq \|f - g\|_2$ if $p \leq 2$ and $\|f - g\|_p^p \leq 2^{p-2}\|f - g\|_2^2$ if $\|f\|_\infty \leq 1$, $\|g\|_\infty \leq 1$ and $p > 2$. Using this, we obtain that, a fortiori, $\mathbb{E}^{\mathcal{A}_{-n}} f \to \mathbb{E}^{\mathcal{A}_{-\infty}} f$ in $L_p(B)$ for every f of the above form, and hence for every f in $L_p(B)$.

We now turn to a.s. convergence. We first replace f by $\tilde{f} = f - \mathbb{E}^{\mathcal{A}_{-\infty}}(f)$ so that we can assume $\mathbb{E}^{\mathcal{A}_{-n}}(f) \to 0$ in $L_p(B)$ and a fortiori in $L_1(B)$. Let $f_n = \mathbb{E}^{\mathcal{A}_{-n}} f$. Now fix $n > 0$ and $k > 0$ and consider the (ordinary sense) martingale

$$M_j = \begin{cases} f_{-n-k+j} & \text{for} \quad j = 0, 1, \ldots, k, \\ f_{-n} & \text{if} \quad j \geq k. \end{cases}$$

Then by Doob's inequality (1.19) applied to (M_j) we have for all $t > 0$

$$t\mathbb{P}\{\sup_{n\leq m\leq n+k} \|f_{-m}\| > t\} \leq \mathbb{E}\|f_{-n}\|$$

therefore

$$t\mathbb{P}\{\sup_{m\geq n} \|f_{-m}\| > t\} \leq \mathbb{E}\|f_{-n}\|$$

and since $\mathbb{E}\|f_{-n}\| \to 0$ when $n \to \infty$, we have $\sup\limits_{m\geq n} \|f_{-m}\| \to 0$ a.s., or equivalently $f_{-n} \to 0$ a.s. when $n \to \infty$. □

As a corollary, we have the following classical application to the strong law of large numbers.

Corollary. *Let $\varphi_1, \ldots, \varphi_n$ be a sequence of independent, identically distributed random variables in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$. Let $S_n = \varphi_1 + \cdots + \varphi_n$. Then $\frac{S_n}{n} \to \mathbb{E}\varphi_1$ a.s. and in $L_1(B)$.*

Proof. Let $\mathcal{A}_{-n}$ be the σ-algebra generated by $(S_n, S_{n+1}, \ldots)$. We claim that $\frac{1}{n}S_n = \mathbb{E}^{\mathcal{A}_{-n}}(\varphi_1)$. Indeed, for every $k \leq n$, since the exchange of φ_1 and φ_k preserves $S_n, S_{n+1}, \ldots$, we have

$$\mathbb{E}^{\mathcal{A}_{-n}}(\varphi_k) = \mathbb{E}^{\mathcal{A}_{-n}}(\varphi_1).$$

Therefore averaging the preceding equality over $k \leq n$ we obtain

$$\mathbb{E}^{\mathcal{A}_{-n}}(\varphi_1) = \frac{1}{n} \sum_{1 \leq k \leq n} \mathbb{E}^{\mathcal{A}_{-n}}(\varphi_k) = \mathbb{E}^{\mathcal{A}_{-n}}\left(\frac{S_n}{n}\right) = \frac{S_n}{n}.$$

Hence $(S_n/n)_{n\geq 1}$ is a reverse martingale satisfying the assumptions of the preceding theorem (we may take say $\mathcal{A}_0 = \mathcal{A}_{-1}$), therefore $\frac{1}{n}S_n \to \mathbb{E}^{\mathcal{A}_{-\infty}}(\varphi_1)$ a.s. and in $L_1(B)$. Finally, let $\mathcal{T} = \bigcap_{n\geq 0} \sigma\{\varphi_n, \varphi_{n+1}, \cdots\}$ be the tail σ-algebra. By the zero-one law, $\mathcal{T}$ is trivial. The limit of S_n/n is clearly $\mathcal{T}$-measurable, hence it must be equal to a constant c, but then $\mathbb{E}(S_n/n) \to c$, so $c = \mathbb{E}(\varphi_1)$. □

1.6 Notes and Remarks

Among the many classical books on Probability that influenced us, we mention [17, 9], see also [26]. As for martingales, the references that considerably influenced us are [48, 25, 20] and the papers [101, 108].

Martingales were considered long before Doob (in particular by Paul Lévy) but he is the one who invented the name and proved their basic almost sure convergence properties using what is now called Doob's maximal inequality.

We give more references in the Appendix relative to continuous time.

In Theorem 1.22, we slightly digress and concentrate on a particular sort of martingale, those that are partial sums of series of independent random vectors. In the symmetric case, it turns out that the maximal inequalities (and the associated almost sure convergence) hold for "martingales" bounded in $L_p(B)$ for $p < 1$. Our presentation of this is inspired by Kahane's book [31].

Chapter 2

Radon Nikodym property

2.1 Martingales, dentability and Radon Nikodym property

To introduce the Radon Nikodym property (in short RNP), we will need to briefly review the basic theory of vector measures. Let B be a Banach space. Let $(\Omega, \mathcal{A})$ be a measure space. Every σ-additive map $\mu\colon\ \mathcal{A} \to B$ will be called a (B-valued) vector measure. We will say that μ is *bounded* if there is a finite positive measure ν on $(\Omega, \mathcal{A})$ such that

$$\forall A \in \mathcal{A} \qquad \|\mu(A)\| \le \nu(A). \tag{2.1}$$

When this holds, it is easy to show that there is a minimal choice of the measure ν. Indeed, for all A in $\mathcal{A}$ let

$$|\mu|(A) = \sup\{\Sigma\|\mu(A_i)\|\}$$

where the supremum runs over all decompositions of A as a disjoint union $A = \cup A_i$ of finitely many sets in $\mathcal{A}$. Using the triangle inequality, one checks that $|\mu|$ is an additive set function, by (2.1) $|\mu|$ must be σ-additive and finite. Clearly, when (2.1) holds, we have

$$|\mu| \le \nu.$$

We define the "total variation norm" of μ as follows

$$\|\mu\| = \inf\{\nu(\Omega) \mid \nu \in M(\Omega, \mathcal{A}), \nu \ge |\mu|\},$$

or equivalently

$$\|\mu\| = |\mu|(\Omega).$$

We will denote by $M(\Omega, \mathcal{A})$ the Banach space of all bounded complex valued measures on $(\Omega, \mathcal{A})$, and by $M_+(\Omega, \mathcal{A})$ the subset of all positive bounded measures. We will denote by $M(\Omega, \mathcal{A}; B)$ the space of all bounded B-valued measures μ on $(\Omega, \mathcal{A})$. When equipped with the preceding norm, it is a Banach space. Let $\mu \in M(\Omega, \mathcal{A}; B)$ and $\nu \in M_+(\Omega, \mathcal{A})$. We will write

$$|\mu| \ll \nu$$

if $|\mu|$ is absolutely continuous (or equivalently admits a density) with respect to ν. This happens iff there is a positive function $w \in L_1(\Omega, \mathcal{A}, \nu)$ such that

$$|\mu| \leq w.\nu$$

or equivalently such that

$$\forall A \in \mathcal{A} \qquad \|\mu(A)\| \leq \int_A w d\nu.$$

Recapitulating, we may state:

Proposition 2.1. *A vector measure μ is bounded in the above sense iff its total variation is finite, the total variation being defined as*

$$V(\mu) = \sup \left(\sum_1^n \|\mu(A_i)\| \right)$$

where the sup runs over all measurable partitions $\Omega = \bigcup_{i=1}^n A_i$ of Ω. Thus, if μ is bounded, we have $V(\mu) = |\mu|(\Omega)$.

Proof. Assuming $V(\mu) < \infty$, let $\forall A \in \mathcal{A} \quad \nu(A) = \sup\left(\sum_1^n \|\mu(A_i)\|\right)$, where the sup runs over all measurable partitions $A = \bigcup_{i=1}^n A_i$ of A. Then ν is a σ-additive finite positive measure on $\mathcal{A}$, and satisfies (2.1). Thus μ is bounded in the above sense (and of course ν is nothing but $|\mu|$). The converse is obvious. $\square$

Remark. It is easy to check that if $d\mu = f.d\nu$ with $f \in L_1(\Omega, \mathcal{A}, \nu; B)$, then

$$d|\mu| = \|f(.)\|_B d\nu, \tag{2.2}$$

and therefore

$$\|f.\nu\|_{M(\Omega,\mathcal{A};B)} = \|f\|_{L_1(\Omega,\mathcal{A},\nu;B)}. \tag{2.3}$$

Indeed, by Jensen's inequality we clearly have

$$\forall A \in \mathcal{A} \qquad \|\mu(A)\| \leq \int_A \|f\| d\nu,$$

hence $d|\mu| \leq \|f(.)\|_B d\nu$. To prove the converse, let $\epsilon > 0$ and let g be a B-valued simple function such that $\int_A \|f - g\| d\nu < \epsilon$. We can clearly assume that g is supported by A, so that we can write $g = \sum_1^n 1_{A_i} x_i$, with $x_i \in B$ and A_i is a disjoint partition of A. We have

$$\Sigma \|\mu(A_i) - \nu(A_i) x_i\| = \Sigma \| \int_{A_i} (f - g) d\nu \| \leq \int_A \|f - g\| d\nu < \epsilon$$

hence

$$\int_A \|g\| d\nu = \Sigma \nu(A_i) \|x_i\| \le \Sigma \|\mu(A_i)\| + \epsilon$$

and finally

$$\int_A \|f\| d\nu \le \int_A \|g\| d\nu + \epsilon \le \Sigma \|\mu(A_i)\| + 2\epsilon,$$

which implies

$$\int_A \|f\| d\nu \le |\mu|(A) + 2\epsilon.$$

This completes the proof of (2.2).

We will use very little from the theory of vector measures, for more details we refer the reader to [16].

Definition. A Banach space B is said to have the Radon Nikodym property (in short RNP) if for every measure space $(\Omega, \mathcal{A})$, for every finite positive measure ν on $(\Omega, \mathcal{A})$ and for every B-valued measure μ in $M(\Omega, \mathcal{A}; B)$ such that $|\mu| \ll \nu$, there is a function f in $L_1(\Omega, \mathcal{A}, \nu; B)$ such that $\mu = f.\nu$ i.e. such that

$$\forall A \in \mathcal{A} \qquad \mu(A) = \int_A f d\nu.$$

We will need the concept of a δ-separated tree.

Definitions. Let $\delta > 0$. A martingale $(M_n)_{n \ge 0}$ in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$ will be called δ-separated if

(i) M_0 is constant,

(ii) Each M_n takes only finitely many values,

(iii) $\forall n \ge 1$, $\forall \omega \in \Omega \quad \|M_n(\omega) - M_{n-1}(\omega)\| \ge \delta$.

Moreover, the set $S = \{M_n(\omega) \mid n \ge 0, \omega \in \Omega\}$ of all possible values of such a martingale will be called a δ-separated tree.

Another perhaps more intuitive description of a δ-separated tree is as a collection of points $\{x_i \mid i \in I\}$ indexed by the set of nodes of a tree-like structure which starts at some origin (0) then separates into N_1 branches which we denote by $(0,1), (0,2), \dots, (0, N_1)$, then each branch itself splits into a finite number of branches, etc. in such a way that each point x_i is a convex combination of its immediate successors, and all these successors are at distance at least δ from x_i. We will also need another more geometric notion.

Definition. Let B be a Banach space. A subset $D \subset B$ is called dentable if for any $\varepsilon > 0$ there is a point x in D such that

$$x \notin \overline{\text{conv}}(D \backslash B(x, \varepsilon))$$

where $\overline{\text{conv}}$ denotes the closure of the convex hull, and where

$$B(x, \varepsilon) = \{y \in B \mid \|y - x\| < \varepsilon\}.$$

Remark 2.2. Let $D \subset B$ be a bounded subset and let C be the closed convex hull of D. If C is dentable, then D is dentable. Moreover, C is dentable iff C admits slices of arbitrarily small diameter. Note in particular that the dentability of all closed bounded convex sets implies that of all bounded sets.

Indeed, the presence of slices of small diameter clearly implies dentability. Conversely, if C is dentable, then for any $\varepsilon > 0$ there is a point x in C that does not belong to the closed convex hull of $C \setminus B(x, \varepsilon)$, and hence by Hahn-Banach separation, there is a slice of C containing x and included in $B(x, \varepsilon)$, therefore with diameter less than 2ε. Now if $C = \overline{\text{conv}}(D)$, then this slice must contain a point in D, exhibiting that D itself is dentable.

The following beautiful theorem gives a geometric sufficient condition for the RNP. We will see shortly that it is also necessary.

Theorem 2.3. *If every bounded subset of a Banach space B is dentable, then B has the RNP.*

Proof. Let $(\Omega, \mathcal{A}, m)$ be a σ-finite measure space and let $\mu\colon\ \mathcal{A} \to B$ be a bounded vector measure such that $|\mu| \ll m$. We will show that μ admits a Radon Nikodym derivative in $L_1(\Omega, \mathcal{A}, m; B)$. Clearly (by replacing m by $|\mu|$) we may as well assume that m is finite and $|\mu| \leq m$. Indeed, let $m' = |\mu| = w.m$ for some w in $L_1(m)$, if we find f' such that $\mu = f'.m'$, we have by (2.2) $|\mu| = \|f'\|.m'$ hence $\|f'\| = 1$ a.s. and therefore if $f = wf'$ we have $\mu = f.m$ and $f \in L_1(m; B)$. Now assume $|\mu| \leq m$ and for every A in $\mathcal{A}$ let $x_A = \frac{\mu(A)}{m(A)}$ and let

$$C_A = \{x_\beta \mid \beta \in \mathcal{A},\ \beta \subset A,\ m(\beta) > 0\}.$$

Note that $\|x_A\| \leq 1$ for all A in $\mathcal{A}$, so that the sets C_A are bounded. We will show that if every set C_A is dentable then the measure admits a Radon Nikodym derivative f in $L_1(\Omega, \mathcal{A}, m; B)$.

Step 1: We first claim that if C_Ω is dentable then $\forall \varepsilon > 0\, \exists A \in \mathcal{A}$ with $m(A) > 0$ such that

$$\text{diam}(C_A) \leq 2\varepsilon.$$

This (as well as the third) step is proved by an exhaustion argument. Suppose that this does not hold, then $\exists \varepsilon > 0$ such that every A with $m(A) > 0$ satisfies $\text{diam}(C_A) > 2\varepsilon$. In particular, for any x in B, A contains a subset β with $m(\beta) > 0$ such that $\|x - x_\beta\| > \varepsilon$. Then, consider a fixed measurable A with $m(A) > 0$ and let (β_n) be a maximal collection of disjoint measurable subsets of A with positive measure such that $\|x_A - x_{\beta_n}\| > \varepsilon$. (Note that since $m(\beta_n) > 0$ and the sets are disjoint, such a maximal collection must be at most countable.) By our assumption, we must have $A = \bigcup \beta_n$, otherwise we could take $A' = A - \bigcup_n \beta_n$ and find a subset β of A' that would contradict the maximality of the family (β_n). But now if $A = \bigcup \beta_n$, we have

$$x_A = \Sigma(m(\beta_n)/m(A))x_{\beta_n} \quad \text{and} \quad \|x_A - x_{\beta_n}\| > \varepsilon.$$

Since we can do this for every $A \subset \Omega$ with $m(A) > 0$ this means that for some $\varepsilon > 0$, every point x of C_Ω lies in the closed convex hull of points in

$C_\Omega - B(x,\varepsilon)$, in other words this means that C_Ω is not dentable, which is the announced contradiction. This proves the above claim and completes step 1. Working with C_A instead of C_Ω, we immediately obtain
Step 2:

$$\forall \varepsilon > 0 \quad \forall A \in \mathcal{A} \qquad \text{with} \quad m(A) > 0$$

$\exists A' \subset A$ with $m(A') > 0$ such that

$$\mathrm{diam}(C_{A'}) \leq 2\varepsilon.$$

Step 3: We use a second exhaustion argument. Let $\varepsilon > 0$ be arbitrary and let (A_n) be a maximal collection of disjoint measurable subsets of Ω with $m(A_n) > 0$ such that $\mathrm{diam}(C_{A_n}) \leq 2\varepsilon$. We claim that, up to a negligible set, we have necessarily $\Omega = \bigcup A_n$. Indeed if not, we could take $A = \Omega - (\bigcup A_n)$ in step 2 and find $A' \subset A$ contradicting the maximality of the family (A_n). Thus $\Omega = \bigcup A_n$. Now let $g_\varepsilon = \Sigma 1_{A_n} x_{A_n}$. Clearly, $g_\varepsilon \in L_1(\Omega, m; B)$ and we have

$$\|\mu - g_\varepsilon.m\|_{M(\Omega,\mathcal{A};B)} \leq 2\varepsilon m(\Omega). \tag{2.4}$$

Indeed, for every A in $\mathcal{A}$ with $m(A) > 0$

$$\mu(A) - \int_A g_\varepsilon dm = \Sigma m(A \cap A_n)[x_{A\cap A_n} - x_{A_n}]$$

hence

$$\begin{aligned}\left\|\mu(A) - \int_A g_\varepsilon dm\right\| &\leq \Sigma m(A \cap A_n)\|x_{A\cap A_n} - x_{A_n}\| \\ &\leq m(A)(2\varepsilon),\end{aligned}$$

which implies (2.4).

This shows that μ belongs to the closure in $M(\Omega, \mathcal{A}, B)$ of the set of all measures of the form $f.m$ for some f in $L_1(\Omega, \mathcal{A}; B)$, and since this set is closed by (2.3) we conclude that μ itself is of this form. Perhaps, a more concrete way to say the same thing is to say that if $f_n = g_{2^{-n}}$ then $f = f_0 + \sum_{n\geq 1} f_n - f_{n-1}$ is in $L_1(\Omega, m; B)$ and we have $\mu = f.m$. (Indeed, note that (2.4) (with (2.2)) implies $\|f_n - f_{n-1}\|_{L_1(B)} \leq 6.2^{-n} m(\Omega)$.) □

To expand on Theorem 2.3, the following simple lemma will be useful.

Lemma 2.4. *Fix $\varepsilon > 0$. Let $D \subset B$ be a subset such that*

$$\forall x \in D \qquad x \in \overline{\mathrm{conv}}(D \backslash B(x,\varepsilon)) \tag{2.5}$$

then the enlarged subset $\widetilde{D} = D + B(0, \varepsilon/2)$ satisfies

$$\forall x \in \widetilde{D} \qquad x \in \mathrm{conv}(\widetilde{D} \backslash B(x, \varepsilon/2)). \tag{2.6}$$

Proof. Consider x in $\widetilde{D}$, $x = x' + y$ with $x' \in D$ and $\|y\| < \varepsilon/2$. Choose $\delta > 0$ small enough so that $\delta + \|y\| < \varepsilon/2$. By (2.5) there are positive numbers $\alpha_1, \dots, \alpha_n$ with $\Sigma\alpha_i = 1$ and $x_1, \dots, x_n \in D$ such that $\|x_i - x'\| \geq \varepsilon$ and $\|x' - \Sigma\alpha_i x_i\| < \delta$. Hence $x' = \Sigma\alpha_i x_i + z$ with $\|z\| < \delta$. We can write $x = x' + y = \Sigma\alpha_i(x_i + z + y)$. Note that $x_i + z + y \in \widetilde{D}$ since $\|z + y\| < \varepsilon/2$ and moreover

$$\begin{aligned}\|x - (x_i + z + y)\| = \|x' - x_i - z\| &\geq \|x' - x_i\| - \|z\| \\ &\geq \varepsilon - \delta \geq \varepsilon/2.\end{aligned}$$

Hence we conclude that (2.6) holds. □

We now come to a very important result which incorporates the converse to Theorem 2.3.

Theorem 2.5. *Fix $1 < p \leq \infty$. The following properties of a Banach space B are equivalent*

(i) B has the RNP.

(ii) Every uniformly integrable martingale in $L_1(B)$ converges a.s. and in $L_1(B)$.

(iii) Every B-valued martingale bounded in $L_1(B)$ converges a.s.

(iv) Every B-valued martingale bounded in $L_p(B)$ converges a.s.

(v) For every $\delta > 0$, B does not contain a bounded δ-separated tree.

(vi) Every bounded subset of B is dentable.

Proof. (i) $\Rightarrow$ (ii). Assume (i). Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $(\mathcal{A}_n)_{n\geq 0}$ be an increasing sequence of σ-subalgebras. Let us assume $\mathcal{A} = \mathcal{A}_\infty$ for simplicity. Let (M_n) be a B-valued uniformly integrable martingale adapted to $(\mathcal{A}_n)_{n\geq 0}$. We can associate to it a vector measure μ as follows. For any A in $\mathcal{A} = \mathcal{A}_\infty$, we define

$$\mu(A) = \lim_{n\to\infty} \int_A M_n d\mathbb{P}. \tag{2.7}$$

We will show that this indeed makes sense and defines a bounded vector measure. Note that if $A \in \mathcal{A}_k$ then by (1.7) for all $n \geq k$ $\int_A M_n d\mathbb{P} = \int_A M_k d\mathbb{P}$, so that the limit in (2.7) is actually stationary. Thus, (2.7) is well defined when $A \in \bigcup_{n\geq 0} \mathcal{A}_n$. Since (M_n) is uniformly integrable, $\forall\varepsilon > 0\,\exists\delta > 0$ such that $\mathbb{P}(A) < \delta \Rightarrow \|\mu(A)\| < \varepsilon$. Using this, it is easy to check that μ extends to a σ-additive vector measure on $\mathcal{A}_\infty$. Indeed, note that (for instance by scalarization) $\mathbb{E}(M_n 1_A) = \mathbb{E}(M_n \mathbb{E}^{\mathcal{A}_n}(1_A))$. Thus the limit in (2.7) is the same as

$$\lim_{n\to\infty} \mathbb{E}(M_n \mathbb{E}^{\mathcal{A}_n}(1_A)). \tag{2.8}$$

To check that this definition makes sense, note that if $\varphi_n = \mathbb{E}^{\mathcal{A}_n}(1_A)$, then

$$\forall n < m \qquad \mathbb{E}(M_n\varphi_n) - \mathbb{E}(M_m\varphi_m) = \mathbb{E}(M_m(\varphi_n - \varphi_m)) \tag{2.9}$$

but by the uniform integrability (since $|\varphi_n - \varphi_m| \le 2$) we also must have $\|\mathbb{E}(M_m(\varphi_n - \varphi_m))\| \to 0$ when $n, m \to \infty$. Indeed, we can write for any $t > 0$

$$\|\mathbb{E}(M_m(\varphi_n - \varphi_m))\| \le 2 \sup_m \int_{\|M_m\|>t} \|M_m\| + t\mathbb{E}|\varphi_n - \varphi_m|,$$

so that $\limsup_{n,m\to\infty} \|\mathbb{E}(M_m(\varphi_n - \varphi_m))\| \le 2\sup_m \int_{\|M_m\|>t} \|M_m\|$ and hence must vanish by the uniform integrability. Thus by (2.9) we conclude that the limit in (2.7) exists by the Cauchy criterion.

By Theorem 1.19, the submartingale $\|M_n\|$ converges in L_1 to a limit w in L_1. Note that for all A in $\mathcal{A}$

$$|\mu|(A) \le \int_A w d\mathbb{P}. \tag{2.10}$$

Indeed, by (2.7) and Jensen's inequality, we have

$$\|\mu(A)\| \le \lim_{n\to\infty} \mathbb{E}(\|M_n\|1_A) = \int_A w d\mathbb{P},$$

and hence also for all $A_1, \ldots, A_m$ in $\mathcal{A}$ disjoint with $A = \cup A_i$

$$\sum_1^m \|\mu(A_i)\| \le \sum_1^m \int_{A_i} w d\mathbb{P} = \int_A w d\mathbb{P}.$$

and taking the supremum of the left hand side, we obtain the above claim (2.10). This shows $|\mu| \ll \mathbb{P}$. By our assumption (i), there is f in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$ such that $\mu(A) = \int_A f d\mathbb{P}$ for all A in $\mathcal{A}$.

Recall that for any $k \ge 0$ and for any A in $\mathcal{A}_k$ we have by (1.7)

$$\forall n \ge k \qquad \mathbb{E}(M_n 1_A) = \mathbb{E}(M_k 1_A)$$

hence by (2.7) $\mu(A) = \mathbb{E}(M_k 1_A)$ for any A in $\mathcal{A}_k$. Therefore we must have

$$\forall k \ge 0 \quad \forall A \in \mathcal{A}_k \qquad \int_A f d\mathbb{P} = \int_A M_k d\mathbb{P}$$

or equivalently, since this property characterizes $\mathbb{E}^{\mathcal{A}_k}(f)$ (see the remarks after (1.5)) $M_k = \mathbb{E}^{\mathcal{A}_k}(f)$. Hence by Theorems 1.5 and 1.14, (M_n) converges to f a.s. and in $L_1(B)$. This completes the proof of (i) $\Rightarrow$ (ii).
(ii) $\Rightarrow$ (iii). This follows from Proposition 1.17.
(iii) $\Rightarrow$ (iv) is obvious. We give below a direct proof that (iv) implies (i).
(iv) $\Rightarrow$ (v) is clear, indeed a bounded δ-separated tree is the range of a uniformly bounded martingale (M_n) which converges nowhere since $\|M_n - M_{n-1}\| \ge \delta$ everywhere.

(vi) $\Rightarrow$ (i) is Theorem 2.3, so it only remains to prove (v) $\Rightarrow$ (vi).

Assume that (vi) fails. We will show that (v) must also fail. Let $D \subset B$ be a bounded nondentable subset. Replacing D by the set $\widetilde{D}$ in Lemma 2.4, we can assume that there is a number $\delta > 0$ such that

$$\forall x \in D \qquad x \in \operatorname{conv}(D - B(x, \delta)).$$

We will then construct a δ-separated tree inside D. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be the Lebesgue interval. We pick an arbitrary point x_0 in D and let $M_0 \equiv x_0$ on $\Omega = [0, 1]$. Then since $x_0 \in \operatorname{conv}(D - B(x_0, \delta))$

$$\exists \alpha_1 > 0, \dots, \alpha_n > 0 \quad \text{with} \quad \sum_1^n \alpha_i = 1 \quad \exists x_1, \dots, x_n \in D$$

such that

$$x_0 = \sum_1^n \alpha_i x_i \quad \text{and} \quad \|x_i - x_0\| \geq \delta. \tag{2.11}$$

We can find in Ω disjoint subsets $A_1, \dots, A_n$ such that $\mathbb{P}(A_i) = \alpha_i$ and $\cup A_i = \Omega$. We then let $\mathcal{A}_0$ be the trivial σ-algebra and let $\mathcal{A}_1$ be the σ-algebra generated by $A_1, \dots, A_n$. Then we define $M_1(\omega) = x_i$ if $\omega \in A_i$. Clearly (2.11) implies $\mathbb{E}^{\mathcal{A}_0} M_1 = M_0$ and $\|M_1 - M_0\| \geq \delta$ everywhere. Since each point x_i is in D, we can continue in this way and represent each x_i as a convex combination analogous to (2.11). This will give M_2, M_3, etc.

We skip the details of the obvious induction argument. This yields a δ-separated martingale and hence a δ-separated tree. This completes the proof of (v) $\Rightarrow$ (vi) and hence of Theorem 2.5.

Finally, as promised, let us give a direct argument for (iv) $\Rightarrow$ (i). Assume (iv) and let μ be a B-valued vector measure such that $|\mu| \ll \nu$ where ν is as in the definition of the RNP. Then, by the classical RN theorem, there is a scalar density w such that $|\mu| = w.\nu$, thus it suffices to produce a RN density for μ with respect to $|\mu|$, so that, replacing ν by $|\mu|$ and normalizing, we may as well assume that we have a probability $\mathbb{P}$ such that

$$\forall A \in \mathcal{A} \quad \|\mu(A)\| \leq \mathbb{P}(A).$$

Then for any finite σ-subalgebra $\mathcal{B} \subset \mathcal{A}$, generated by a finite partition $A_1, \cdots, A_N$ of Ω, we consider the $\mathcal{B}$-measurable (step) function $f_{\mathcal{B}} \colon \Omega \to B$ that is equal to $\mu(A_j)\mathbb{P}(A_j)^{-1}$ on each atom A_j of $\mathcal{B}$. It is then easy to check that $\{f_{\mathcal{B}} \mid \mathcal{B} \subset \mathcal{A}, |\mathcal{B}| < \infty\}$ is a martingale indexed by the directed set of all such $\mathcal{B}$'s. By the above Remark 1.7, if (iv) holds then the resulting net converges in $L_p(B)$, and a fortiori in $L_1(B)$ to a limit $f \in L_1(B)$. By the continuity of $\mathbb{E}^{\mathcal{C}}$, for each fixed finite $\mathcal{C}$, $\mathbb{E}^{\mathcal{C}}(f_{\mathcal{B}}) \to \mathbb{E}^{\mathcal{C}}(f)$ in $L_1(B)$, and $\mathbb{E}^{\mathcal{C}}(f_{\mathcal{B}}) = f_{\mathcal{C}}$ when $\mathcal{C} \subset \mathcal{B}$, therefore we must have $\mathbb{E}^{\mathcal{C}}(f) = f_{\mathcal{C}}$ for any finite $\mathcal{C}$. Applying this to an arbitrary $A \in \mathcal{A}$, taking for $\mathcal{C}$ the σ-subalgebra generated by A (and its complement), we obtain (recall that $f_{\mathcal{C}}$ is constant on A, equal to $\mu(A)\mathbb{P}(A)^{-1}$)

$$\mathbb{E}(1_A f) = \mathbb{E}(1_A f_{\mathcal{C}}) = \mathbb{P}(A) \times \mu(A)\mathbb{P}(A)^{-1} = \mu(A),$$

so that we conclude that $f.\mathbb{P} = \mu$, i.e. we obtain (i). □

Remark. If the preceding property (vi) is weakened by considering only dyadic trees (i.e. martingales relative to the standard dyadic filtration on say $[0,1]$), or k-regular trees, then it does not imply the RNP: Indeed, by [89] there is a Banach space B (isometric to a subspace of L_1) that does not contain any bounded δ-separated *dyadic* tree, but that fails the RNP. Actually, that same paper shows that for any given sequence $(K(n))$ of integers, there is a Banach space B failing the RNP but not containing any δ-separated tree relative to a filtration such that $|\mathcal{A}_n| \le K(n)$ for all n.

Corollary 2.6. *If for some $1 \le p \le \infty$ every B-valued martingale bounded in $L_p(B)$ converges a.s. then the same property holds for all $1 \le p \le \infty$.*

Remark 2.7. Note that for $1 < p < \infty$, if a B-valued martingale (M_n) is bounded in $L_p(B)$ and converges a.s. to a limit f, then it automatically also converges to f in $L_p(B)$. Indeed, by the maximal inequalities (1.20) the convergence of $\|M_n - f\|^p$ to zero is dominated, hence by Lebesgue's theorem $\int \|M_n - f\|^p d\mathbb{P} \to 0$.

Corollary 2.8. *The RNP is separably determined, that is to say: if every separable subspace of a Banach space B has the RNP, then B also has it.*

Proof. This follows from Theorem 2.5 by observing that a B-valued martingale in $L_1(B)$ must "live" in a separable subspace of B. Alternately, note that any δ-separated tree is included in a separable subspace. □

Corollary 2.9. *If a Banach space B satisfies either one of the properties (ii)–(v) in Theorem 2.5 for martingales adapted to the standard dyadic filtration on [0,1], then B has the RNP.*

Proof. It is easy to see by a suitable approximation that if B contains a bounded δ-separated tree, then it contains one defined on a subsequence $\{\mathcal{A}_{n_k} \mid k \ge 1\}$, $(n_1 < n_2 < \ldots)$ of the dyadic filtration $(\mathcal{A}_n)$ in [0,1]. This yields the desired conclusion. □

Corollary 2.10. *If a Banach space B satisfies the property in Definition 2.1 when (Ω, ν) is the Legesgue interval $([0,1], dt)$, then B has the RNP.*

Corollary 2.11. *Any reflexive Banach space and any separable dual have the RNP.*

Proof. Since the RNP is separably determined by Corollary 2.8, it suffices to prove that separable duals have the RNP. So assume $B = X^*$, and that B is separable. Note that X is necessarily separable too and the closed unit ball of B is a metrizable compact set for $\sigma(X^*, X)$. Let $\{M_n\}$ be a martingale with values in the latter unit ball. For any ω, let $f(\omega)$ be a cluster point for $\sigma(X^*, X)$ of $\{M_n(\omega) \mid n \ge 0\}$. Let $D \subset X$ be a countable dense subset of the unit ball of X.

For any d in D, the bounded scalar martingale $\langle d, M_n\rangle$ converges almost surely to a limit which has to be equal to $\langle d, f(\omega)\rangle$. Hence since D is countable, there is $\Omega' \subset \Omega$ with $\mathbb{P}(\Omega') = 1$ such that

$$\forall\ \omega \in \Omega' \quad \forall\ d \in D \qquad\qquad \langle d, M_n(\omega)\rangle \to \langle d, f(\omega)\rangle.$$

In other words we have $M_n(\omega) \xrightarrow{\sigma(X^*,D)} f(\omega)$ or equivalently (since we are in the unit ball of B) $M_n(\omega) \xrightarrow{\sigma(X^*,X)} f(\omega)$ for any ω in Ω'. Notice that we did not discuss the measurability of f yet. But now we know that $\omega \to \langle x, f(\omega)\rangle$ is measurable for any x in X, hence since X is separable for any $x_0 \in B$, $\omega \to \|x_0 - f(\omega)\|$ is measurable, so $f^{-1}(\beta) = \{\omega \mid f(\omega) \in \beta\}$ is measurable for any open (or closed) ball $\beta \subset X^*$, and finally since X^* is separable, for any open set $U \subset X^*$, the set $f^{-1}(U)$ must be measurable, so f is Borel measurable.

We claim that this implies that f is Bochner measurable. This (and the desired conclusion) follows from Phillips' theorem (see Appendix 2). Alternatively we can conclude the proof by the same trick as in Appendix 2, as follows.
For any x_0 in B we have

$$\|x_0 - M_n\| = \sup_{d\in D, \|d\|\le 1} |\langle d, x_0 - M_n\rangle| = \sup_{d\in D, \|d\|\le 1} |\mathbb{E}_n\langle d, x_0 - f\rangle| \le \mathbb{E}_n\|x_0 - f\|$$

(note that $\omega \to \|x_0 - f(\omega)\|$ is bounded and measurable, so that $\mathbb{E}_n\|x_0 - f\| \xrightarrow{\text{a.s.}} \|x_0 - f\|$). Hence $\limsup_n \|x_0 - M_n\| \overset{\text{a.s.}}{\le} \|x_0 - f\|$. We can assume that this holds on the same set of probability one for all x_0 in a countable dense subset of B, hence actually for all x_0 in B. But then taking $x_0 = f(\omega)$ we have for almost all ω, $\limsup_{n\to\infty} \|f(\omega) - M_n(\omega)\| = 0$. Thus we conclude by Theorem 2.5 that B has the RNP. $\square$

Remark. The above examples of divergent martingales show that the separable Banach spaces $L_1([0, 1])$ and c_0 fail the RNP.

Remark. The RNP is clearly stable by passing to subspaces but obviously not to quotients Indeed, ℓ_1, being a separable dual, has the RNP but any separable space (e.g. c_0) is a quotient of it.

Notation. By analogy with the Hardy space case, let us denote by

$$h_p(\Omega, (\mathcal{A}_n)_{n\ge 0}, \mathbb{P}; B)$$

the (Banach) space of all B-valued martingales $M = \{M_n \mid n \ge 0\}$ that are bounded in $L_p(B)$, equipped with the norm

$$\|M\| = \sup_{n\ge 0} \|M_n\|_{L_p(B)}.$$

Remark 2.12. Note that, by Theorem 1.5, the mapping

$$f \to \{E_n(f) \mid n \ge 0\}$$

defines an isometric embedding of

$$L_p(\Omega, \mathcal{A}_\infty, \mathbb{P}; B) \text{ into } h_p(\Omega, (\mathcal{A}_n)_{n\ge 0}, \mathbb{P}; B).$$

Remark 2.13. Let $1 < p \leq \infty$. With this notation, Theorem 2.5 says that B has the RNP iff

$$h_p(\Omega, (\mathcal{A}_n)_{n\geq 0}, \mathbb{P}; B) = L_p(\Omega, \mathcal{A}_\infty, \mathbb{P}; B).$$

We now turn to the identification of the dual of $L_p(B)$.
Let p' be the conjugate exponent such that $p^{-1} + {p'}^{-1} = 1$.
Suppose that we are given a filtration $\mathcal{A}_0 \subset \dots \mathcal{A}_n \subset \mathcal{A}_{n+1} \subset \dots$ of finite σ-subalgebras and let us assume $\mathcal{A} = \mathcal{A}_\infty$. Let $L_p(B) = L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$ with $1 \leq p < \infty$. Let φ be a bounded linear form on $L_p(B)$. By restriction to $L_p(\Omega, \mathcal{A}_n, \mathbb{P}; B), \varphi$ defines a linear form φ_n in $L_p(\Omega, \mathcal{A}_n, \mathbb{P}; B)^*$. But, since $\mathcal{A}_n$ is finite, we have $L_p(\Omega, \mathcal{A}_n, \mathbb{P}; B)^* = L_{p'}(\Omega, \mathcal{A}_n, \mathbb{P}; B^*)$ isometrically, hence φ_n corresponds to an element M_n in $L_p(\Omega, \mathcal{A}_n, \mathbb{P}; B^*)$. Moreover, since φ_n is the restriction of φ_{n+1} it is easy to see that $M_n = \mathbb{E}_n(M_{n+1})$, i.e. that $\{M_n\}$ is a B^*-valued martingale. Moreover, we have

$$\sup_n \|M_n\|_{L_{p'}(B^*)} = \|\varphi\|_{L_p(B)^*}.$$

Proposition 2.14. *In the above situation, the correspondence*

$$\varphi \to (M_n)_{n\geq 0}$$

is an isometric isomorphism from $L_p(\Omega, \mathcal{A}, \mathbb{P}; B)^*$ to the space

$$h_{p'}(\Omega, (\mathcal{A}_n)_{n\geq 0}, \mathbb{P}; B^*).$$

Proof. Indeed, it is easy to see conversely that given any martingale $\{M_n\}$ in the unit ball of $h_{p'}(\Omega, (\mathcal{A}_n)_{n\geq 0}, \mathbb{P}; B^*)$, M_n defines an element φ_n in $L_p(\Omega, \mathcal{A}_n, \mathbb{P}; B)^*$ so that φ_{n+1} extends φ_n, and $\|\varphi_n\| \leq 1$. Hence by density of the union of the spaces $L_p(\Omega, \mathcal{A}_n, \mathbb{P}; B)$ in $L_p(B)$, we can extend the φ_n's to a (unique) functional φ in $L_p(B)^*$ with $\|\varphi\| \leq 1$. Thus, it is easy to check that the correspondence is one-to-one and isometric. □

Remark 2.15. By Remark 2.12, we have an isometric embedding

$$L_{p'}(\Omega, \mathcal{A}_\infty, \mathbb{P}; B^*) \subset L_p(\Omega, \mathcal{A}_\infty, \mathbb{P}; B)^*.$$

Theorem 2.16. *A dual space B^* has the RNP iff for any countably generated measure space and any $1 \leq p < \infty$ we have (isometrically)*

$$L_p(\Omega, \mathcal{A}, \mathbb{P}; B)^* = L_{p'}(\Omega, \mathcal{A}, \mathbb{P}; B^*).$$

Moreover for B^* to have the RNP it suffices that this holds for some $1 \leq p < \infty$ and for the Lebesgue interval.

Proof. If $\mathcal{A}$ is countably generated we can assume $\mathcal{A} = \mathcal{A}_\infty$ with $\mathcal{A}_\infty$ associated to a filtration of finite σ-algebras $(\mathcal{A}_n)$ as above. Then Theorem 2.16 follows from Proposition 2.14 and Remark 2.13. The second assertion folllows from Corollary 2.9. □

Remark. Of course the preceding isometric duality holds for any dual space B^* when the measure space is discrete (i.e. atomic).

Remark. The preceding theorem does remain valid for $p = 1$. Note however that, if $\dim(B) = \infty$, the B-valued step functions are, of course, *not dense* in the space $L_\infty(\Omega, \mathcal{A}, \mathbb{P}; B)$. This is in sharp contrast with the finite dimensional case. But if $\dim(B) = \infty$ the unit ball is no longer compact, there is no finite ε net for small ε, so, in general, we cannot uniformly approximate even the nicest bounded continuous functions by step functions, i.e. functions taking only finitely many values. Recall that instead, we defined the space $L_\infty(\Omega, \mathcal{A}, \mathbb{P}; B)$ (in Bochner's sense) as the space of B-valued Bochner-measurable functions f (see Appendix 2) such that $\|f(.)\|_B$ is in L_∞, equipped with its natural norm. This definition makes sense for any measure space $(\Omega, \mathcal{A}, \mathbb{P})$, and, with it, the preceding theorem is valid for $p = 1$.

A function $f\colon\ \Omega \to B^*$ will be called weak* scalarly measurable if for every b in B the scalar valued function $\langle f(.), b\rangle$ is measurable. Assume B separable. Let us denote by $\Lambda_p(\Omega, \mathcal{A}, \mathbb{P}; B^*)$ the space of (equivalence classes of) scalarly measurable functions $f\colon\ \Omega \to B^*$ such that the function $\omega \mapsto \|f(\omega)\|_{B^*}$ (which is measurable since B is separable) is in L_p. We equip this space with the obvious norm

$$\|f\| = (\int \|f(\omega)\|^p_{B^*})^{1/p}.$$

We have then

Theorem 2.17. *Assume B separable. Then for any countably generated measure space and any $1 \le p < \infty$ we have (isometrically)*

$$L_p(\Omega, \mathcal{A}, \mathbb{P}; B)^* = \Lambda_{p'}(\Omega, \mathcal{A}, \mathbb{P}; B^*).$$

Proof. We assume as before that $\mathcal{A}$ is generated by a filtration of finite algebras. By Proposition 2.14, it suffices to show how to identify $h_{p'}(\Omega, (\mathcal{A}_n)_{n\ge 0}, \mathbb{P}; B^*)$ with $\Lambda_{p'}(\Omega, \mathcal{A}, \mathbb{P}; B^*)$. Consider a martingale (f_n) in $h_{p'}(\Omega, (\mathcal{A}_n)_{n\ge 0}, \mathbb{P}; B^*)$. By the maximal inequality, (f_n) is bounded a.s. and hence a.s. weak* compact. Let $f(\omega)$ be a weak* cluster point of (f_n). Then for any fixed $b \in B$, the scalar martingale $\langle f_n(.), b\rangle$ converges a.s.. Its limit must necessarily be equal to $\langle f(.), b\rangle$. This shows that f is weak* scalarly measurable. Let D be a countable dense subset of the unit ball of B. Since D is countable, and $\langle f(.), b\rangle = \lim_{n\to\infty}\langle f_n(.), b\rangle$ for any $b \in D$, we have a.s.

$$\|f\| = \sup_{b\in D} |\langle f, b\rangle| \le \lim_{n\to\infty} \|f_n\|$$

and hence by Fatou's lemma

$$\|f\|_{\Lambda_{p'}} \le \|(f_n)\|_{h_{p'}}.$$

Conversely, consider now $f \in \Lambda_{p'}(\Omega, \mathcal{A}, \mathbb{P}; B^*)$. Fix n. Let A be an atom of $\mathcal{A}_n$. Then $b \mapsto \mathbb{P}(A)^{-1}\int_A \langle b, f\rangle$ is a continuous linear form on B with norm

$\leq \mathbb{P}(A)^{-1}\int_A \|f\|_{B^*}$. Let us denote it by f_A. Let f_n be the B^*-valued function that is equal to f_A on each atom $A \in \mathcal{A}_n$. We have clearly $\mathbb{E}_n(\langle b, f\rangle) = \langle b, f_n\rangle$ and hence $\mathbb{E}_n(\langle b, f_{n+1}\rangle) = \langle b, f_n\rangle$ for any b in D. Since D separates points, this shows that (f_n) is a martingale, and moreover $\|f_n\| = \sup_{b\in D} |\langle b, f_n\rangle| \leq \mathbb{E}_n\|f\|$. It follows that

$$\|(f_n)\|_{h_{p'}} \leq (\int \|f\|^{p'})^{1/p'} = \|f\|_{\Lambda_{p'}}.$$

This shows that the correspondence $(f_n) \mapsto f$ is an isometric isomorphism from $h_{p'}$ to $\Lambda_{p'}$. □

Remark 2.18. The notion of "quasi-martingale" is useful to work with random sequences which are obtained by perturbation of a martingale. An adapted sequence $(F_n)_{n\geq 0}$ in $L_1(B)$ is said to be a quasi-martingale if

$$\sum_1^\infty \|\mathbb{E}_{n-1}(F_n - F_{n-1})\|_{L_1(B)} < \infty.$$

Given such a sequence, let

$$f_n = F_n - \sum\nolimits_1^n \mathbb{E}_{k-1}(F_k - F_{k-1}),$$

so that

$$df_n = dF_n - \mathbb{E}_{n-1}(dF_n).$$

Clearly (f_n) is then a martingale and for all $m < n$ we have pointwise

$$\|(f_n - f_m) - (F_n - F_m)\|_B \leq \sum\nolimits_{m<k\leq n} \|E_{k-1}(F_k - F_{k-1})\|_B$$

and hence

$$\|(f_n - f_m) - (F_n - F_m)\|_{L_1(B)} \leq \sum\nolimits_{m<k\leq n} \|E_{k-1}(F_k - F_{k-1})\|_{L_1(B)}.$$

Note that (F_n) is bounded in $L_1(B)$ (resp. uniformly integrable) iff the same is true for (f_n). Therefore, if this holds and if B has the RNP, (F_n) converges a.s. (resp. and in $L_1(B)$).

The following complements the panorama of the interplay between martingale convergence and Radon-Nikodym theorems. This statement is valid for general Banach spaces, but we should emphasize for the reader that the ω-a.s. convergence of the variables $\omega \mapsto \|f_n(\omega)\|$ is considerably weaker than that of the sequence $(f_n(\omega))$ itself. The latter requires the RNP by Theorem 2.5.

Proposition. *Let B be an arbitrary Banach space. Consider $\mu \in M(\Omega, \mathcal{A}; B)$ such that $|\mu| = w \cdot \mathbb{P}$ where $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{A})$ and $w \in L_1(\Omega, \mathcal{A}, \mathbb{P})$. Let $(\mathcal{A}_n)_{n\geq 0}$ be a filtration such that $\mathcal{A}_\infty = \mathcal{A}$, and such that, for each n, $\mu_{|\mathcal{A}_n}$ admits a RN density f_n in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$ (for instance this is automatic if $\mathcal{A}_n$ is finite or atomic). Then $\|f_n\| \to w$ a.s.*

Proof. By Proposition 2.1, for each fixed $\varepsilon > 0$ we can find unit vectors $\xi_1, \ldots, \xi_N$ in B^* such that the vector measure

$$\mu_N \colon \ \mathcal{A} \to \ell_\infty^N$$

defined by $\mu_N(A) = (\xi_j(\mu(A)))_{j \le N}$ satisfies $|\mu_N|(\Omega) > |\mu|(\Omega) - \varepsilon = 1 - \varepsilon$. Assume $|\mu|(\Omega) = \int w \, d\mathbb{P} = 1$ for simplicity. Note that $|\mu_{|\mathcal{A}_n}| \le |\mu|_{|\mathcal{A}_n} = w_n \cdot \mathbb{P}$ where $w_n = \mathbb{E}^{\mathcal{A}_n} w$. Therefore $\|f_n\| \le w_n$. By the martingale convergence Theorem 1.5, $w_n \to w$ a.s. and in L_1, and hence

$$\limsup \|f_n\| \le w \quad \text{a.e.}$$

and $\int \limsup \|f_n\| \le \int w = 1$. We claim that

$$\int \liminf \|f_n\| \ge \int \liminf \sup_{j \le N} |\xi_j(f_n)| = |\mu_N|(\Omega) > 1 - \varepsilon.$$

Indeed, being finite dimensional, ℓ_∞^N has the RNP and hence $\mu_N = \varphi_N \cdot \mathbb{P}$ for some φ_N in $L_1(\Omega; \mathcal{A}, \mathbb{P}; \ell_\infty^N)$. This implies (by (2.2)) $|\mu_N| = \|\varphi_N\| \cdot \mathbb{P}$. Clearly $\mathbb{E}^{\mathcal{A}_n} \varphi_N = (\xi_j(f_n))_{j \le N}$ and hence

$$\sup\nolimits_{j \le N} |\xi_j(f_n)| \to \|\varphi_N\| \quad \text{a.s. and}$$

in L_1. Thus

$$\mathbb{E} \liminf \sup\nolimits_{j \le N} |\xi_j(f_n)| = \int \|\varphi_N\| d\mathbb{P} = |\mu_N|(\Omega) > 1 - \varepsilon,$$

proving the above claim.

Using this claim, we conclude easily: We have $\liminf \|f_n\| \le \limsup \|f_n\| \le w$ but $\int \liminf \|f_n\| d\mathbb{P} > \int w \, d\mathbb{P} - \varepsilon$, so we obtain $\liminf \|f_n\| = \limsup \|f_n\| = w$ a.e. □

2.2 The Krein Milman property

Recall that a point x in a convex set $C \subset B$ is called extreme in C if whenever x lies inside a segment $S = \{\theta y + (1-\theta) z \mid 0 < \theta < 1\}$ with endpoints y, z in C, then we must have $y = z = x$. Equivalently $C \backslash \{x\}$ is convex. See [16] and [8] for more information.

Definition. We will say that a Banach space B has the Krein Milman property (in short KMP) if every closed bounded convex set in B is the closed convex hull of its extreme points.

We will show below that RNP $\Rightarrow$ KMP.

The converse remains a well known important open problem (although it is known that RNP is equivalent to a stronger form of the KMP, see below). We will use the following beautiful fundamental result due to Bishop and Phelps, but we will skip the proof (see e.g. [16, p. 189]).

Theorem 2.19 (Bishop–Phelps). *Let $C \subset B$ be a closed bounded convex subset of a Banach space B. Then the set of functionals in B^* that attain their supremum on C is dense in B^*.*

Remark 2.20. (i) Let $x^* \in B^*$ be a functional attaining its supremum on C, so that if $\alpha = \sup\{x^*(b) \mid b \in C\}$, the set $F = \{b \mid x^*(b) = \alpha\}$ is non-void. We will say that F is a *face* of C. We need to observe that a face enjoys the following property: If a point in F is inside the segment joining two points in C, then this segment must entirely lie in F.

(ii) In particular, any extreme point of F is an extreme point of C.

(iii) Now assume that we have been able to produce a decreasing sequence of sets $\cdots \subset F_n \subset F_{n-1} \subset \cdots F_0 = C$ such that F_n is a face of F_{n-1} for any $n \geq 1$ and the diameter of F_n tends to zero. Then, by the Cauchy criterion, the intersection of the F_n's contains exactly one point x_0 in C. We claim that x_0 is an extreme point of C. Indeed, if x_0 sits inside a segment S joining two points in C, then by (i) we have $S \subset F_1$, hence (since F_2 is a face in F_1 and $x_0 \in F_2$) $S \subset F_2$ and so on. Hence $S \subset \cap F_n = \{x_0\}$, which shows that x_0 is extreme in C.

(iv) Assume that every closed bounded convex subset $C \subset B$ has at least one extreme point. Then B has the KMP. Indeed, let $C_1 \subset C$ be the closed convex hull of the extreme points of C. We must have $C_1 = C$. Indeed, otherwise there is x in $C \backslash C_1$ and by Hahn–Banach there is x^* in B^* such that $x^*_{|C_1} < \beta$ and $x^*(x) > \beta$. Assume first that this functional achieves its supremum $\alpha = \sup\{x^*(b) \mid b \in C\}$. This case is easier. Note $\alpha > \beta$. Then let $F = \{b \in C \mid x^*(b) = \alpha\}$, so that F is a face of C disjoint from C_1. But now F is another non-void closed bounded convex set that, according to our assumption, must have an extreme point. By (ii) this point is also extreme in C, but this contradicts the fact that F is disjoint from C_1.

In general, x^* may not achieve its norm, but we can use the Bishop–Phelps Theorem 2.19 to replace x^* by a small perturbation of itself that will play the same role in the preceding argument.

Indeed, by Theorem 2.19, for any $\varepsilon > 0$ there is y^* in B^* with $\|x^* - y^*\| < \varepsilon$ that achieves its sup on C. We may assume $\|b\| \leq r$ for any b in C. Let $\gamma = \sup\{y^*(b) \mid b \in C\}$ and note that $\gamma > \alpha - r\varepsilon$; and hence $y^*(b) = \gamma$ implies $x^*(b) > \alpha - 2r\varepsilon$. Hence if ε is chosen so that $\alpha - 2r\varepsilon > \beta$, we are sure that $\widetilde{F} = \{b \in C \mid y^*(b) = \gamma\}$ is included in $\{b \mid x^*(b) > \beta\}$ hence is disjoint from C_1. We now repeat the preceding argument: $\widetilde{F}$ must have an extreme point, by (ii) it is extreme in C hence must be in C_1, but this contradicts $\widetilde{F} \cap C_1 = \emptyset$.

(v) The preceding argument establishes the following general fact: let $\mathcal{S}$ be a slice of C, i.e. we assume given x^* in B^* and a number β so that

$$\mathcal{S} = \{b \in C \mid x^*(b) > \beta\},$$

then if $\mathcal{S}$ is non-void it must contain a (non-void) face of C.

Theorem 2.21. *The RNP implies the KMP.*

Proof. Assume B has the RNP. Let $C \subset B$ be a bounded closed convex subset. Then by Theorem 2.5, C is dentable. So for any $\varepsilon > 0$, there is x in C such that $x \notin \overline{\text{conv}}(C \backslash B(x, \varepsilon))$. By Hahn–Banach separation, there is x^* in B^* and a number β such that the slice $\mathcal{S} = \{b \in C \mid x^*(b) > \beta\}$ contains x and is disjoint from $C \backslash B(x, \varepsilon)$. In particular, we have $\|b - x\| \leq \varepsilon$ for any b in $\mathcal{S}$, so the diameter of $\mathcal{S}$ is $\leq 2\varepsilon$. By Remark 2.20 (v), $\mathcal{S}$ must contain a face F_1 of C, a fortiori of diameter $\leq 2\varepsilon$.

Now we can repeat this procedure on F_1 : we find that F_1 admits a face F_2 of arbitrary small diameter, then F_2 also admits a face of small diameter, and so on. Thus, adjusting $\varepsilon > 0$, we find a sequence of (non-void) sets $\cdots \subset F_{n+1} \subset F_n \subset \cdots \subset F_1 \subset F_0 = C$ such that F_{n+1} is a face of F_n and $\text{diam}(F_n) < 2^{-n}$. Then, by Remark 2.20 (iii), the intersection of $\{F_n\}$ contains an extreme point of C. By Remark 2.20 (iv), we conclude that B has the KMP. $\square$

Let $C \subset B$ be a convex set. A point x in C is called "exposed" if there is a functional x^* such that $x^*(x) = \sup\{x^*(b) \mid b \in C\}$ and x is the only point of C satisfying this. (Equivalently, if the singleton $\{x\}$ is a face of C.) The point x is called "strongly exposed" if the functional x^* can be chosen such that, in addition, the diameter of the slice

$$\{b \in C \mid x^*(b) > \sup_C x^* - \varepsilon\}$$

tends to zero when $\varepsilon \to 0$. Clearly, the existence of such a point implies that C is dentable. More precisely, if C is the closed convex hull of a bounded set D, then D is dentable because every slice of C contains a point in D (see Remark 2.2).

We will say that B has the "strong KMP" if every closed bounded convex subset $C \subset B$ is the closed convex hull of its strongly exposed points. It is clear (by (vi) $\Rightarrow$ (i) in Theorem 2.5) that the strong KMP implies the RNP. That the converse also holds is a very beautiful and deep result due to Bob Phelps [225]:

Theorem 2.22. *The RNP is equivalent to the strong KMP.*

2.3 Edgar's Choquet Theorem

2.4 Notes and Remarks

For vector measures and Radon–Nikodym theorems, a basic reference is [16]. A more recent, much more advanced, but highly recommended reading is Bourgain's Lecture Notes on the RNP [81].

For the Banach space valued case, the first main reference is Chatterji's paper [112] where the equivalence of (i), (ii), (iii) and (iv) in Theorem 2.5 is proved. The statements numbered from 2.6 to 2.16 all follow from Chatterji's result but some of them were probably known before.

Rieffel introduced dentability and proved that it suffices for the RNP. The converse is (based on work by Maynard) due to Davis–Phelps and Huff independently. The Lewis–Stegall theorem in §**??** comes from [188]. Theorem 2.21 is due to Joram Lindenstrauss and Theorem thm1.31a to Phelps [225]. See [16] for a more detailed history of the RNP and more precise references.

Our presentation of the RNP is limited to the basic facts. We will now briefly survey additional material.

In §2.3 we present Edgar's theorem (improving Theorem 2.21) that the RNP implies a Choquet representation theorem. This is proved using the martingale convergence theorem and a basic measure theoretic result (namely von Neumann measurable lifting theorem). See also [183] for more illustrations of the use of Banach valued martingales.

Charles Stegall [252] proved the following beautiful characterization of duals with the RNP:

Stegall's Theorem ([252])
Let B be a separable Banach space. Then B^* has the RNP iff it is separable. More generally, a dual space B^* has the RNP iff for any separable subspace $X \subset B$, the dual X^* is separable.

In the 80's, a lot of work was devoted (notably at the impulse of H.P. Rosenthal and Bourgain) to "semi-embeddings". A Banach space X is said to semi-embed in another one Y if there is an injective linear mapping $u\colon\ X \to Y$ such that the image of the closed unit ball of X is closed in B (and such a u is then called a semi-embedding). The relevance of this notion lies in

Proposition 2.23. *If X is separable and semi-embeds in a space Y with the RNP, then X has the RNP.*

Proof. One way to prove this is to consider a martingale (f_n) with values in the closed unit ball B_X of X. Let $u\colon\ X \to Y$ be a semi-embedding. If Y has RNP then the martingale $g_n = u(f_n)$ converges in Y to a limit g_∞ such that $g_\infty(\cdot) \in \overline{u(B_X)} = u(B_X)$. Let now $f(\omega) = u^{-1}(g_\infty(\omega))$. We will show that f is Borel measurable. Let U be any open set in X. By separability, there is a sequence $\{\beta_n\}$ of closed balls in X such that $U = \cup\beta_n$. Then

$$\{\omega \mid f(\omega) \in U\} = \cup_n\{\omega \mid g_\infty(\omega) \in u(\beta_n)\}$$

but since $u(\beta_n)$ is closed and g_∞ measurable we find that $f^{-1}(U)$ is measurable. This shows that f is Borel measurable. By Phillips' theorem, f is Bochner measurable. Now, since $g_n = \mathbb{E}_n(g_\infty) = \mathbb{E}_n(u(f)) = u(\mathbb{E}_n(f))$ we have

$$f_n = u^{-1}(g_n) = \mathbb{E}_n(f),$$

and hence f_n converges to f a.s. This shows that X has the RNP (clearly one could use a vector measure instead of a martingale and obtain the RNP a bit more directly). $\square$

We refer to [90] for work on semi-embeddings. More generally, an injective linear map $u\colon X \to Y$ is called a G_δ-embedding if the image of any closed bounded subset of X is a G_δ-subset of Y.

We refer to [140, 141, 143, 144, 149] for Ghoussoub and Maurey's work on G_δ-embeddings. To give the flavor of this work, let us quote the main result of [141]: A separable Banach space X has the RNP iff there is a G_δ-embedding $u\colon X \to \ell_2$ such that $u(B_X)$ is a countable intersection of open sets with convex complements.

The proof of the above Proposition 2.23 shows that the RNP is stable under G_δ-embedding.

As mentioned in the text, it is a famous open problem whether KMP implies RNP. It was proved for dual spaces by Huff and Morris using the above theorem of Stegall [252], see [8, p. 91], and also for Banach lattices by Bourgain and Talagrand ([8, p. 423]). See also Chu's paper [113] for preduals of von Neumann algebras. Schachermayer [245] proved that it is true for Banach spaces isomorphic to their square. See also [246, 247, 248] for related work by the same author.

We should mention that one can define the RNP for *subsets* of Banach spaces. One can then show that weakly compact sets are RNP sets. See [8, 50] for more on RNP sets.

A Banach space X is called an Asplund space if every continuous convex function defined on a (non-empty) convex open subset $D \subset E$ is Fréchet differentiable on a dense G_δ-subset of D. Stegall [253] proved that X is Asplund iff X^* has the RNP. We refer the reader to [50] for more information in this direction.

Chapter 3

Super-reflexivity

3.1 Finite representability and Super-properties

The notion of "finite representability" is the basis for that of "super-property."

Definition. A Banach space X is said to be finitely representable (f.r. in short) in another Banach space Y if for any finite dimensional subspace $E \subset X$ and for any $\varepsilon > 0$ there is a subspace $\widetilde{E} \subset Y$ that is $(1+\varepsilon)$-isomorphic to E (i.e. there is an isomorphism $u\colon\ E \to \widetilde{E}$ with $\|u\|\|u^{-1}\| \le 1+\varepsilon$).

In other words, X f.r. Y means that, although Y may not contain an isomorphic copy of the whole of X, it contains an almost isometric copy of any finite dimensional subspace of X. In Appendix 1 to this chapter devoted to background on ultraproducts, we show that X is f.r. Y iff X embeds isometrically in an ultraproduct of Y.

The following simple perturbation argument will be used repeatedly.

Lemma 3.1. *Let X, Y be Banach spaces. Let $E_0 \subset \cdots \subset E_n \subset E_{n+1} \subset \cdots$ be a sequence (or a family directed by inclusion) of finite dimensional subspaces of X such that $\overline{\cup E_n} = X$. Then for X to be f.r. in Y it suffices that for any $\varepsilon > 0$ and any n there is a subspace $\widetilde{E}_n \subset Y$ that is $(1+\varepsilon)$-isomorphic to E_n.*

Proof. Consider $E \subset X$ with $\dim(E) < \infty$. It suffices to show that for any fixed $\varepsilon > 0$ there is n and $\widehat{E} \subset E_n$ such that E is $(1+\varepsilon)$ isomorphic to $\widehat{E}$. Let $\delta > 0$ to be specified later. Let $x_1, \ldots, x_d$ be a linear basis of E. Choose n and $\hat{x}_1, \ldots, \hat{x}_d$ in E_n such that $\|x_j - \hat{x}_j\| < \delta$ for all $j = 1, \ldots, d$. Let $v\colon\ E \to \widehat{E}$ be the linear map determined by $v(x_j) = \hat{x}_j$. For any $(\alpha_j) \in \mathbb{K}^d$ we have by the triangle inequality

$$\left|\left\|v\left(\sum \alpha_j x_j\right)\right\| - \left\|\sum \alpha_j x_j\right\|\right| \le \delta \sum |\alpha_j|, \tag{3.1}$$

but since all norms are equivalent on $\mathbb{R}^d$ there is a constant C_E such that

$$\sum |\alpha_j| \le C_E \left\|\sum \alpha_j x_j\right\|.$$

Thus (3.1) implies

$$(1-\delta C_E)\left\|\sum \alpha_j x_j\right\| \le \left\|\sum \alpha_j \hat{x}_j\right\| \le (1+\delta C_E)\left\|\sum \alpha_j x_j\right\|,$$

and hence $\widehat{E}$ is λ-isomorphic to E with $\lambda = (1+\delta C_E)(1-\delta C_E)^{-1}$. To conclude we simply choose δ small enough so that $\lambda < 1+\varepsilon$. $\square$

Remark. The preceding Lemma shows in particular that L_p is f.r. in ℓ_p for any $1 \le p < \infty$, and that L_∞ is f.r. in c_0.

Definition. Consider a property P for Banach spaces. We say that a Banach space Y has "super-P" if every Banach space X that is f.r. in Y has P.

Remark. In particular Y is super-reflexive (resp. has the super-RNP) if every X f.r. in Y is reflexive (resp. has the RNP). The passage from P to super-P is a fruitful way to associate to an *infinite dimensional* property (such as e.g. reflexivity) its finite dimensional counterpart. If the property P is already stable by finite representability, then P and super P are the same. Such properties are usually called "local." The "local theory" of Banach spaces designates the part of the theory that studies infinite dimensional spaces through the collection of their finite dimensional subspaces.

Remark 3.2. Let B be a complex Banach space. If B is super-reflexive as a real Banach space then it is also super-reflexive as a complex space. Indeed, any complex space X that is f.r. in B must be reflexive as a real space, but this is the same as reflexive as a complex space. Conversely, if B is super-reflexive as a complex space, it is also as a real space, but this is a bit less obvious. It follows e.g. from (i) $\Leftrightarrow$ (iii) in Theorem 3.22 below, since the notion of separated tree is the same in the real or complex cases. It also follows from Proposition 3.8 below.

The following result called the "local reflexivity principle" is classical.

Theorem 3.3 ([191])**.** *The bidual B^{**} of an arbitrary Banach space B is f.r. in B.*

To study super-reflexivity, we will need the following elementary fact.

Lemma 3.4. *Let B be a Banach space. Then for any b^{**} in B^{**} any $\varepsilon > 0$ and any finite subset $\xi_1, \ldots, \xi_n$ in B^* there is b in B with $\|b\| \le (1+\varepsilon)\|b^{**}\|$ such that*

$$\langle \xi_i, b^{**}\rangle = \langle \xi_i, b\rangle \qquad \forall i = 1, \ldots, n.$$

Proof. Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ be the scalar field. We may clearly assume ξ_i linearly independent. Assume $\|b^{**}\| = 1$ for simplicity. Let $C \subset \mathbb{K}^n$ be the convex set $\{(\langle \xi_i, b\rangle)_{i\le n} \mid b \in B \ \|b\| \le 1\}$. Clearly, since b^{**} is in the $\sigma(B^{**}, B^*)$ closure of the unit ball of B, we know that $(\langle \xi_i, b^{**}\rangle)_{i\le n} \in \overline{C}$. But (since we assumed the ξ_i's independent) C has nonempty interior hence $\overline{C} \subset (1+\varepsilon)C$ for any $\varepsilon > 0$. Thus we conclude that $(\langle \xi_i, b^{**}\rangle)_{i\le n} \in (1+\varepsilon)C$. $\square$

The following result is classical. It combines several known facts, notably (iv) $\Rightarrow$ (iii) goes back to R.C. James [162].

Theorem 3.5. *The following properties of a Banach space B are equivalent:*

(i) *Every Banach space is f.r. in B.*

(ii) *c_0 is f.r. in B.*

(iii) *For any $\lambda > 1$ and any $n \geq 1$ there are $x_1^n, \dots, x_n^n$ in B satisfying*

$$\forall (\alpha_j) \in \mathbb{K}^n \qquad \sup |\alpha_j| \leq \left\| \sum\nolimits_1^n \alpha_j x_j^n \right\| \leq \lambda \sup |\alpha_j|. \tag{3.2}$$

(iv) *For some $\lambda > 1$, for any $n \geq 1$ there are $x_1^n, \dots, x_n^n$ in B satisfy* (3.2).

(v) *For some $\lambda > 1$, for any $n \geq 1$ there are $x_1^n, \dots, x_n^n$ in B with norm ≥ 1 and such that*

$$\sup \left\{ \left\| \sum\nolimits_1^n \varepsilon_j x_j^n \right\| \mid \varepsilon_j = \pm 1 \right\} \leq \lambda.$$

Proof. (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv)$\Rightarrow$ (v) are trivial. We will show that (iii) $\Rightarrow$ (i) and (v) $\Rightarrow$ (iii). Assume (iii). Let $E \subset B_1$ be a finite dimensional subspace in an arbitrary Banach space B_1. Let $S = \{s_1, \dots, s_n\}$ be an ε-net in the unit sphere of E. For each $s_j \in S$, choose $\xi_j \in E^*$ such that $\langle \xi_j, s_j \rangle = 1 = \|\xi_j\|$. We define $u\colon\ E \to \ell_\infty^n$ by setting $u(x) = (\xi_j(x))_{j \leq n}$. We have $\|u\| = 1$ and

$$\forall s \in S \qquad \|u(s)\|_{\ell_\infty^n} = 1.$$

Therefore by Lemma 3.47

$$\forall x \in E \qquad (1-\varepsilon)\|x\| \leq \|u(s)\|_{\ell_\infty^n} \leq \|x\|.$$

This shows that E embeds $(1-\varepsilon)^{-1}$-isomorphically into ℓ_∞^n. Thus (iii) implies that B_1 is f.r. in B, or equivalently (iii) $\Rightarrow$ (i).

The proof that (v) $\Rightarrow$ (iii) is a well known "blocking trick" . Assume (v). Let $C(n)$ be the smallest constant C such that for any $x_1, \dots, x_n$ in B we have

$$\inf\nolimits_{j \leq n} \|x_j\| \leq C \sup_{\varepsilon_j = \pm 1} \left\| \sum\nolimits_1^n \varepsilon_j x_j \right\|.$$

A simple blocking argument shows that $C(nk) \leq C(n)C(k)$ for all n, k. Since we assume (v), we have $\inf_n C(n) \geq \lambda^{-1}$, but by the submultiplicativity of $C(n)$ this implies $C(n) \geq 1$ for all n. Therefore, for any n and any $\lambda > 1$ we can find $x_1, \dots, x_n$ in B such that $\sup_{\varepsilon_j = \pm 1} \|\sum \varepsilon_j x_j\| \leq \lambda$ and $\inf_{j \leq n} \|x_j\| \geq 1$. For each k, choose $\xi_k \in B^*$ such that $\|\xi_k\| = 1$ and $\xi_k(x_k) \geq 1$. Note that if ε_j is the sign of $\xi_k(x_j)$ we have

$$\sum |\xi_k(x_j)| = \left\langle \xi_k, \sum \varepsilon_j x_j \right\rangle \leq \left\| \sum \varepsilon_j x_j \right\| \leq \lambda.$$

Consequently

$$\sum_{j\neq k}|\xi_k(x_j)|\leq\lambda-1. \tag{3.3}$$

Let C be the set of real scalars $\alpha_1,\ldots,\alpha_n$ with $\sup|\alpha_j|\leq 1$. Note that the maximum value on C of $\|\sum\alpha_j x_j\|$ is attained on an extreme point of C (of the form $\alpha_j=\pm1$), so we have $\|\sum\alpha_j x_j\|\leq\lambda\sup|\alpha_j|$ for any $(\alpha_1,\ldots,\alpha_n)$ in $\mathbb{R}^n$. Let $x=\sum\alpha_j x_j$. Choose k so that $|\alpha_k|=\sup_j|\alpha_j|$. By (3.3) we have

$$\sup|\alpha_j|=|\alpha_k|=\left|\xi_k\left(\sum\alpha_j x_j\right)-\sum_{j\neq h}\alpha_j\xi_k(x_j)\right|\leq\|x\|+(\lambda-1)\sup|\alpha_j|$$

and hence we find $(2-\lambda)\sup|\alpha_j|\leq\|x\|$. Thus we conclude

$$\sup|\alpha_j|\leq(2-\lambda)^{-1}\|x\|\leq(2-\lambda)^{-1}\lambda\sup|\alpha_j|,$$

and since $(2-\lambda)^{-1}\lambda$ is arbitrarily close to 1 this shows that (v) $\Rightarrow$ (iii), at least in the real case. To check the complex case, note that

$$\sup_{z_j\in\mathbb{C}|z_j|=1}\left\|\sum z_j x_j\right\|\leq 2\sup_{\varepsilon_j=\pm1}\left\|\sum\varepsilon_j x_j\right\|.$$

From this one sees that in (v) we may replace the choices of signs by unimodular complex numbers and complete the proof of (v) $\Rightarrow$ (iii) exactly as in the real case.

□

Recall from §4.4:

Definition. We say that B contains ℓ_∞^n's uniformly if it satisfies (iii) in Theorem 3.5. We sometimes say λ-uniformly if we wish to keep track of the constant.

A property P (of Banach spaces) is called a super-property if super $-P\Leftrightarrow P$.

Corollary 3.6. *Let P be a non-universal super-property, meaning that there is at least one Banach space failing it. Then a Banach space with property P cannot contain ℓ_∞^n's uniformly.*

The reader will find background on ultrafilters, ultraproducts and ultrapowers in the appendix to this chapter.

Proposition 3.7. *Let X,Y be Banach spaces. Then X is finitely representable in Y (in short X f.r. Y) iff X embeds isometrically into an ultrapower of Y.*

Proof. Assume that X embeds isometrically into an ultrapower $Y^I/\mathcal{U}$ of Y. By Lemma 3.48, for any Y, $Y^I/\mathcal{U}$, and hence a fortiori X, is f.r. in Y, proving the "if" part. Conversely assume X f.r. in Y. Let I be the set of pairs (E,ε) where $E\subset Y$ is a finite dimensional subspace and $\varepsilon>0$. We equip I with the order defined by $i = (E_1, \varepsilon_1) \leq j = (E_2, \varepsilon_1)$ if $E_1 \subset E_2$ and $\varepsilon_2 < \varepsilon_1$. Note that obviously for any x in X there is $i = (E, \varepsilon)$ in I such that $x \in E$. Since X f.r. Y, for any $i = (E, \varepsilon)$ there is a linear map $u_i\colon\ E \to Y$ such that

$$\forall x \in E \qquad \|x\| \leq \|u_i(x)\| \leq (1+\varepsilon)\|x\|. \tag{3.4}$$

Then we define $u\colon\ X \to Y^I/\mathcal{U}$ as follows: for any x in X we set $u(x) = \overbrace{u_i(x_i))_{i\in I}}^{\cdot}$ where $x_{(E,\varepsilon)} = x$ whenever $x \in E$ and (say) $x_{(E,\varepsilon)} = 0$ if $x \notin E$. By the observation after (3.48), this indeed defines a *linear* map $u\colon\ X \to Y^I/\mathcal{U}$. Let ε_i denote the second coordinate of i so that $i = (E, \varepsilon_i)$. Note that $\lim \varepsilon_i = 0$ and hence $\lim_{\mathcal{U}} \varepsilon_i = 0$. Therefore, by (3.48) and (3.4) for any x in X

$$\|ux\| = \lim_{\mathcal{U}} \|u_i(x)\| = \|x\|.$$

This shows that u is an isometric embedding of X into $Y^I/\mathcal{U}$. $\square$

The following is an immediate consequence of Proposition 3.7:

Proposition 3.8. *Let P be a Banach space property. A Banach space B has super-P iff any space isometric to a subspace of an ultrapower of B has P.*

Proposition 3.9. *Let P be a Banach space property that is stable under isomorphism (for example reflexivity). Then super-P is also stable under isomorphisms.*

Proof. Indeed, if $B_1 \simeq B$ (isomorphically) then, for any $(I, \mathcal{U})$, we have obviously $B_1^I/\mathcal{U} \simeq B^I/\mathcal{U}$ (isomorphically). By Proposition 3.8, if B has super-P then any subspace of $B^I/\mathcal{U}$ has P, and hence (by the stability under isomorphism) any subspace of $B_1^I/\mathcal{U}$ has P, so that B_1 has super-P. $\square$

3.2 Super-reflexivity and inequalities for basic sequences

We will make crucial use of the following beautiful theorem due to V. Ptak [239]. This was later rediscovered by several authors, among which R.C. James who made an extremely deep contribution ([163, 164, 165]) to the subject of reflexivity and weak compactness.

Theorem 3.10. *The following properties of a Banach space B are equivalent:*

(i) *B is not reflexive.*

(ii) *For any $0 < \theta < 1$, there is a sequence $(x_n, \xi_n)_{n\geq 1}$ in $B \times B^*$ with $\|x_n\| \leq 1$, $\|\xi_n\| \leq 1$ for all n such that*

$$\xi_j(x_i) = 0 \qquad \forall i < j \tag{3.5}$$

$$\xi_j(x_i) = \theta \qquad \forall i \geq j. \tag{3.6}$$

(ii)′ *For some* $0<\theta<1$*, there is* $(x_n,\xi_n)_{n\ge 1}$ *as in (ii).*

(iii) *For any* $0<\theta<1$*, there is a sequence* (x_n) *in* B *such that for any finitely supported scalar sequence* (α_n) *we have*

$$\theta\sup_j\left|\sum_{i\ge j}\alpha_i\right|\le\left\|\sum\alpha_i x_i\right\|\le\sum|\alpha_i|. \tag{3.7}$$

(iii)′ *For some* $0<\theta<1$*, the same as (iii) holds.*

(iv) *For any* $0<\theta<1$*, there is a sequence* (y_n) *in* B *such that for any finitely supported scalar sequence* (β_n) *we have*

$$\theta\sup_n|\beta_n|\le\left\|\sum\beta_n y_n\right\|\le\sum_{n\ge 0}|\beta_n-\beta_{n+1}|. \tag{3.8}$$

(iv)′ *For some* $0<\theta<1$*, the same as (iv) holds.*

(v) *The inclusion mapping* $v_1\to\ell_\infty$ *(where* v_1 *denotes the space of scalar sequences* (β_n) *with* $\sum_{n\ge 0}|\beta_n-\beta_{n+1}|<\infty$*) factors through* B*.*

Proof. (ii) $\Rightarrow$ (ii)′ is trivial and (ii)′ $\Rightarrow$ (i) is easy. Indeed, if (ii)′ holds and if x^{**} is a $\sigma(B^{**},B^*)$ cluster point of (x_n), we must have $\xi_j(x^{**})=\theta$ by (3.6). Let $\xi\in B^*$ be a $\sigma(B^*,B)$ cluster point of (ξ_n). Then by (3.5) we must have $\xi(x_i)=0$. If $x^{**}\in B$, on one hand this implies $\xi(x^{**})=0$ but on the other hand $\xi_j(x^{**})=\theta$ implies $\xi(x^{**})=\theta$. This contradiction shows that $x^{**}\notin B$ and hence that B is not reflexive.

The main point is to show (i) $\Rightarrow$ (ii). Assume (i). Fix $0<\theta<1$ and $\varepsilon>0$. Pick $x^{**}\in B^{**}$ with $\|x^{**}\|=1$ such that $\mathrm{dist}(x^{**},B)>\theta$. (Obviously, such an x^{**} must exist, otherwise a simple iteration argument would show that $B^{**}=B$.)

Since $\|x^{**}\|=1$, there is ξ_1 in B^* with $\|\xi_1\|\le 1$ such that $x^{**}(\xi_1)=\theta$. Hence (see Lemma 3.4), for any $\varepsilon>0$, there is x_1 in B with $\|x_1\|\le 1+\varepsilon$ such that $x_1(\xi_1)=\theta$. We will now prove by induction the existence of a sequence as in (ii) except that we will find $\|x_n\|\le 1+\varepsilon$, but a posteriori we may renormalize (x_n), so this is unimportant.

Let E_1 be the subspace spanned by $\{x_1\}$. Since $\mathrm{dist}(x^{**},E_1)>\theta$ (and since $B^{**}/E_1=(B/E_1)^{**}=(E_1^{\perp})^*)$, there is ξ_2 in B^* with $\|\xi_2\|\le 1$ such that $\xi_2\in E_1^{\perp}$ and $x^{**}(\xi_2)=\theta$. Then, by Lemma 3.4, there is x_2 in B with $\|x_2\|\le 1+\varepsilon$ such that $x_2(\xi_1)=\theta$ and $x_2(\xi_2)=\theta$ and so on. To check the induction step, assume we have constructed $(x_1,\cdots,x_n)$, $(\xi_1,\cdots,\xi_n)$ satisfying (3.5) and (3.6). Let $E_n=\mathrm{span}\{x_1,\ldots,x_n\}$, we find $\xi_{n+1}\in E_n^{\perp}$ with $\|\xi_{n+1}\|\le 1$ such that $x^{**}(\xi_{n+1})=\theta$, and (using Lemma 3.4) we find x_{n+1} in B with $\|x_{n+1}\|\le 1+\varepsilon$ such that $x_{n+1}(\xi_i)=\theta\ \forall i\le n+1$. This completes the induction step and also the proof that (i) implies (ii).

It is an easy exercise to see that (ii) $\Leftrightarrow$ (iii) and (ii)′ $\Leftrightarrow$ (iii)′.

The equivalences (iii)$\Leftrightarrow$ (iv) and (iii)$'$ $\Leftrightarrow$ (iv)$'$ are obvious : just note the identity ("Abel summation") $\sum \alpha_i x_i = \sum \beta_n y_n$ where $y_0 = x_0$ and $y_n = x_n - x_{n-1}$ (or equivalently $x_n = y_0 + \cdots + y_n$), $\alpha_n = \beta_n - \beta_{n+1}$ (or equivalently $\beta_n = \sum_{i \geq n} \alpha_i$).

Lastly, (iv)$\Leftrightarrow$ (v) is easy: (iv) can be interpreted as a factorization $v_1 \to Y \to \ell_\infty$ of the inclusion $v_1 \to \ell_\infty$ through the closed span Y of (y_n) but using Hahn-Banach extensions of the functionals $\sum \beta_n y_n \mapsto \beta_n$ we can extend the second map $Y \to \ell_\infty$ to one from B to ℓ_∞, and this gives the factorisation in (v). Conversely if (v) holds i.e. we have a factorisation $v_1 \to B \to \ell_\infty$ (with bounded maps) then (iv)$'$ is immediate. □

Theorem 3.11. *The super-RNP is equivalent to super-reflexivity.*

Proof. From reflexive $\Rightarrow$ RNP, we deduce trivially super-reflexive $\Rightarrow$ super-RNP. To show the converse, it suffices obviously to prove that super RNP $\Rightarrow$ reflexive. Equivalently it suffices to show that if B is a non-reflexive space then there is a space X that is f.r. in B failing the RNP. Assume B non-reflexive. Then by the preceding Theorem there is a sequence (x_n) in B such that for any finitely supported scalar sequence (α_n) we have $\xi_j(\sum \alpha_i x_i) = \theta \sum\limits_{i \geq j} \alpha_i$, hence

$$\theta \sup_j \left| \sum_{i \geq j} \alpha_i \right| \leq \left\| \sum \alpha_i x_i \right\| \leq \sum |\alpha_i|. \tag{3.9}$$

We will now construct a space X that will be f.r. in B and will contain a $\theta/2$-separated dyadic tree, and hence will fail the RNP. The space X will be defined as the completion of L_1 with respect to the norm $|||\cdot|||$ defined below.

The underlying model for the construction is this: When (x_i) is the canonical basis of ℓ_1 (which satisfies (3.9) with $\theta = 1$) then the construction produces L_1 as the space X.

Let $(\mathcal{A}_n)$ be the dyadic filtration in $L_1 = L_1([0,1])$. For any f in L_1 we introduce the semi-norm

$$\|f\|_{(n)} = \left\| \sum_{0 \leq k < 2^n} \int\limits_{k2^{-n}}^{(k+1)2^{-n}} f(t)dt \cdot x_k \right\|.$$

Let $\mathcal{U}$ be a nontrivial ultrafilter on $\mathbb{N}$, i.e. an ultrafilter adapted to $\mathbb{N}$ (see Appendix 1). We set

$$|||f||| = \lim_{n,\mathcal{U}} \|f\|_{(n)}.$$

We have by (3.9) for all f in L_1

$$\theta \sup_{0 \leq s \leq 1} \left| \int_s^1 f(t)dt \right| \leq |||f||| \leq \int |f(t)| \, dt \tag{3.10}$$

Indeed, (3.9) implies this on the left side with the supremum over s of the form $s = k2^{-n}$, hence (3.10) follows by continuity of $s \to \int_s^1 f(t)dt$.

Let X be the completion of $(L_1, |||\cdot|||)$. By a routine argument one can check that this space X embeds in an ultraproduct of copies of B and hence is f.r. in B. By Lemma 3.12 below, the unit ball of X contains an infinite $\theta/2$-separated dyadic tree and hence fails the RNP. □

Lemma 3.12. *Let X be a Banach space. Assume that there is a linear map $J\colon L_1([0,1], dt) \to X$ such that for some $\theta > 0$ we have for all f in L_1*

$$\theta \sup_{0\le s\le 1} \left| \int_s^1 f(t)dt \right| \le \|J(f)\| \le \int_0^1 |f(t)|dt.$$

Then the unit ball of X contains a $\theta/2$-separated dyadic tree and hence X fails the RNP.

Proof. Fix n. To any $(\varepsilon_1, \dots, \varepsilon_n)$ in $\{-1,1\}^n$ we associate the interval $I(\varepsilon_1, \dots, \varepsilon_n)$ defined by induction as follows we set $I(1) = [0, \frac{1}{2}[$, $I(-1) = [\frac{1}{2}, 1]$ and if $I(\varepsilon_1, \dots, \varepsilon_n)$ is given we define $I(\varepsilon_1, \dots, \varepsilon_n, +1)$ as the left half of $I(\varepsilon_1, \dots, \varepsilon_n)$ and $I(\varepsilon_1, \dots, \varepsilon_n, -1)$ as its right half.

Note that $|I(\varepsilon_1, \dots, \varepsilon_n)| = 2^{-n}$. Let then $\Omega = \{-1,1\}^{\mathbb{N}}$. Let $(M_n)_{n\ge 0}$ be the L_1-valued martingale defined for $\varepsilon = (\varepsilon_n)_n \in \Omega$ by $M_0 \equiv 1$ and

$$M_n(\varepsilon) = 2^n \cdot 1_{I(\varepsilon_1,\dots,\varepsilon_n)}.$$

Note that $\|M_n(\varepsilon)\|_{L^1} = 1$ for all ε in Ω and since

$$M_n(\varepsilon) - M_{n-1}(\varepsilon) = 2^{n-1}\varepsilon_n(1_{I(\varepsilon_1,\dots,\varepsilon_{n-1},1)} - 1_{I(\varepsilon_1,\dots,\varepsilon_{n-1},-1)})$$

for all $n \ge 1$ we have

$$\sup_s \left| \int_s^1 (M_n - M_{n-1})(t)\ dt \right| \ge 1/2.$$

Hence the martingale $(J(M_n(\cdot)))$ is a B-valued $\theta/2$-separated dyadic martingale with range in the unit ball of B. □

Remark 3.13. The proof of Theorem 3.11 shows that, if B is not reflexive, then, for any $\theta < 1$ there is a space X f.r. in B satisfying the condition in Lemma 3.12. In the case of real valued scalars, this will be refined in (3.36) below, but the proof of this improvement is much more delicate.

Remark 3.14. By [89] (see also [82]), there are Banach spaces without RNP that do *not* contain any δ-separated infinite *dyadic* tree, whatever $\delta > 0$ may be. This gives an example of a space X failing the RNP but also failing the assumption of Lemma 3.12.

Definition. Fix a number $\lambda \ge 1$. A finite sequence $\{x_1, \dots, x_N\}$ in a Banach space B is called λ-basic if for any N-tuple of scalars $(\alpha_1, \dots, \alpha_N)$ we have

$$\sup_{1\le n\le N} \left\| \sum_{j=1}^n \alpha_j x_j \right\| \le \lambda \left\| \sum\nolimits_1^N \alpha_j x_j \right\|. \tag{3.11}$$

An infinite sequence (x_n) is called λ-basic if $\{x_1, \ldots, x_N\}$ is λ-basic for all $N \geq 1$. The case $\lambda = 1$ has already been distinguished in the preceding chapter: 1-basic sequences are called "monotone basic" sequences.

Note that (3.11) trivially implies by the triangle inequality

$$\sup |\alpha_j| \|x_j\| \leq 2\lambda \left\| \sum \alpha_j x_j \right\| \tag{3.12}$$

If $\overline{\text{span}}[x_n] = B$, the sequence $\{x_n\}$ is said to be a basis (sometimes called a Schauder basis) of B. Then any x in B has a *unique* representation as the sum of a convergent series $\sum_1^\infty \alpha_j x_j$ with uniquely determined scalar coefficients. Conversely any sequence (x_n) with this property must be λ-basic for some $\lambda \geq 1$ by the classical Banach–Steinhaus principle. Indeed, this property implies that there are biorthogonal functionals x_n^* in B^* such that any b in B can be written as $b = \sum_1^\infty x_n^*(b) x_n$. Let $P_N(b) = \sum_1^N x_n^*(b) x_n$ so that, for any b in B, $P_N(b) \to b$ and hence $\sup_N \|P_N(b)\| < \infty$. By the Banach–Steinhaus principle, we must have $\sup\limits_N \|P_N\| < \infty$, so that (x_n) is λ-basic with $\lambda = \sup_N \|P_N\|$.

Obviously, a λ-basic sequence is a basis for the closed subspace it spans. This justifies the term "basic."

The natural basis of ℓ_p $(1 \leq p < \infty)$ or c_0 is of course a basis in the above sense. Let B be any Banach space. In the sequel we will use repeatedly the observation that a sequence of martingale differences (df_n) in $L_p(B)$ is a monotone (i.e. λ-basic with $\lambda = 1$) basic sequence in $L_p(B)$.

Definition 3.15. A basis (x_n) is called boundedly complete if for any scalar sequence (α_n) such that $\sup_N \| \sum_1^N \alpha_n x_n \| < \infty$ the sum $S_N = \sum_1^N \alpha_n x_n$ converges in B.

Note that if $S_N \to b$ we have automatically $x_n^*(S_N) \to x_n^*(b)$ for each n and hence $\alpha_n = x_n^*(b)$ for all n. Let $P_N \colon B \to B$ be, as above, the projection defined by $P_N(b) = \sum_1^N x_n^*(b) x_n$.

Definition 3.16. A basis (x_n) is called shrinking if for any x^* in B^* we have $\|x^* - P_N^* x^*\| \to 0$. Equivalently, this means that the biorthogonal functionals (x_n^*) form a basis in B^*.

The following classical theorem due to R.C. James characterizes reflexive Banach spaces with a basis.

Theorem 3.17. *Let B be a Banach space with a basis (x_n). Then B is reflexive iff (x_n) is both boundedly complete and shrinking.*

Proof. We may assume that (x_n) is λ-basic for some $\lambda \geq 1$. Assume that B is reflexive. Let $S_N = \sum_1^N \alpha_n x_n$. If $\{S_N\}$ is bounded, by weak compactness of the closed balls, there is a subsequence weakly converging to a limit b in B. Then, for any fixed n, $x_n^*(S_N) \to x_n^*(b)$ (along a subsequence), but $\alpha_n = x_n^*(S_N)$ for all $N > n$, therefore $\alpha_n = x_n^*(b)$ for any n and hence (see the remarks preceding

Definition 3.15) $S_N = P_N(b)$ tends to b when $N \to \infty$. This shows that (x_n) is boundedly complete. Since $P_N(b) \to b$ for any b in B, we have $x^*(P_N(b)) \to x^*(b)$ for any x^* in B^* and hence $P_N^* x^* \to x^*$ with respect to $\sigma(B^*, B)$. If B is reflexive, $\sigma(B^*, B) = \sigma(B^*, B^{**})$ is the weak topology on B^*, and hence by Mazur's theorem x^* lies in the norm closure of $\operatorname{conv}\{P_N^* x^* \mid N \geq 1\}$. This yields: $\forall \varepsilon > 0\ \exists m\ \exists \xi \in \operatorname{conv}\{P_N^* x^* \mid 1 \leq N \leq m\}$ with $\|\xi - x^*\| < \varepsilon$. Clearly $P_m^* \xi = \xi$ (since $P_N P_m = P_m P_N = P_N\ \forall N \leq m$) we have

$$\|(1 - P_m^*)(x^*)\| \leq \|(1 - P_m^*)(x^* - \xi)\| \leq (1 + \lambda)\varepsilon$$

and hence we conclude that (x_n) is shrinking. Conversely, assume that (x_n) is boundedly complete and shrinking. Consider x^{**} in B^{**}. We can write $P_N^{**}(x^{**}) = \sum_1^N x^{**}(x_n^*) x_n$. We have

$$\sup \left\| \sum\nolimits_1^N x^{**}(x_n^*) x_n \right\| \leq \sup \|P_N\| \leq \lambda.$$

Since (x_n) is assumed boundedly complete, $\sum_1^N x^{**}(x_n^*) x_n$ converges to an element b in B. But now for any fixed n we have

$$x^{**}(x_n^*) = x_n^* \left(\sum\nolimits_1^N x^{**}(x_n^*) x_n \right) \to x_n^*(b) \text{ when } N \to \infty$$

and hence $x^{**}(x_n^*) = b(x_n^*)$ for any n. Finally, if (x_n) is assumed shrinking, $\{x_n^*\}$ is norm total in B^*, so this last equality implies $x^{**}(x^*) = b(x^*)$ for any x^* in B^* which means $x^{**} = b$. Thus we conclude that B is reflexive. □

Remark 3.18. Let $p > 1$ (resp. $q < \infty$). Let (e_n) be a basic sequence in a Banach space B. We say that (e_n) satisfies an upper p-estimate (resp. a lower q-estimate) if there is a constant C such that for any finite sequence $x_1, \ldots, x_N$ of disjoint consecutive (finite) blocks on (e_n) we have

$$\left\| \sum x_j \right\| \leq C \left(\sum \|x_j\|^p \right)^{1/p} \left(\text{resp. } \left(\sum \|x_j\|^q \right)^{1/q} \leq C \left\| \sum x_j \right\| \right).$$

If this holds, then (e_n) is shrinking (resp. is boundedly complete).

Indeed, let P_N denote the projection onto $\operatorname{span}[e_0, \ldots, e_N]$. Consider $\xi \in B^*$. Dualizing our hypothesis we find that for any increasing sequence $0 = n(0) < n(1) < \ldots$ we have

$$\left(\sum \|(P_{n(k)} - P_{n(k-1)})^* \xi\|^{p'} \right)^{1/p'} \leq C\|\xi\|.$$

This implies $(P_{n(k)} - P_{n(k-1)})^* \xi \to 0$ when $k \to \infty$. But we may choose the sequence $n(k)$ inductively so that (say) $\|(P_{n(k+1)} - P_{n(k)})^* \xi\| > (1/2)\|(I - P_{n(k)})^* \xi\|$ so we conclude that $\|\xi - P_N^* \xi\| \to 0$ when $N \to \infty$. The boundedly complete case is similar. We leave the details to the reader. □

Remark 3.19. Fix $1 < p < \infty$. By Theorem 2.5, B has the RNP iff any martingale difference sequence in $L_p(B)$ is boundedly complete when viewed as a monotone basic sequence.

We will use two variants of Theorem 3.10 as follows.

Remark 3.20. If B is not reflexive then for any $\lambda > 1$ there is a sequence (x_n, ξ_n) satisfying (ii) in Theorem 3.10 but moreover such that the sequence (ξ_n) is λ-basic.
Indeed, choose numbers $1 < \lambda_n < \lambda$ such that $\prod \lambda_n < \lambda$. It is easy to modify the induction step to obtain this: at each step where we have produced $(x_j, \xi_j)_{j \le n}$ we can find a finite subset F_n of the unit ball of B such that for any ξ in $\mathrm{span}(\xi_1, \ldots, \xi_n)$ we have $\|\xi\| \le \lambda_n \sup\{|\xi(x)| \mid x \in F_n\}$. Suppose we have produced $x_1, \ldots, x_n$, $\xi_1, \ldots, \xi_n$. We then replace E_n by $\mathrm{span}(E_n, F_n)$ to find ξ_{n+1} in $E_n^\perp \cap F_n^\perp$ with otherwise the same properties, so we may continue and find x_{n+1} with $\|x_{n+1}\| \le 1 + \varepsilon$ such that $x_{n+1}(\xi_j) = \theta$ for all $j \le n+1$. The fact that $\xi_{n+1} \in F_n^\perp$ guarantees that for any $\xi \in \mathrm{span}[\xi_1, \ldots, \xi_n]$ and any scalar α, we have

$$\forall x \in F_n \qquad \xi(x) = (\xi + \alpha \xi_{n+1})(x)$$

and hence

$$\|\xi\| \le \lambda_n \sup_{x \in F_n} |\xi(x)| \le \lambda_n \|\xi + \alpha \xi_{n+1}\|. \tag{3.13}$$

Now if we choose our sequence $\lambda_1, \ldots, \lambda_n, \ldots$ as announced so that $\prod \lambda_n < \lambda$, we clearly deduce from (3.13) that (ξ_n) is λ-basic.

Remark 3.21. By an analogous refinement, if B is not reflexive then for any $\lambda > 4$ there is a sequence (x_n, ξ_n) satisfying (ii) in Theorem 3.10 but moreover such that the two sequences (x_n) and $(x_1, x_2 - x_1, x_3 - x_2, \ldots)$ are λ-basic. Let x^{**} be as in the proof of Theorem 3.10. Suppose given $x_1, \ldots, x_n$ and $E_n = \mathrm{span}(x_1, \ldots, x_n)$. Since $d(x^{**}, E_n) > \theta$, we have for any x in E_n and any scalar α

$$\theta|\alpha| \le \|x + \alpha x^{**}\|$$

and hence by the triangle inequality

$$\|x\| \le (1 + \theta^{-1}) \|x + \alpha x^{**}\|. \tag{3.14}$$

Let $\varepsilon > 0$. Let G_n be a finite subset of $B_{E_n^*}$ such that

$$\forall x \in E_n \qquad \|x\| \le (1 + \varepsilon) \sup\{|\xi(x)| \mid \xi \in G_n\}.$$

By (3.14), each ξ in $B_{E_n^*}$ (in particular each ξ in G_n) extends to a linear form $\hat{\xi}$ of norm $\le 1 + \theta^{-1}$ on the span of x^{**} and E_n that vanishes on x^{**}. Then we claim that $\hat{\xi}$ extends to $\tilde{\xi} \in B^*$ with $\|\tilde{\xi}\| \le (1 + \theta^{-1})(1 + \varepsilon)$. Indeed, since $\mathrm{span}[E_n, x^{**}] \subset B^{**}$, the Hahn–Banach theorem a priori gives us $\tilde{\xi}$ in B^{***}, extending $\hat{\xi}$ to the whole of B^{**}, but we can use Lemma 3.4, applied to B^* instead of B, to find $\tilde{\xi}$ in B^*. In any case, note that $\tilde{\xi}(x^{**}) = 0$.

Let $\widetilde{G}_n = \{\tilde{\xi} \mid \xi \in G_n\}$. Then in the induction step, we may select x_{n+1} so that $\xi(x_{n+1}) = 0$ for all ξ in $\widetilde{G}_1 \cup \cdots \cup \widetilde{G}_n$. This guarantees that, for any k, all the x_i's for $i > k$ vanish on $\widetilde{G}_k$, so for any x in E_k, say $x = \sum_1^k \alpha_j x_j$ we have

$$\begin{aligned}\|x\| &\le (1+\varepsilon)\sup\{|\xi(x)| \mid \xi \in \widetilde{G}_k\}\\ &\le (1+\varepsilon)\sup\left\{\left|\xi\left(x + \sum_{i>k}\alpha_i x_i\right)\right| \;\middle|\; \xi \in \widetilde{G}_k\right\}\\ &\le (1+\theta^{-1})(1+\varepsilon)^2\left\|x + \sum_{i>k}\alpha_i x_i\right\|.\end{aligned}$$

Thus, if we choose θ and ε so that $(1+\theta^{-1})(1+\varepsilon)^2 \le \lambda'$, we obtain (x_n) λ'-basic (so we can obtain it λ'-basic for any $\lambda' > 2$). In addition, we will show that the sequence (z_i) defined by $z_1 = x_1$ and $z_i = x_i - x_{i-1}$ is $\lambda'(1+\theta^{-1}(1+\varepsilon))$-basic (so we can obtain it λ-basic for any $\lambda > 4$). Indeed, consider scalars (β_i) and let

$$x = \sum\nolimits_1^j \beta_i z_i, \qquad y = \sum\nolimits_{j+1}^n \beta_i z_i.$$

Note that $x \in E_j$ but our problem is that $z_{j+1} = x_{j+1} - x_j$ involves x_j. We must show

$$\|x\| \le \lambda'(1+\theta^{-1}(1+\varepsilon))\|x+y\|.$$

We have by (3.9)

$$\theta|\beta_{j+1}| \le \|x+y\|, \tag{3.15}$$

and hence by the triangle inequality

$$\|x + (y + \beta_{j+1}x_j)\| \le \|x+y\| + |\beta_{j+1}|(1+\varepsilon) \le (1+\theta^{-1}(1+\varepsilon))\|x+y\|.$$

But now since (x_i) is λ'-basic and $y + \beta_{j+1}x_j$ is in the span of $\{x_{j+1}, x_{j+2}, \cdots\}$

$$\|x\| \le \lambda'\|x + (y + \beta_{j+1}x_j)\|$$

and hence the announced result

$$\|x\| \le \lambda'(1+\theta^{-1}(1+\varepsilon))\|x+y\|.$$

□

To state the next result it will be convenient to introduce two sequences of positive numbers attached to a Banach space B, as follows.

For each $n \ge 1$, we set

$$\mathrm{bio}_n(B) = \inf\left\{\sup_{i\le n}\left\|\sum_{j\le i} y_j\right\|_B \sup_{j\le n}\|\xi_j\|_{B^*}\right\}$$

where the infimum runs over all biorthogonal systems $(y_i, \xi_i)_{i\le n}$ in $B \times B^*$ (biorthogonal means here that $\xi_i(y_j) = 0$ if $i \ne j$ and $= 1$ if $i = j$).

Note that obviously $\mathrm{bio}_n(B) \le \mathrm{bio}_{n+1}(B)$ for all $n \ge 1$. Let $c = \sup \|\xi_i\|$. Replacing ξ_i by $c^{-1}\xi_i$ and y_i by cy_i we may assume $c = 1$. By the Hahn–Banach theorem, $(y_1, \dots, y_n)$ admits a biorthogonal system (ξ_j) with $\sup \|\xi_j\|_{B^*} \le 1$ iff for any scalar n-tuple $(\alpha_1, \dots, \alpha_n)$ we have

$$\sup |\alpha_j| \le \left\| \sum \alpha_j y_j \right\|. \tag{3.16}$$

Thus, we can equivalently define $\mathrm{bio}_n(B)$ as the infimum of $\sup\limits_{i\le n} \left\| \sum\limits_{j\le i} y_j \right\|$ over all (y_j) satisfying (3.16).

Equivalently, setting $x_i = \theta \sum_{j\le i} y_j$, we have

$$\mathrm{bio}_n(B) = \inf\{\theta^{-1}\} \tag{3.17}$$

where the infimum runs over all $\theta \le 1$ for which there is an n-tuple $(x_1, \dots, x_n)$ in B satisfying for any scalar n-tuple $(\alpha_1, \dots, \alpha_n)$

$$\theta \sup_j \left| \sum_{i\ge j} \alpha_i \right| \le \left\| \sum \alpha_i x_i \right\| \le \sum |\alpha_i|.$$

Note that (3.17) clearly shows that if a space X is f.r. in B then necessarily

$$\mathrm{bio}_n(B) \le \mathrm{bio}_n(X) \qquad \forall n \ge 1. \tag{3.18}$$

In particular, this shows by Theorem 3.3, that $\mathrm{bio}_n(B) \le \mathrm{bio}_n(B^{**})$. The converse is obvious: since $B \subset B^{**}$ we must have $\mathrm{bio}_n(B^{**}) \le \mathrm{bio}_n(B)$. Thus we obtain

$$\mathrm{bio}_n(B^{**}) = \mathrm{bio}_n(B) \qquad \forall n \ge 1. \tag{3.19}$$

Moreover, (3.17) also shows that for any quotient space, say B/S (with $S \subset B$ a closed subspace), we have

$$\mathrm{bio}_n(B) \le \mathrm{bio}_n(B/S). \tag{3.20}$$

Indeed, one verifies this by a trivial lifting of $x_i \in B/S$ up in B.

Note one more equivalent definition of $\mathrm{bio}_n(B)$:

$$\mathrm{bio}_n(B) = \inf\{\theta^{-1}\} \quad \text{where} \quad \theta \le 1 \quad \text{runs over} \tag{3.21}$$

all the numbers for which there is a n-tuple $(x_j, \xi_j)_{j\le n}$ in $B \times B^*$ with $\|x_j\| \le 1$, $\|\xi_j\| \le 1$ such that

$$\xi_j(x_i) = \theta \quad \forall i \ge j \quad \text{and} \quad \xi_j(x_i) = 0 \quad \forall i < j. \tag{3.22}$$

Indeed, if (y_i, ξ_i) is as in the original definition, if we set $x_i = \theta \sum_{j \le i} y_j$ with $\theta = \left(\sup_i \left\| \sum_{j\le i} y_j \right\| \right)^{-1}$ and $\sup \|\xi_i\| = 1$, we obtain (3.22). Conversely, given (3.22), if we set $y_j = \theta^{-1}(x_j - x_{j-1})$, $y_1 = \theta^{-1}x_1$ we find $\sup \left\| \sum_{j\le i} y_j \right\| \le \theta^{-1}$.

From (3.21), it is immediate (replacing (y_j, ξ_j) by (ξ_{n+1-j}, y_{n+1-j}), $1 \le j \le n$) that

$$\mathrm{bio}_n(B^*) \le \mathrm{bio}_n(B).$$

Hence also $\mathrm{bio}_n(B^{**}) \le \mathrm{bio}_n(B^*)$, and since we already saw $\mathrm{bio}_n(B^{**}) = \mathrm{bio}_n(B)$, we conclude that $\mathrm{bio}_n(B)$ is self-dual:

$$\mathrm{bio}_n(B) = \mathrm{bio}_n(B^*) \quad \forall n \ge 1. \tag{3.23}$$

We also introduce

$$t_n(B) = \inf\{\sup_\omega \|M_n(\omega)\|_B\}$$

where the infimum runs over all dyadic martingales $(M_k)_{k\ge0}$ such that $\|M_k(\omega) - M_{k-1}(\omega)\| \ge 1$ for all ω and all $1 \le k \le n$. Again we have obviously $t_n(B) \le t_{n+1}(B)$ for all n. Note that $t = \sup_n t_n(B) < \infty$ iff B contains for some $\delta > 0$ arbitrarily long δ-separated finite dyadic trees in its unit ball.

Again we have $t_n(B) \le t_n(X)$ if X f.r. B. Moreover, by an easy lifting argument, this also holds when X is isometric to a quotient of B.

Theorem 3.22. *The following properties of a Banach space B are equivalent:*

(i) *B is super-reflexive.*

(i)′ *B^* is super-reflexive.*

(ii) *$bio_n(B) \to \infty$ when $n \to \infty$.*

(iii) *$t_n(B) \to \infty$ when $n \to \infty$.*

(iv) *For any $\lambda > 1$, there is $q < \infty$ and a constant C such that for any N and any λ-basic sequence $(y_1, \dots, y_N)$ in B or in any quotient of B we have*

$$\left(\sum \|y_j\|^q \right)^{1/q} \le C \left\| \sum y_j \right\|. \tag{3.24}$$

(iv)′ *For some $\lambda > 1$, the same as (iv) holds.*

(v) *For any $\lambda > 1$, there is $p > 1$ and a constant C such that for any N and any λ-basic sequence $(y_1, \dots, y_N)$ in B we have*

$$\left\| \sum y_j \right\| \le C \left(\sum \|y_j\|^p \right)^{1/p}. \tag{3.25}$$

(v)′ *For some $\lambda > 1$, the same as (v) holds.*

Proof. The proofs that (i) $\Rightarrow$ (ii) or that (i) $\Rightarrow$ (iii) are similar to the proof of Theorem 3.11. Assume that $\mathrm{bio}_n(B)$ (resp. $t_n(B)$) remains bounded when $n \to \infty$. We will show that there is a space X (resp. Y) that is f.r. in B and that is not reflexive (resp. fails the RNP). This will show that (i) $\Rightarrow$ (ii) (resp. (i) $\Rightarrow$ (iii)). Let us outline the argument for (i) $\Rightarrow$ (ii). Assume that (ii) fails i.e. that $\mathrm{bio}_n(B) < C$ for all $n \geq 1$. Then, for each n we have (by homogeneity) a biorthogonal system $(y_i^n, \xi_i^n)_{i \leq n}$ such that $\sup_{i \leq n} \| \sum_{j \leq i} y_j^n \| \leq C$ and

$$\|\xi_i^n\| = 1 \quad \text{for} \quad i = 1, 2, \ldots, n.$$

We will define the Banach space X as the completion of $\mathbb{K}^{(\mathbb{N})}$ for the norm $\|\cdot\|_X$ defined as follows. For each n we set for any finitely supported scalar sequence $\alpha = (\alpha_k)$

$$\|\alpha\|_n = \left\| \sum_{1 \leq k \leq n} \alpha_k \sum_{j \leq k} y_j^n \right\|.$$

Then we fix a nontrivial ultrafilter $\mathcal{U}$ on $\mathbb{N}$ and we set:

$$\|\alpha\|_X = \lim_{n,\mathcal{U}} \|\alpha\|_n.$$

Let $x = \sum_{1 \leq k \leq n} \alpha_k \sum_{j \leq k} y_j^n$. We have clearly by biorthogonality $\xi_i^n(x) = \sum_{k \geq i} \alpha_k$ hence

$$\sup_{i \leq n} \left| \sum_{k \geq i} \alpha_k \right| \leq \|\alpha\|_n \leq C \sum |\alpha_i|$$

and hence

$$\sup_{i} \left| \sum_{k \geq i} \alpha_k \right| \leq \|\alpha\|_X \leq C \sum |\alpha_i|.$$

By (i) $\Leftrightarrow$ (iii) in Theorem 3.10 we see that X is not reflexive, but since X manifestly embeds in an ultraproduct of subspaces of B, X is f.r. in B. This completes the proof that (i) $\Rightarrow$ (ii).

The proof that (i) $\Rightarrow$ (iii) is similar: if $t_n(B) < C$ for all n we produce Y f.r. in B and containing in its unit ball an infinite δ-separated dyadic tree with $\delta = 1/C$ (see the proof of Theorem 3.11); we leave the details to the reader. Note that (ii) $\Rightarrow$ (i) follows from Theorem 3.10. Indeed, by the latter theorem if B is not reflexive $\mathrm{bio}_n(B)$ is bounded; therefore (ii) implies B reflexive. But by (3.18), if B satisfies (ii) then any X f.r. in B also satisfies (ii) and hence must be reflexive. This shows that (ii) $\Rightarrow$ (i).

Similarly, we have (iii) $\Rightarrow$ (i). Indeed, it suffices to show (iii) implies B reflexive. But if B is not reflexive, Remark 3.13 (and $t_n(B) \leq t_n(X)$ if X f.r. B) clearly shows that $t_n(B)$ remains bounded when $n \to \infty$; this shows (iii) $\Rightarrow$ (i). Thus we have proved (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$(iii), and hence by (3.23), (i) $\Leftrightarrow$ (i)′.

We will now show that (ii) $\Rightarrow$ (iv). Fix $\lambda > 1$. We will show that if (iv) fails for this λ then (ii) also fails. We will argue as we did in the preceding section for monotone basic sequences. Let $b(N, \lambda)$ be the smallest constant b such that for any λ-basic $(y_1, \ldots, y_N)$ in a quotient of B we have

$$\inf_{1 \le k \le N} \|y_k\| \le b \left\| \sum\nolimits_1^N y_k \right\|.$$

Clearly (see the proof of Theorem 4.9)

$$b(NK, \lambda) \le b(N, \lambda) b(K, \lambda) \quad \text{for all} \quad N, K, \tag{3.26}$$

and also $b(K, \lambda) \le \lambda b(N, \lambda)$ for any $K > N$. Therefore, if $b(N, \lambda) < 1$ for some $N > 1$ we find $r < \infty$ and C such that $b(N, \lambda) \le C N^{-1/r}$ for all N and this leads to (see the proof of (4.10))

$$\left(\sum \|y_j\|^q \right)^{1/q} \le C \left\| \sum y_j \right\| \quad \text{for} \quad q > r$$

and some constant C. This argument shows that if (iv) fails for some $\lambda > 1$ we must have $b(n, \lambda) \ge 1$ for all $n > 1$. Equivalently, for any $\varepsilon > 0$ there is $(y_1, \ldots, y_n)$ λ-basic in a quotient of B, say B/S for some subspace $S \subset B$, such that $\|\sum_1^n y_j\| \le 1 + \varepsilon$ but $\|y_j\| > 1$ for all $1 \le j \le n$. By (3.12), there are functionals (ξ_i) biorthogonal to y_j with $\|\xi_i\| \le 2\lambda$, and by (3.11) we have

$$\sup_i \left\| \sum_{j \le i} y_j \right\| \le \lambda \left\| \sum\nolimits_1^n y_j \right\| \le \lambda(1 + \varepsilon)$$

hence we obtain $\mathrm{bio}_n(B/S) \le \lambda(1+\varepsilon)2\lambda$, but by (3.20) we know that $\mathrm{bio}_n(B) \le \mathrm{bio}_n(B/S)$, therefore (ii) fails. This completes the proof that (ii) $\Rightarrow$ (iv).

We now show (ii) $\Rightarrow$ (v). Assume (ii). Then, as we already mentioned, by (3.23), B^* satisfies (ii) and hence, using the already proved implication (ii) $\Rightarrow$ (iv), B^* satisfies (iv), and actually all quotient spaces of B^* satisfy (iv). Then let $(x_1, \ldots, x_n)$ be λ-basic in B. Let $E = \mathrm{span}\{x_1, \ldots, x_n\}$. We have $E^* = B^*/E^\perp$ and the biorthogonal functionals $(x_1^*, \ldots, x_n^*)$ are λ-basic in E^*. By (iv) applied in E^*, we have for any scalar n-tuple (note that $(\alpha_i x_i^*)_{i \le n}$ is also basic if $\alpha_i \ne 0$)

$$\left(\sum |\alpha_i|^q \|x_i^*\|^q \right)^{1/q} \le C \left\| \sum \alpha_i x_i^* \right\|$$

hence, by duality, if $p > 1$ is conjugate to q we find

$$\begin{aligned}\left\|\sum x_i\right\| &= \sup\left\{\left|\left(\sum x_i\right)(x^*)\right| \;\middle|\; x^* \in E^*,\ \|x^*\| \le 1\right\} \\ &= \sup\left\{\left|\sum \alpha_i\right| \;\middle|\; \left\|\sum \alpha_i x_i^*\right\| \le 1\right\} \\ &\le \sup\left\{\left(\sum(|\alpha_i|\|x_i^*\|)^q\right)^{1/q}\left(\sum \|x_i^*\|^{-p}\right)^{1/p} \;\middle|\; \left\|\sum \alpha_i x_i^*\right\| \le 1\right\} \\ &\le C\left(\sum \|x_i^*\|^{-p}\right)^{1/p} \\ &\le C\left(\sum \|x_i\|^p\right)^{1/p}\end{aligned}$$

where for the last line we used $1 = x_i^*(x_i) \le \|x_i^*\|\|x_i\|$. This completes the proof that (ii) $\Rightarrow$ (v).

Note that (iv) $\Rightarrow$ (iv)$'$ and (v) $\Rightarrow$ (v)$'$ are trivial. Now, we prove (v)$'$ $\Rightarrow$ (i): Since (v)$'$ is clearly a super-property, it suffices to show (v)$'$ implies B reflexive. But if B is not reflexive, by Remark 3.20, for any $\lambda > 1$, we can find a λ-basic sequence (ξ_n) with $\|\xi_n\| \le 1$ satisfying (3.6) for some (x_n) in the unit ball of B. This implies

$$\theta n = \sum_{j\le n} \xi_j(x_n) \le \left\|\sum\nolimits_1^n \xi_j\right\|$$

but now (v)$'$ implies $\left\|\sum_1^n \xi_j\right\| \le Cn^{1/p}$ with $p > 1$ which is impossible when $n \to \infty$. This contradiction shows that (v)$'$ implies the reflexivity of B, concluding the proof of (v)$'$ $\Rightarrow$ (i).

It only remains to show (iv)$'$ $\Rightarrow$ (v)$'$. Since the finite dimensional subspaces of B^* are the duals of the finite dimensional quotients of B, by duality (iv)$'$ implies that B^* satisfies (v)$'$. Applying the (just proved) implication (v)$'$ $\Rightarrow$ (i) to the space B^*, we conclude that B^* must be super-reflexive, and hence (recall (3.23)) B itself satisfies (ii), and we already proved (ii) $\Rightarrow$ (v) $\Rightarrow$ (v)$'$. So we conclude (iv)$'$ $\Rightarrow$ (v)$'$. □

Remark. Returning to Remark 3.21, recall that the sequence (z_i) (defined by $z_1 = x_1$ and $z_i = x_i - x_{i-1}$ for $i > 1$) can be found λ-basic with $\lambda > 4$, and also $\|z_i\| \ge \xi_i(z_i) = \theta$. But then $\sum_1^n z_i = x_n$ hence $\left\|\sum_1^n z_i\right\| \le 1$, which contradicts any estimate of the form $\left(\sum \|z_i\|^q\right)^{1/q} \le C\left\|\sum z_i\right\|$. This shows that if B itself (without its quotients) satisfies (iv) then B is super-reflexive.

Corollary 3.23. *If B is super-reflexive then for any $\lambda > 1$ there are $p > 1$ and $q < \infty$ and positive constants C' and C'' such that any λ-basic sequence $(x_1, \ldots, x_N)$ is B satisfies*

$$(C')^{-1}\left(\sum \|x_i\|^q\right)^{1/q} \le \left\|\sum x_i\right\| \le C''\left(\sum \|x_i\|^p\right)^{1/p}.$$

Proof. This is immediate from Theorem 3.22 since we can replace B by B^* in (v). □

3.3 Uniformly non-square and J-convex spaces

We start this section by a remarkable result discovered by R.C. James [162].

Theorem 3.24. *In any non-reflexive Banach space B, there is, for any $\delta > 0$, a pair x, y in the unit sphere of B such that*

$$\|x \pm y\| \geq 2 - \delta.$$

Remark. Banach spaces that fail the conclusion of Theorem 3.24 are called uniformly non-square. More precisely, B is "uniformly non-square" if there is $\delta > 0$ such that for any x, y in the unit ball we have either $\|(x + y)/2\| \leq 1 - \delta$ or $\|(x - y)/2\| \leq 1 - \delta$. This is a weak form of uniform convexity. In fact, this is the same as saying that the uniform convexity modulus $\delta_B(\varepsilon)$ is > 0 for *some* $\varepsilon > 0$ (while uniform convexity is the same but for *all* $\varepsilon > 0$).

Remark. Let $\alpha, \beta \in \mathbb{R}$ such that $|\alpha| + |\beta| = 1$. Assume $\|x\|, \|y\| \leq 1$ and $\|x \pm y\| \geq 2 - \delta$. Then for some $\varepsilon = \pm 1$ we have

$$\begin{aligned}\|\alpha x + \beta y\| = \| \, |\alpha| x + \varepsilon |\beta| y\| &\geq \|x + \varepsilon y\| - \|(1 - |\alpha|)x + \varepsilon(1 - |\beta|)y\| \\ &\geq 2 - \delta - (1 - |\alpha| + 1 - |\beta|) = 1 - \delta.\end{aligned}$$

Therefore by homogeneity we have

$$\forall \alpha, \beta \in \mathbb{R} \qquad (1 - \delta)(|\alpha| + |\beta|) \leq \|\alpha x + \beta y\| \leq |\alpha| + |\beta|.$$

In particular, any non-reflexive Banach space contains for any $\delta > 0$ a 2-dimensional subspace $(1 + \delta)$-isometric to $\ell_1^{(2)}$.

In the real case, $\ell_1^{(2)}$ is the same (isometrically) as $\ell_\infty^{(2)}$. Explicitly: Given x, y as above, let $a = (x + y)/2$ and $b = (x - y)/2$. Then

$$\forall \alpha, \beta \in \mathbb{R} \qquad (1 - \delta) \max\{|\alpha|, |\beta|\} \leq \|\alpha a + \beta b\| \leq \max\{|\alpha|, |\beta|\}.$$

Note however that this is no longer valid in the complex case.

Thus we have

Corollary 3.25. *The 2-dimensional space $\ell_1^{(2)}$ (over the reals) is finitely representable in every non-reflexive real Banach space.*

Corollary 3.26. *Any uniformly non-square Banach space is super-reflexive.*

By Proposition 3.9 we can "automatically" strengthen the preceding statement:

Corollary 3.27. *Any Banach space isomorphic to a uniformly non-square one is super-reflexive.*

Naturally the question was raised whether $\ell_1^{(2)}$ could be replaced by $\ell_1^{(n)}$ for $n > 2$ in particular for $n = 3$, but, in a 1973 tour de force, James himself gave a counterexample ([166], see also [125, 168]). We will give different and simpler examples of the same kind in Chapter 7. In the positive direction, one can generalize Theorem 3.24 as follows. This is also due to James (see [162, 169]).

Theorem 3.28. *Let B be a non-reflexive space. Then for any $n \geq 1$ and any $\delta > 0$ there are $x_1, \ldots, x_n$ in the unit sphere of B such that for any choice of signs $\varepsilon_j = \pm 1$ where the $+$ signs all precede the $-$ signs (we call these "admissible" choices of signs) we have*

$$\|\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n\| \geq n - \delta.$$

More explicitly we have for any $j = 1, \ldots, n-1$,

$$\|x_1 + \cdots + x_j - x_{j+1} - \cdots - x_n\| \geq n - \delta \quad \text{and also} \quad \|x_1 + \cdots + x_n\| \geq n - \delta.$$

Definition. A Banach space B is called J-convex if there is an integer $n > 1$ and a number $\delta > 0$ such that for any $x_1, \ldots, x_n$ in the unit ball of B

$$\inf \|\sum \varepsilon_k x_k\| \leq n(1 - \delta)$$

where the infimum runs over all admissible choice of signs i.e. such that $\varepsilon_k = \pm 1$ and all the $+$ signs appear before the $-$ signs (if any).

Note that if B is J-convex then any space f.r. in B is automatically J-convex. Using this, Theorem 3.28 can then be rephrased as follows.

Corollary 3.29. *Any J-convex Banach space is reflexive (and actually super-reflexive).*

The next result will be deduced rather easily from this last one.

Corollary 3.30. *J-convexity and super-reflexivity are equivalent properties.*

Remark 3.31. The girth of the unit ball of a real Banach space B is the infimum of the lengths of centrally symmetric simple closed rectifiable curves on its surface. It is proved in [169] (see also [249]) that a Banach space is super-reflexive if and only if the girth of its unit ball is (strictly) more than 4. In sharp contrast, the girth of ℓ_1, c_0 or ℓ_∞ is equal to 4. This is closely connected to the fact that super-reflexivity is equivalent to J-convexity.

The original proofs of both Theorems 3.24 and 3.28 are rather delicate. We follow a simpler approach due to Brunel and Sucheston [110].

We will need the following notion.

Definition. A sequence $(\hat{x}_n)$ in a Banach space will be called subsymmetric if for any integer N, for any $(\alpha_1, \ldots, \alpha_N)$ in $\mathbb{R}^N$ and for any increasing sequence $n(1) < n(2) < \cdots < n(N)$ we have

$$\left\|\sum\nolimits_1^N \alpha_j \hat{x}_j\right\| = \left\|\sum\nolimits_1^N \alpha_j \hat{x}_{n(j)}\right\|.$$

The sequence $(\hat{x}_n)$ will be called "additive" if for any finite sequence of real scalars (α_j) and for any $m \geq 1$, the preceding term $\left\|\sum_1^N \alpha_j \hat{x}_j\right\|$ is equal to:

$$\left\|\alpha_1 \sum_{0<j\leq m} \hat{x}_j/m + \alpha_2 \sum_{m<j\leq 2m} \hat{x}_j/m + \cdots + \alpha_N \sum_{(N-1)m<j\leq Nm} \hat{x}_j/m\right\|.$$

We will also need

Lemma 3.32. *Let (x_n) be a subsymmetric sequence such that $x_1 \neq x_2$ in a Banach space B. Then the sequence $d_j = x_{2j-1} - x_{2j}$ $(j \geq 1)$ is an unconditional basic sequence with constant 2. More precisely, for any finitely supported sequence of scalars (α_j) and any subset $\beta \subset \mathbb{N}$ we have*

$$\left\| \sum\nolimits_{j \in \beta} \alpha_j d_j \right\| \leq \left\| \sum \alpha_j d_j \right\| \tag{3.27}$$

and hence

$$\sup_{\pm} \left\| \sum \pm \alpha_j d_j \right\| \leq 2 \left\| \sum \alpha_j d_j \right\| . \tag{3.28}$$

Proof. Clearly (3.27) implies (3.28) by considering the index sets β_+ and β_- where the sign is $+$ or $-$. By an elementary iteration, it suffices to prove (3.27) when β is the complement of a singleton $\{j\}$. Equivalently, it suffices to prove

$$\|\alpha_1 d_1 + \cdots + \widehat{\alpha_j d_j} + \cdots + a_N d_N\| \leq \left\| \sum \alpha_j d_j \right\|$$

where the hat marks the absence. But now by subsymmetry for any m and any $0 < p \leq m$

$$\left\| \sum \alpha_j d_j \right\| = \left\| \sum\nolimits_{k=1}^{j-1} \alpha_k d_k + \alpha_j D_{j+p} + \sum\nolimits_{k=j+1}^{\infty} \alpha_k d_{k+m} \right\| , \tag{3.29}$$

where $D_{j+p} = x_{2(j-1)+p} - x_{2(j-1)+p+1}$. Note $m^{-1}(D_{j+1} + \cdots + D_{j+m}) \to 0$ when $m \to \infty$ (telescoping sum). Averaging (3.29) over $0 < p \leq m$ and letting $m \to \infty$ we obtain (3.27). □

Notation: Consider a bounded function $f\colon\ I \times I \to \mathbb{R}$. For each fixed $k \in I$, we can define $\lim_{\mathcal{U}} f(k,i)$ but also $\lim_{\mathcal{U}} f(i,k)$ and of course these differ in general. To avoid ambiguity we will denote by $\lim\limits_{i\ \mathcal{U}} f(i,j)$ the limit (relative to i) when j is kept fixed, and we denote by $\lim\limits_{j\ \mathcal{U}} f(i,j)$ the limit (relative to j) when i is kept fixed. Similarly, given a function $f\colon\ I^N \to \mathbb{R}$ we can define the iterated limits

$$\lim_{i(1)\ \mathcal{U}} (\lim_{i(2)\ \mathcal{U}} \ldots (\lim_{i(N)\ \mathcal{U}} f(i(1), \ldots, i(N)) \ldots).$$

Lemma 3.33. *If B is non-reflexive then there is a subsymmetric and additive sequence (x_n) satisfying (3.7) for some $\theta > 0$ and such that the closed span of $[x_n]$ is f.r. in B.*

Proof. By Theorem 3.10, B contains a sequence (x_n) satisfying (3.7). Let (e_n) be the canonical basis in the space $\mathbb{K}^{(\mathbb{N})}$ of finitely supported sequences of scalars ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$). For any N and any (α_j) in $\mathbb{K}^N$ we define

$$\|\alpha_1 e_1 + \cdots + \alpha_N e_N\| = \lim_{i(1)\ \mathcal{U}} (\lim_{i(2)\ \mathcal{U}} \ldots (\lim_{i(N)\ \mathcal{U}} \|\alpha_1 x_{i(1)} + \cdots + \alpha_N x_{i(N)}\|) \ldots).$$

Let B_1 be the completion of $\mathbb{K}^{(\mathbb{N})}$ equipped with this norm. Since $\mathrm{span}[e_1, \ldots, e_N]$ is (by definition) a subspace of an N-times iteration of ultrapowers starting with one of B, it must be f.r. in B (see Lemma 3.48). Therefore B_1 itself is f.r. in B. Clearly, if we replace (x_n) by $(x_{i(1)}, x_{i(2)}, \ldots)$ with $i(1) < i(2) < \cdots$ then (3.7) remains valid, therefore (e_n) itself still satisfies (3.7). Lastly, it takes a moment of thought to check that (e_n) is subsymmetric. We will now modify (e_n) to obtain a sequence that is also additive. Consider again a finitely supported sequence of scalars $(\alpha_1, \ldots, \alpha_N, 0, 0, \ldots)$, we define

$$\|(\alpha_j)\|_{(m)} = \left\|\alpha_1 m^{-1} \sum_{j=1}^{m} e_j + \alpha_2 m^{-1} \sum_{j=m+1}^{2m} e_j + \cdots + \alpha_N m^{-1} \sum_{j=m(N-1)+1}^{Nm} e_j\right\|.$$

We claim that $\|(\alpha_j)\|_{(m)}$ converges when $m \to \infty$. By subsymmetry of (e_n) and the triangle inequality, we have obviously

$$\|(\alpha_j)\|_{(m)} \le \|(\alpha_j)\|_{(1)} = \left\|\sum \alpha_j e_j\right\|.$$

More generally, for any pair of integer k, m we have

$$\|(\alpha_j)\|_{(mk)} \le \|(\alpha_j)\|_{(m)}. \tag{3.30}$$

Thus for any $n \ge m$, dividing n by m we can write $n = mk + p$ with $p < m$ and we easily check (again by the triangle inequality) that

$$\|(\alpha_j)\|_{(n)} \le \frac{mk}{n}\|(\alpha_j)\|_{(mk)} + \frac{p}{n}\|(\alpha_j)\|_{(p)}.$$

This gives us by (3.30)

$$\forall m \ge 1 \qquad \overline{\lim_{n\to\infty}}\|(\alpha_j)\|_{(n)} \le \|(\alpha_j)\|_{(m)},$$

and hence $\overline{\lim}_{n\to\infty}\|(\alpha_j)\|_{(n)} = \inf_m \|(\alpha_j)\|_{(m)}$. This proves the announced claim. We now define a norm $|||\cdot|||$ on $\mathbb{R}^{(\mathbb{N})}$ by setting

$$|||(\alpha_j)||| = \lim_{m\to\infty}\|(\alpha_j)\|_{(m)}. \tag{3.31}$$

Let B_2 be the completion of $(\mathbb{K}^{(\mathbb{N})}, |||\cdot|||)$. Let us denote by $(\hat{x}_n)$ the basis (e_n) viewed as sitting in B_2. Then, an easy verification shows that $(\hat{x}_n)$ is subsymmetric and still satisfies (3.7). Moreover, using (3.30) it is easy to see that $(\hat{x}_n)$ is additive. Lastly, note that (3.31) implies that B_2 is f.r. in B_1 and a fortiori in B. □

Proof of Theorem 3.24. By Lemma 3.33, we may assume that B contains a subsymmetric additive sequence (x_n) satisfying (3.7) for some $\theta > 0$. The idea of the proof (going back to [162]) can be roughly outlined as follows: Consider two long sequences of coefficients equal to ± 1 as follows,

$$\begin{array}{ccccccccccccc} 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & \ldots & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & \ldots & 0 & 1 & 0 & -1 \end{array}$$

where the second sequence is obtained from the first one by a single right shift. Then if these represent x and y we have $\|x\| = \|y\|$ by subsymmetry and $\|x + y\| = 2\|x\|$ by additivity. But moreover (this needs more care) $x - y$ is very similar (up to 2 digits) to $2x$ so we also can get $\|x-y\| \simeq 2\|x\|$. More precisely, let (these depend on m but for the moment we deliberately keep m silent):

$$z_1 = x_1 - x_3 + x_5 - x_7 + \cdots + x_{4m-3} - x_{4m-1} \tag{3.32}$$

$$z_2 = x_2 - x_4 + x_6 - x_8 + \cdots + x_{4m-2} - x_{4m}. \tag{3.33}$$

Let $r(m) = \|z_1\|$. Note that $\|z_1\| = \|z_2\| = r(m)$. Observe that the sequence of signs appearing in $z_1 + z_2$ is $(+ + - - + + - - \cdots)$. Therefore, by additivity, we have

$$\|z_1 + z_2\| = 2\|z_1\| = 2r(m).$$

As for $z_1 - z_2$ the sequence of signs is

$$(+ - - + + - - \cdots - - +).$$

This is as before except for the first and last sign. From this we easily deduce

$$\|z_1 - z_2\| \geq 2\|z_1\| - \| - e_1 + e_2\| = 2r(m) - \|e_2 - e_1\|.$$

We then distinguish two cases.

Case 1. $r(m)$ is unbounded when $m \to \infty$. Let $x = z_1/\|z_1\|$ and $y = z_2/\|z_2\|$. We have $\|x+y\| = 2$ and $\|x-y\| \geq 2-\delta(m)$ where $\delta(m) = \|e_2-e_1\|r(m)^{-1} \to 0$ when $m \to \infty$, so the proof is complete in this case.

Case 2. $\sup_m r(m) < \infty$. By Lemma 3.32, we have

$$\sup_{\pm} \left\|\sum\nolimits_1^m \pm(x_{2j-1} - x_{2j})\right\| \leq 2 \left\|\sum\nolimits_1^m x_{2j-1} - x_{2j}\right\| = 2r(m).$$

Thus we find for any (α_j) in $\mathbb{R}^m$

$$\left\|\sum\nolimits_1^m \alpha_j(x_{2j-1} - x_{2j})\right\| \leq 2r(m) \sup|\alpha_j|.$$

Moreover by (3.27) we have for any j

$$|\alpha_j|\|x_1 - x_2\| \leq \left\|\sum \alpha_j(x_{2j-1} - x_{2j})\right\|$$

and hence $\sup|\alpha_j|\|x_1 - x_2\| \leq \|\sum \alpha_j(x_{2j-1} - x_{2j})\|$. Thus we conclude in this case that $\overline{\text{span}}[x_{2j-1} - x_{2j}]$ is isomorphic to c_0, and hence that B contains ℓ_∞^n's uniformly. By Theorem 3.5 any Banach space (in particular ℓ_1) is f.r. in B. Thus we obtain the desired conclusion in this case also. $\square$

Proof of Theorem 3.28. The idea is similar to that of the preceding proof. We keep the same notation. We define for $k = 1, 2, \ldots, n$

$$z_k = x_k - x_{n+k} + x_{2n+k} - x_{3n+k} + \cdots + x_{(2m-2)n+k} - x_{(2m-1)n+k}.$$

Note that again, for any $k = 1, \ldots, n$, we have by subsymmetry

$$\|z_k\| = \|x_1 - x_2 + \cdots + x_{2m-1} - x_{2m}\| = r(m),$$

and by additivity

$$\|z_1 + \cdots + z_n\| = n\|z_1\| = nr(m).$$

Consider now $z_1 + \cdots + z_j - (z_{j+1} + \cdots + z_n)$ with $1 \leq j < n$.

The ordered sequence of nonzero basis coefficients of that vector is

$$\overbrace{+\cdots+}^{j}\ \overbrace{-\cdots-}^{n}\ \overbrace{+\cdots+}^{n}\ \cdots\ \overbrace{-\cdots-}^{n}\ \overbrace{+\cdots+}^{n-j}$$

where in the middle we have $2m-1$ series of n equal signs. This implies by additivity that

$$\|z_1 + \cdots + z_j - (z_{j+1} + \cdots + z_n)\| \geq nr(m) - (n-j)\|e_1 - e_2\|.$$

We may assume that we are in case 1, i.e. $r(m) \to \infty$. Let then

$$z'_j = z_j/\|z_j\|.$$

We find $\|z'_1 + \cdots + z'_n\| = n$ and

$$\|z'_1 + \cdots + z'_j - (z'_{j+1} + \cdots + z'_n)\| \geq n - \delta'(m)$$

with $\delta'(m) = (n/m)\|e_1 - e_2\| \to 0$ when $m \to \infty$. □

Corollary 3.34. *Let B be a non-reflexive or merely a non-J-convex Banach space. Then there is a Banach space $\widetilde{B}$ f.r. in B that contains a sequence (x_n) such that (here we deliberately insist on real scalars)*

(3.34)

$$\forall (\alpha_j) \in \mathbb{R}^{(\mathbb{N})} \qquad \sup_j \left\{ \left|\sum_{i<j} \alpha_i\right| + \left|\sum_{i\geq j} \alpha_i\right| \right\} \leq \left\|\sum \alpha_i x_i\right\| \leq \sum |\alpha_i|.$$

Equivalently, there are ξ_j in $\widetilde{B}^*$ with $\|\xi_j\| \leq 1$ such that $\xi_j(x_i) = 1$ for all $i < j$ and $\xi_j(x_i) = -1$ for all $i \geq j$.

Proof. Choose a sequence δ_n tending to 0, say $\delta_n = 1/n$. By Theorem 3.28, we may assume that B is not J-convex, so that for any $n \geq 1$, there are $x_1^{(n)}, \ldots, x_n^{(n)}$ in the unit sphere of B such that

$$\left\|\sum\nolimits_1^n \varepsilon_j x_j^{(n)}\right\| \geq n - \delta_n$$

for all the admissible choices of signs. Note that this implies obviously

$$\forall k \leq n \qquad \left\|\sum\nolimits_1^k \varepsilon_j x_j^{(n)}\right\| \geq k - \delta_n. \tag{3.35}$$

For any $(\alpha_1, \alpha_2, \ldots, \alpha_N, 0, 0 \ldots)$ in $\mathbb{K}^{(\mathbb{N})}$ we define

$$\left|\left|\left|\sum \alpha_j e_j\right|\right|\right| = \lim_{n\mathcal{U}} \left\|\sum_1^n \alpha_j x_j^{(n)}\right\|.$$

Let $\widetilde{B}$ be the completion of $(\mathbb{K}^{(\mathbb{N})}, |||\cdot|||)$ and let $x_j = e_j$ viewed as an element of $\widetilde{B}$. By Lemma 3.48 we know that $\widetilde{B}$ is f.r. in B. Then by (3.35) we have $|||\sum_1^k \varepsilon_j x_j||| \geq k$ (and hence this is $= k$) for any k and any admissible choice of signs. For each n and $j \leq n$ let $\xi_j^{(n)}$ in the unit sphere of $\widetilde{B}^*$ be such that $\xi_j^{(n)}(x_1 + \cdots + x_j - x_{j+1} - \cdots - x_n) = n$. Clearly we must have $\xi_j^{(n)}(x_i) = 1$ for all $i \leq j$ and $= -1$ for all i such that $j < i \leq n$ (if any). Let ξ_j be a $\sigma(\widetilde{B}^*, \widetilde{B})$ cluster point of $\{\xi_j^{(n)} \mid n \geq 1\}$ (or let $\xi_j = \lim_{n\,\mathcal{U}} \xi_j^{(n)}$). Clearly, (ξ_j) satisfies the property in Corollary 3.34. Then (3.34) follows since (here we deliberately insist on *real* scalars)

$$\xi_n\left(\sum \alpha_j x_j\right) = \sum_{j\leq n} \alpha_j - \sum_{j>n} \alpha_j$$

and also

$$\xi_{N+1}\left(\sum_1^N \alpha_j x_j\right) = \sum_1^N \alpha_j. \qquad \square$$

Proof of Corollary 3.30. By Corollary 3.29 we know that J-convexity implies super-reflexivity, and Corollary 3.34 implies the converse. Indeed, the space $\widetilde{B}$ appearing in Corollary 3.34 satisfies (3.7) and hence is not reflexive. $\square$

Remark. We suspect that Corollary 3.34 fails in the complex case. More precisely, there might exist non-reflexive complex Banach spaces that do not contain almost isometric copies of the complex version of "squares", i.e. do not contain almost isometrically the space $\mathbb{C}^2$ equipped with the norm $\|(x, y)\| = |x| + |y|$.

Corollary 3.35. *Let B be a non-reflexive real Banach space. Then there is a space X f.r. in B admitting a linear map $J\colon\ L_1([0,1], dt; \mathbb{R}) \to X$ such that for any f (real valued) in L_1*

$$\sup_{0\leq s\leq 1} \left|\int_0^s f(f)dt\right| + \left|\int_s^1 f(t)dt\right| \leq \|J(f)\| \leq \int_0^1 |f(t)|dt. \tag{3.36}$$

Proof. Let (x_n) be the sequence in the preceding corollary and let X be the associated space via the construction described in the proof of Theorem 3.11. We clearly have the announced property. $\square$

Corollary 3.36. *Let B be a non-reflexive space. Then there is a space X f.r. in B such that there is a dyadic martingale (f_n) in $L_\infty(X)$ satisfying for all $n \geq 1$ and all $\omega \in \{-1, 1\}^n$*

$$\|f_n(\omega)\| \leq 1, \quad \text{but} \quad \|f_n(\omega) - f_{n-1}(\omega)\| = 1.$$

In addition, for all $n \geq 1$ and all $\omega \neq \omega' \in \{-1, 1\}^n$ we have

$$\|f_n(\omega) - f_n(\omega')\| = 2, \quad \|f_n(\omega) - f_{n-1}(\omega')\| = 2.$$

In particular, the unit ball of X contains a 1-separated infinite dyadic tree.

Proof. We just repeat the argument for Lemma 3.12. Then the stronger property (3.36) yields the announced result. □

See Remark 1.25 for concrete examples of infinite trees as described in the preceding statement.

3.4 Super-reflexivity and uniform convexity

The main result of this section is the following.

Theorem 3.37. *The following properties of a Banach space B are equivalent.*

(i) *B is super-reflexive.*

(ii) *There is an equivalent norm on B for which the associated modulus of uniform convexity δ satisfies for some $2 \le q < \infty$*

$$\inf_{0<\varepsilon\le 2} \delta(\varepsilon)/\varepsilon^q > 0.$$

(ii)' *There is an equivalent norm on B for which the associated modulus of uniform smoothness ρ satisfies for some $1 < p \le 2$*

$$\sup_{t>0} \rho(t)/t^p < \infty.$$

(iii) *B is isomorphic to a uniformly convex space.*

(iii)' *B is isomorphic to a uniformly smooth space.*

(iv) *B is isomorphic to a uniformly nonsquare space.*

The equivalence of (i),(iii),(iii)' and (iv) is a beautiful result due to Enflo [131]. As in the preceding chapter, we will follow the martingale inequality approach of [227] and prove directly that (i) ⇒ (ii) (or equivalently since super-reflexivity is self-dual (i) ⇒ (ii)').

The proof will use martingale inequalities in $L_2(B)$. So we first need to replace B by $L_s(B)$. This is the content of the next two statements.

Lemma 3.38. *Let $1 < s < \infty$. Then a Banach space B is J-convex iff there are $n > 1$ and $\alpha < 1$ such that for any $x_1, \ldots, x_n$ in B we have*

$$\left(n^{-1} \sum_{\xi\in A(n)} \left\|\sum \xi_j x_j\right\|^s\right)^{1/s} \le \alpha n^{\frac{1}{s'}} \left(\sum \|x_j\|^s\right)^{1/s}, \tag{3.37}$$

where $A(n) \subset \{-1,1\}^n$ is the subset formed of the n admissible choices of signs.

Proof. Assume B J-convex, so $\exists n\ \exists\delta>0$ such that $\forall x_1,\dots,x_n\in B$

$$\inf_{\xi\in A(n)}\left\|\sum\nolimits_1^n \xi_j x_j\right\|\le n(1-\delta)\max\|x_j\| \tag{3.38}$$

Fix $1<s<\infty$. We claim that there is $\delta'>0$ such that $\forall x_1,\dots,x_n\in B$

$$\inf_{\xi\in A(n)}\left\|\sum \xi_j x_j\right\|\le n^{1/s'}(1-\delta')\left(\sum\|x_j\|^s\right)^{1/s}. \tag{3.39}$$

Indeed, if not then $\exists x_1,\dots,x_n$ such that

$$(1-\delta')n^{1/s'}\left(\sum\|x_j\|^s\right)^{1/s}<\inf\left\|\sum\xi_j x_j\right\|\le\sum\|x_j\|. \tag{3.40}$$

Moreover we may assume by homogeneity $\sum\|x_j\|=n$. But (3.40) contains an approximate reverse Hölder inequality, so an elementary reasoning shows that (3.40) implies

$$\max\{|\|x_i\|-\|x_j\||\mid 1\le i,j\le n\}\le\varphi_n(\delta')$$

with $\varphi_n(\delta')\to 0$ when $\delta'\to 0$. Since $\sum_1^n\|x_j\|=n$, we obtain

$$\max\|x_j\|\le 1+\varphi_n(\delta')\quad\text{and}\quad\min\|x_j\|\ge\sup\|x_j\|-\varphi_n(\delta')\ge 1-\varphi_n(\delta').$$

But then (3.38) and (3.40) together imply

$$n(1-\delta')(1-\varphi_n(\delta'))<n(1-\delta)(1+\varphi_n(\delta')),$$

and here $\delta>0$ is fixed while δ' and $\varphi_n(\delta')$ tend to zero, so this is impossible. This establishes (3.39). Then we note that (3.39) trivially implies (3.37) with $\alpha=(n^{-1}((1-\delta')^s+(n-1))^{1/s}$ and $\delta'>0$ ensures $\alpha<1$.
Conversely if (3.37) holds then a fortiori $\inf_{\xi\in A(n)}\|\sum\xi_j x_j\|\le\alpha n\sup\|x_j\|$ and hence B is J-convex. $\square$

Proposition 3.39. *Let $1<s<\infty$ and let $(\Omega,\mathcal{A},\mu)$ be any measure space. If B is super-reflexive, then $L_s(\mu;B)$ also is.*

Proof. By Corollary 3.30 it suffices to show that B is J-convex iff $L_s(\mu;B)$ also is. By integration at the s-th power, it is clear that B satisfies (3.37) iff $L_s(\mu;B)$ also does. $\square$

Corollary 3.40. *Fix $1<s<\infty$. If B is super-reflexive, then there are $1<p\le 2\le q<\infty$ (a priori depending on s) and positive constants C and C' such that any B-valued martingale (f_n) satisfies:*

$$C^{-1}\left(\sum\nolimits_0^\infty\|df_n\|^q_{L_s(B)}\right)^{1/q}\le\sup\|f_n\|_{L_s(B)}\le C'\left(\sum\nolimits_0^\infty\|df_n\|^p_{L_s(B)}\right)^{1/p}. \tag{3.41}$$

Proof. By the preceding Proposition, we may apply Theorem 3.22 to $L_s(B)$. Note that martingale difference sequences are monotone basic sequences in $L_s(B)$. Thus the Corollary follows from (i) $\Rightarrow$ (iv) and (v) in Theorem 3.22. $\square$

We can outline the proof of Theorem 3.37 like this: if B is super-reflexive, so is $L_2(B)$, so that all monotone basic sequences in $L_2(B)$ satisfy a lower q-estimate of the form (3.24). Applying this to B-valued martingales we find that there is $q < \infty$ and a constant C such that all B-valued martingales $(f_n)_{n\geq 0}$ satisfy (recall the convention $df_0 = f_0$ and $df_k = f_k - f_{k-1}$ for all $k \geq 1$)

$$\forall N \geq 1 \qquad \left(\sum\nolimits_0^N \|df_n\|^q_{L_2(B)}\right)^{1/q} \leq C \left\|\sum\nolimits_0^N df_n\right\|_{L_2(B)}. \tag{3.42}$$

The technical problem that we solved in the preceding chapter is to pass from (3.42) to an inequality of the form (4.8). By Theorem 4.51 and Remark 4.58, (4.8) is equivalent to an estimate of the form

$$\forall N \geq 1 \qquad \left\|\left(\sum \|df_n\|^q\right)^{1/q}\right\|_2 \leq C' \left\|\sum\nolimits_0^N df_n\right\|_{L_2(B)}. \tag{3.43}$$

The difficulty here is that when $2 \leq q < \infty$ we have always

$$\left(\sum \|df_n\|^q_{L_2(B)}\right)^{1/q} \leq \left\|\left(\sum \|df_n\|^q\right)^{1/q}\right\|_2 \tag{3.44}$$

but not conversely! So the inequality we need appears significantly stronger than (3.42). However, in the context of martingales there are frequent situations where a priori weak inequalities actually imply stronger ones. The proof of Lemma 4.13 in the preceding chapter illustrates this principle.

Proof of Theorem 3.37. We first prove the equivalence of (i)-(iv). The implications (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) are trivial and (iv) $\Rightarrow$ (i) is Corollary 3.27. Thus it suffices to show (i) $\Rightarrow$ (ii). Assume (i). By Proposition 3.39 $L_2(B)$ is super-reflexive. By Theorem 3.22, there is a constant C and $s < \infty$ such that any finite martingale (f_n) in $L_2(B)$ satisfies (4.8). By Lemma 4.12 and Corollary 4.7 we obtain (ii).

We now turn to (ii)' and (iii)'. Note that B satisfies (ii)' (resp. (iii)') iff B^* satisfies (ii) (resp. (iii)) in Theorem 3.37. Thus since B is super-reflexive iff B^* also is (see Theorem 3.22) we can deduce (i) $\Leftrightarrow$ (ii)' $\Leftrightarrow$ (iii)' from the part of Theorem 3.37 that we just proved above. However, the reader will surely observe that a direct argument for the main point (i) $\Rightarrow$ (ii)' can alternatively be obtained by combining together (i) $\Rightarrow$ (v) in Theorem 3.22 applied to $L_2(B)$, Lemma 4.13 and Corollary 4.22. □

Corollary 3.41. *If a Banach space B is super-reflexive, there are $p > 1$ and $q < \infty$ and a single equivalent norm $|\cdot|$ satisfying both* (4.1) *and* (4.24) *for some constants $\delta, C > 0$.*

Proof. Assume first that B is a *complex* Banach space. Then by the preceding Proposition the complex interpolation method applied between the two norms appearing respectively in (ii) and (ii)' in Theorem 3.37 produces an interpolated norm (of course still equivalent to the original one) that satisfies the desired property. Indeed, if the first norm, say $\| \;\|_0$, is q-uniformly convex and the second one, say $\| \;\|_1$, is p-uniformly smooth, then, by the following Proposition 3.42, the interpolated norm $\| \;\|_\theta$, is both q_θ-uniformly convex and p_θ-uniformly smooth, where $q_\theta^{-1} = (1-\theta)q^{-1} + \theta\infty^{-1}$ and $p_\theta^{-1} = (1-\theta)1^{-1} + \theta p^{-1}$. If B is a real space, its complexification (e.g. $B(H, B)$ with $H = \mathbb{C}$ viewed as a two dimensional real Hilbert space) inherits the super-reflexivity of B. Indeed, by Proposition 3.39), it is isomorphic to a super-reflexive (real) space (namely the ℓ_2-sense direct sum $B \oplus B$) and hence (see Remark 3.2) it is super-reflexive as a complex space. Therefore the real case reduces to the complex one. □

Let $1 \le p \le 2 \le q \le \infty$. Recall that a Banach space is q-uniformly convex (resp. p-uniformly smooth) with constant C if for all x, y in B we have

$$\|(x+y)/2\|^q + C^{-q}\|(x-y)/2\|^q \le 2^{-1}(\|x\|^q + \|y\|^q) \tag{3.45}$$

$$(\text{resp. } 2^{-1}(\|x+y\|^p + \|x-y\|^p) \le \|x\|^p + C^p\|y\|^p).$$

Note that any Banach space is trivially 1-uniformly smooth and ∞-uniformly convex with constant 1. The next result describes the stability of these notions under complex interpolation.

Proposition 3.42. *Let (B_0, B_1) be a compatible couple of complex Banach spaces. Let $0 < \theta < 1$ and let $B_\theta = (B_0, B_1)_\theta$.*

(i) *Let $2 \le q_0, q_1 \le \infty$. If B_j is q_j-uniformly convex with constant C_j ($j = 0, 1$) then B_θ is q_θ-uniformly convex with constant $C_\theta = C_0^{1-\theta}C_1^\theta$ where $q_\theta^{-1} = (1-\theta)q_0^{-1} + \theta q_1^{-1}$.*

(ii) *Let $1 \le p_0, p_1 \le 2$. If B_j is p_j-uniformly smooth with constant C_j ($j = 0, 1$) then B_θ is p_θ-uniformly smooth with constant $C_\theta = C_0^{1-\theta}C_1^\theta$ where $p_\theta^{-1} = (1-\theta)p_0^{-1} + \theta p_1^{-1}$.*

Proof. Let $Y(q_j)$ denote the direct sum $B_j \oplus B_j$ equipped with the norm $\|(x, y)\| = ((\|x\|^{q_j} + \|y\|^{q_j})/2)^{1/q_j}$. Let $X(q_j)$ denote $B_j \oplus B_j$ equipped with the norm $\|(x, y)\|_{X(q_j)} = (\|x\|^{q_j} + C_j^{-q_j}\|y\|^{q_j})^{\frac{1}{q_j}}$.
By ??, we have both $(Y(q_0), Y(q_1))_\theta = Y(q_\theta)$ and $(X(q_0), X(q_1))_\theta = X(q_\theta)$ isometrically for any $1 \le q_0, q_1 \le \infty$.
Consider the operator T defined by

$$T(x, y) = \Big(\frac{x+y}{2}, \frac{x-y}{2}\Big).$$

Note that by our assumption in (i) we have $\|T\colon\ Y(q_j) \to X(q_j)\| \le 1$ both for $j = 0$ and $j = 1$. Therefore by the interpolation Theorem

$$\|T\colon\ Y(q_\theta) \to X(q_\theta)\| \le 1.$$

This proves (i). The proof of (ii) is similar (or can be deduced by duality). □

Remark.

Problem: If B is both isomorphic to a p-uniformly smooth space and isomorphic to a q-uniformly convex one, is B isomorphic to a space that is both q-uniformly convex and p-uniformly smooth?

Note that the interpolation argument in Corollary 3.41 yields a norm that is both q_θ-uniformly convex and p_θ-uniformly smooth but with "worse" values $q_\theta > q$ and $p_\theta < p$ and in such a way that $q_\theta \to \infty$ when $p_\theta \to p$ (and $p_\theta \to 1$ when $q_\theta \to q$).

We now return to the strong law of large numbers, this time for (Banach space valued) martingales.

Lemma 3.43. *Fix an integer $n \geq 1$. Let $\Omega = \{-1,1\}^n$, let $\varepsilon_k\colon\ \Omega \to \{-1,1\}$ denote as usual the k-th coordinate, let $\mathcal{A}_0 = \{\phi, \Omega\}$ be the trivial σ-algebra and let $\mathcal{A}_k = \sigma(\varepsilon_1, \ldots, \varepsilon_k)$ for $k = 1, 2, \ldots, n$. Fix an integer $n \geq 1$. The following properties of a finite dimensional Banach space B are equivalent:*

(i) *There is a B-valued martingale $(f_0, \ldots, f_n)$ adapted to $(\mathcal{A}_0, \ldots, \mathcal{A}_n)$ such that for all $1 \leq k \leq n$ and all $\omega \in \Omega$*

$$\|df_k(\omega)\| = 1 \quad \text{and} \quad \|f_n(\omega)\| = 1.$$

(ii) *There is a B^*-valued martingale $(g_0, \ldots, g_n)$ adapted to $(\mathcal{A}_0, \ldots, \mathcal{A}_n)$, with $g_0 = 0$ such that for all $1 \leq k \leq n$ and all $\omega \in \Omega$*

$$\|g_n(\omega)\| = n \quad \text{and} \quad \|dg_k(\omega)\| = 1.$$

Proof. We start by observing that for any B-valued dyadic martingale $(f_0, \ldots, f_n)$ on $(\mathcal{A}_0, \ldots, \mathcal{A}_n)$ we have (pointwise):

$$\|df_k(\omega)\| \leq \|f_k(\omega)\| \text{ for all } k = 1, \ldots, n.$$

Indeed since $df_k = \varepsilon_k \psi_{k-1}$ with ψ_{k-1} being $\mathcal{A}_{k-1}$-measurable we have $f_k(\omega) = f_{k-1}(\omega) \pm df_k(\omega)$ if $\varepsilon_k(\omega) = \pm 1$, so this observation follows from the triangle inequality. A fortiori, we have

$$\|df_k\|_{L_p(B)} \leq \|f_k\|_{L_p(B)} \leq \|f_n\|_{L_p(B)}.$$

Note however that this is special to the dyadic filtration, the general case requires an extra factor 2.

Assume (i). Since $1 \leq \|df_k(.)\|$ there is φ_k in the unit ball of $L_\infty(\mathcal{A}_n, B^*)$ such that $1 \leq \langle df_k(.), \varphi_k(.)\rangle$ and a fortiori $1 \leq \mathbb{E}\langle df_k, \varphi_k\rangle$.

Let $g_n = \sum_1^n (\mathbb{E}_k - \mathbb{E}_{k-1})(\varphi_k)$. We have

$$n \leq \sum\nolimits_1^n \mathbb{E}\langle df_k, \varphi_k\rangle = \mathbb{E}\langle f_n, g_n\rangle$$

and hence $n \leq \mathbb{E}\|g_n\|$, but since (by the preceding observation for $p = \infty$) $\|dg_k\|_{L_\infty(B^*)} = \|(\mathbb{E}_k - \mathbb{E}_{k-1})\varphi_k\|_{L_\infty(B^*)} \leq \|\varphi_k\|_{L_\infty(B^*)} \leq 1$ we have $\|g_n\|_{L_\infty(B^*)} \leq n$ and hence $\mathbb{E}\|g_n\| \geq n$ forces $\|g_n(\omega)\| = n$ for all ω. Similarly, since $1 \leq \mathbb{E}\langle df_k, \varphi_k\rangle = \mathbb{E}\langle df_k, dg_k\rangle \leq \mathbb{E}\|dg_k\|$, the fact that $\|dg_k\|_{L_\infty(B^*)} \leq 1$ forces $\|dg_k(\omega)\| = 1$ for all ω.

Conversely, assume (ii). Since $\mathbb{E}\|g_n\| \geq n$ there is f_n in the unit ball of $L_\infty(\mathcal{A}_n, B)$ such that $\mathbb{E}\langle f_n, g_n\rangle \geq n$, and hence $\sum_1^n \mathbb{E}\langle df_k, dg_k\rangle \geq n$. The latter implies $\sum_1^n \mathbb{E}\|df_k\| \geq n$ but (by the preceding observation again with $p = \infty$) we have

$$\|df_k\|_{L_\infty(B)} \leq 1 \quad \text{and hence} \quad \sum\nolimits_1^n \|df_k(\omega)\| \leq n$$

for all ω. It follows that $\|df_k(\omega)\| = 1$ for all $k = 1, \ldots, n$ and all ω. In addition, since we have $\|df_n(\omega)\| \leq \|f_n(\omega)\|$, we also obtain $\|f_n(\omega)\| = 1$ for all ω. □

Lemma 3.44. *If a Banach space B is not super-reflexive then for each $n \geq 1$ and any $0 < \theta < 1$ there is a B-valued martingale $(\tilde{g}_0, \ldots, \tilde{g}_n)$ adapted to $\mathcal{A}_0, \ldots, \mathcal{A}_n$ with $\tilde{g}_0 = 0$ such that*

$$\inf_{\omega \in \Omega} \|\tilde{g}_n(\omega)\| \geq \theta n \quad \text{and} \quad \sup_{1 \leq k \leq n} \sup_{\omega \in \Omega} \|d\tilde{g}_k(\omega)\| \leq 1.$$

Proof. By Theorem 3.22 we may assume that B^* is not super-reflexive. By Corollary 3.36, for each n there is a finite dimensional space E f.r. in B^* containing an E-valued martingale satisfying (i) in Lemma 3.43. Fix $\varepsilon > 0$. Since E is $(1+\varepsilon)$-isomorphic to a subspace of B^*, E^* is $(1+\varepsilon)$-isomorphic to a quotient of B. Thus, E^* contains the range of a martingale (g_n) satisfying (ii) in Lemma 3.43. We have $dg_k = \varepsilon_k \psi_{k-1}$ with ψ_{k-1} in the unit ball of $L_\infty(\mathcal{A}_{k-1}, E^*)$. Fix θ so that $0 < \theta < (1+\varepsilon)^{-1}$. Let U_B denote the unit ball of B. Let $Q\colon\ B^* \to E^*$ be a surjection of norm 1 such that $Q(U_B) \supset \theta U_{E^*}$. Then there is $\widetilde{\psi}_{k-1}$ in $L_\infty(\mathcal{A}_{k-1}, B)$ with $\|\widetilde{\psi}_{k-1}\|_{L_\infty(B)} \leq \theta^{-1}$ lifting ψ_{k-1}, i.e. such that $Q(\widetilde{\psi}_{k-1}) = \psi_{k-1}$. Let then $\tilde{g}_n = \theta \sum_1^n \varepsilon_k \widetilde{\psi}_{k-1}$. We have $\|df_k\|_{L_\infty(B)} = \theta \|\widetilde{\psi}_{k-1}\|_{L_\infty(B)} \leq 1$ and $Q(\tilde{g}_n) = \theta g_n$ therefore

$$\theta n = \|\theta g_n(\omega)\| \leq \|\tilde{g}_n(\omega)\|$$

for all ω in Ω. □

The strong law of large numbers yields one more characterization of super-reflexivity:

Theorem 3.45. *Fix $1 < s \leq \infty$. The following properties of a Banach space B are equivalent:*

(i) *B is super-reflexive.*

(ii) *For any martingale (f_n) in $L_s(B)$ such that $\sup_n \|df_n\|_{L_s(B)} < \infty$, we have $n^{-1} f_n \to 0$ almost surely.*

(iii) *For any dyadic B-valued martingale such that $\sup\limits_n \|df_n\|_{L_\infty(B)} < \infty$ we have $n^{-1} f_n \to 0$ almost surely.*

(iv) *For any dyadic B-valued martingale such that $\sup_n \|df_n\|_{L_\infty(B)} \leq 1$ we have $\limsup_{n\to\infty} n^{-1}\|f_n\| < 1$ almost surely.*

Proof. Assume (i). By Corollary 3.40 there is $p > 1$ and C such that (3.41) holds. If $\sup \|df_n\|_{L_s(B)} < \infty$, this implies that $\sum n^{-1} df_n$ converges in $L_s(B)$ and hence (cf. Theorem 1.14) almost surely. By a classical (elementary) lemma

due to Kronecker any sequence $\{x_n\}$ in B such that $\sum n^{-1}x_n$ converges must satisfy $n^{-1}\sum_1^n x_k \to 0$. Therefore we obtain (ii) and (ii) $\Rightarrow$ (iii)$\Rightarrow$ (iv) are trivial. Conversely, assume (iv). If B is not super-reflexive, we will construct a dyadic B-valued martingale $(F_n)_{n\geq 0}$ with $\|dF_n\|_{L_\infty(B)} \leq 1$ for all n and such that $\limsup_{n\to\infty} n^{-1}\|F_n\| = 1$ a.s., thus contradicting (iv). This shows that (iv) $\Rightarrow$ (i). We now turn to the announced construction:
Our basic building block will be this: Let $0 < \theta < 1$. By Lemma 3.44, for any N there is a dyadic martingale $g_1^{(N)}, \dots, g_N^{(N)}$ with $g_0^{(N)} = 0$ such that $\|dg_k^{(N)}\|_{L_\infty(B)} \leq 1$ and $\inf_\omega \|g_N^{(N)}(\omega)\| \geq \theta N$.

Now let $0 < \theta_n < 1$ and $\xi_n > 0$ be sequences such that

$$\lim_n \theta_n = 1 \quad \text{and} \quad \lim_n \xi_n = 0.$$

Let $N(1) < N(2) < \cdots < N(n) < \cdots$ be increasing sufficiently fast so that

$$\frac{N(1) + \cdots + N(n-1)}{N(n)} < \xi_n \quad \text{for all} \quad n \geq 1. \tag{3.46}$$

Let $S(n) = N(1) + \cdots + N(n)$. Let $g_1^{[n]}, \dots, g_{N(n)}^{[n]}$ (with $g_0^{[n]} = 0$) be the product of our basic building block, when we take $N = N(n)$ and $\theta = \theta_n$. We define a martingale $(F_{S(n)})_{n\geq 1}$ adapted to $(\mathcal{A}_{S(n)})$ as follows: we set $F_{S(1)} = g_{N(1)}^{[1]}$, then $F_{S(2)} - F_{S(1)} = g_{N(2)}^{[2]}(\varepsilon_{S(1)+1}, \dots, \varepsilon_{S(1)+N(2)}) \dots$ and

$$F_{S(n)} - F_{S(n-1)} = g_{N(n)}^{[n]}(\varepsilon_{S(n-1)+1}, \dots, \varepsilon_{S(n-1)+N(n)}). \tag{3.47}$$

Since $\mathbb{E}g_N^{(N)} = g_0^{(N)} = 0$ for all N, $(F_{S(n)})_{n\geq 0}$ is indeed a martingale adapted to $(\mathcal{A}_{S(n)})_{n\geq 1}$. For any $k \leq S(n)$ we set $F_k = \mathbb{E}^{\mathcal{A}_k}(F_{S(n)})$. Then $(F_k)_{k\geq 1}$ is a (dyadic) martingale adapted to $(\mathcal{A}_k)_{k\geq 1}$ and of course $F_k = F_{S(n)}$ if $k = S(n)$. Note that $\|F_k\|_{L_\infty(B)} \leq \sum_1^k \|dF_j\|_{L_\infty(B)} \leq k$ for all $k \geq 1$. We have by (3.47) for any ω

$$\|F_{S(n)}(\omega)\| \geq \theta_n N(n) - \|F_{S(n-1)}\| \geq \theta_n N(n) - S(n-1) \geq \theta_n S(n) - (1+\theta_n)S(n-1)$$

and hence by (3.46)

$$\|F_{S(n)}(\omega)\| \geq S(n)(\theta_n - (1+\theta_n)\xi_n).$$

Thus, since $\theta_n - (1+\theta_n)\xi_n \to 1$, for any ω

$$\limsup_{n\to\infty} \|n^{-1}F_n(\omega)\| \geq 1$$

as announced. □

3.5 Notes and Remarks

The notion of "finitely representable" and "super-property" are due to R.C. James [163]. The local reflexivity principle (Theorem 3.3) goes back to Lindenstrauss and Rosenthal [191]. As we mentioned in the text, Ptak's paper

[239] seems to be the earliest reference for Theorem 3.10 but it was independently proved (slightly later) by D. Milman and V. Milman and by James. The reformulations in terms of factorizations such as (v) in Theorem 3.10 were emphasized in Lindenstrauss and Pełczyński's influential paper [190].

Theorem 3.11 was stated in [227]. Theorems 3.17 and 3.22 are due to R.C. James as well as Corollary 3.23 and essentially all the results in §3.3. James first proved in [162] that uniformly non-square implies reflexive. In the same paper, he notes that the extension from pairs to triples of vectors leads to a proof that if $n = 3$ for any $\varepsilon > 0$ any J-(n, ε) convex space is reflexive. Later on in [170], the authors observe that the same proof works for any integer $n \geq 2$, thus showing that J-convex implies reflexive. Since J-convex is a super-property, this shows that J-convex implies super-reflexive. But the converse was an easy consequence of James early ideas on reflexivity. Therefore this yielded the equivalence of "J-convex" and "super-reflexive". In the mean time, in [163], having observed the implications (isomorphic to uniformly non-square) $\Rightarrow$ super-reflexive and (isomorphic to uniformly convex) $\Rightarrow$ super-reflexive, James asked whether the converses hold. In his remarkable paper [131], Enflo proved that indeed the converses are true. In Theorem 3.37, this corresponds to the equivalence of (i), (iii), (iii)$'$ and (iv) which all come from [131]. The equivalence with (ii) and (ii)$'$ (i.e. the existence of moduli of power type) was proved later in [227].

We follow [227] throughout §3.4. The strong law of large numbers for super-reflexive spaces given in Theorem 3.43 (essentially from [227]) is modeled on Beck's strong law of large numbers ([73]) for B-convex Banach spaces, that is restricted to martingales with independent increments.

Appendix 1: Ultrafilters. Ultraproducts

Let I be a "directed set". By this we mean that I is a partially ordered set such that for any i, j in I there exists k in I such that $k \geq i$ and $k \geq j$.

If $(x_i)_{i \in I}$ is a family in a metric space, we view $(x_i)_{i \in I}$ as a "generalized sequence" so that $x_i \to x$ means that $\forall \varepsilon > 0\ \exists j$ such that $\forall i \geq j\ d(x_i, x) < \varepsilon$.

Definition. Consider a linear form $\mathcal{U} \in \ell_\infty(I)^*$ that is also a $*$-homomorphism

$$\text{(i.e.}\quad \forall x, y \in \ell_\infty(I) \qquad \mathcal{U}(xy) = \mathcal{U}(x)\mathcal{U}(y) \quad \text{and} \quad \mathcal{U}(\bar{x}) = \overline{\mathcal{U}(x)}).$$

We will say that $\mathcal{U}$ is an ultrafilter adapted to I if for any (x_i) in $\ell_\infty(I)$ such that $x_i \to x$ we have $\mathcal{U}((x_i)) = x$.

Remark 3.46. The existence of ultrafilters adapted to I is easy to check: let $\delta_i \in \ell_\infty(I)^*$ be the evaluation homomorphism defined by $\delta_i(x) = x_i$. Let F_j be the pointwise closure of the set $\{\delta_i \mid i \geq j\}$. Since I is a directed index set, the intersection of finitely many of the F_j's is non-empty. Thus, by the weak-$*$ compactness of the unit ball of $\ell_\infty(I)^*$, the intersection of the whole family of sets $\{F_j\}$ is non void and it is formed of ultrafilters in the above sense.

We will denote by convention

$$\lim_{\mathcal{U}} x_i = \mathcal{U}((x_i)_{i\in I}).$$

Given a family of Banach spaces $(B_i)_{i\in I}$, let $\mathcal{B} = \left(\oplus \sum_{i\in I} B_i\right)_\infty$, i.e. $\mathcal{B}$ is formed of families $b = (b_i)_{i\in I}$ with $b_i \in B_i$ for all i such that $\|b\|_{\mathcal{B}} = \sup_{i\in I} \|b_i\| < \infty$. For any b in $\mathcal{B}$ we set

$$p_{\mathcal{U}}(b) = \lim_{\mathcal{U}} \|b_i\|_{B_i}.$$

Then $p_{\mathcal{U}}$ is a semi-norm on $\mathcal{B}$. The ultraproduct $\prod_{i\in I} B_i/\mathcal{U}$ is defined as the Banach space quotient $\mathcal{B}/\ker(p_{\mathcal{U}})$. Fix an element x in $\prod_{i\in I} B_i/\mathcal{U}$. It is important to observe that for any representative $(b_i)_{i\in I}$ of the equivalence class of x modulo $\ker(p_{\mathcal{U}})$ we have

$$\|x\|_{\prod_{i\in I} B_i/\mathcal{U}} = \lim_{\mathcal{U}} \|b_i\|_{B_i}. \tag{3.48}$$

We will denote by $\dot{b}$ the element of $\prod_{i\in I} B_i/\mathcal{U}$ determined by $b = (b_i)_{i\in I}$ so we can rewrite (3.48) as $\|\dot{b}\| = \lim_{\mathcal{U}} \|b_i\|_{B_i}$. Another useful observation is that if for some j we have $b_i = b'_i$ $\forall i \geq j$ then $\dot{b} = \dot{b}'$. Indeed, this implies $\|b_i - b'_i\| \to 0$ (relative to the directed set I) and hence $\lim_{\mathcal{U}} \|b_i - b'_i\| = 0$.

Remark. Let K be a compact subset of a locally convex space L. Let $(y_i)_{i\in I}$ be a family of elements of K. Clearly there is a unique y in K such that for any linear form $\xi \in L^*$ we have $\xi(y) = \lim_{\mathcal{U}} \xi(y_i)$. In that case also we will denote $y = \lim_{\mathcal{U}} y_i$.

When $B_i = B$ for all $i \in I$, we say that $\prod B_i/\mathcal{U}$ is an ultrapower and we denote it by $B^I/\mathcal{U}$.

The following elementary lemma will be useful

Lemma 3.47. *Let E, Y be Banach spaces. let S be an ε-net in the unit sphere of E and let $u\colon\ E \to Y$ be a linear operator such that*

$$\forall s \in S \qquad 1 - \delta \leq \|u(s)\| \leq 1 + \delta.$$

Then

$$\forall x \in E \qquad \left(\frac{1-\delta-2\varepsilon}{1-\varepsilon}\right)\|x\| \leq \|u(x)\| \leq \left(\frac{1+\delta}{1-\varepsilon}\right)\|x\|$$

Proof. Assume $\dim(E) < \infty$ (this is the only case we will use). Consider $x \in E$ with $\|x\| = 1$ and $\|u\| = \|ux\|$. Choose $s \in S$ such that $\|x - s\| \leq \varepsilon$. Then $\|u\| = \|ux\| \leq \|us\| + \|u(x-s)\| \leq (1+\delta) + \varepsilon\|u\|$ and hence $\|u\| \leq (1+\delta)(1-\varepsilon)^{-1}$. In the converse direction, if $\|x\| = 1$ we have

$$\|ux\| \geq \|us\| - \varepsilon\|u\| \geq 1 - \delta - \varepsilon(1+\delta)(1-\varepsilon)^{-1} = (1-\delta-2\varepsilon)(1-\varepsilon)^{-1}.$$

The argument can be easily adapted to the infinite dimensional case. $\square$

Lemma 3.48. *Assume that each space in the family $(B_i)_{i\in I}$ is f.r. in a Banach space B. Then the ultraproduct $\prod B_i/\mathcal{U}$ is f.r. in B. In particular, any ultrapower $B^I/\mathcal{U}$ of B is f.r. in B.*

Proof. Let $E \subset \prod B_i/\mathcal{U}$ be a finite dimensional subspace. Note that since its unit sphere is compact it admits a *finite* ε-net S. Let $(\dot{e}_1, \dots, \dot{e}_n)$ be a linear basis of E with representatives $(e_1(i))_{i\in I}, \dots, (e_n(i))_{i\in I}$. Any $x \in E$ can be uniquely written as $x = \sum_1^n \alpha_j \dot{e}_j$ ($\alpha_j \in \mathbb{K}$). We define $u_i\colon\ E \to B_i$ by setting $u_i(x) = \sum_1^n \alpha_j e_j(i)$ for each i in I. Note that $\forall x \in E$, $\overbrace{(u_i(x))}^{\cdot}{}_{i\in I} = x$. Therefore by (3.48) we have

$$\forall x \in E \qquad \lim_{\mathcal{U}} \|u_i(x)\| = \|x\|.$$

Fix $\delta > 0$. Since S is finite there is j such that

$$\forall i \geq j\ \forall s \in S \qquad 1 - \delta < \|u_i(s)\| < 1 + \delta$$

and hence by Lemma 3.47 we have

$$\forall x \in E \qquad (1 - \delta - 2\varepsilon)(1 - \varepsilon)^{-1}\|x\| \leq \|u_i(x)\| \leq (1 + \delta)(1 - \varepsilon)^{-1}\|x\|.$$

Thus we conclude that E is $(1 + f(\varepsilon, \delta))$-isometric to $u_i(E) \subset B_i$ for some function $(\varepsilon, \delta) \mapsto f(\varepsilon, \delta)$ tending to 0 when ε and δ tend to 0. Since each B_i is f.r. in B we conclude that $\prod B_i/\mathcal{U}$ is f.r. in B. $\square$

Chapter 4

Uniformly convex Banach space valued martingales

4.1 Uniform convexity

This chapter is based mainly on [227]. The main result is:

Theorem 4.1. *Any uniformly convex Banach space B admits an equivalent norm $|\cdot|$ satisfying for some constant $\delta > 0$ and some $2 \le q < \infty$*

$$(4.1) \qquad \forall x, y \in B \qquad |\frac{x+y}{2}|^q + \delta|\frac{x-y}{2}|^q \le \frac{|x|^q + |y|^q}{2},$$

or equivalently

$$(4.2) \qquad \forall x, y \in B \qquad |x|^q + \delta|y|^q \le \frac{|x+y|^q + |x-y|^q}{2}.$$

In other words, B with its new norm is at least as uniformly convex as L^q (for some $2 \le q < \infty$). The argument crucially uses martingale inequalities, but the relevant inequalities (see Corollary 4.7 below) are "weaker" than those expressing the UMD property.

We recall:

Definition 4.2. A Banach space B is called uniformly convex if for any $0 < \varepsilon \le 2$ there is a $\delta > 0$ such that for any pair x, y in B the following implication holds

$$(\|x\| \le 1, \|y\| \le 1, \|x-y\| \ge \varepsilon) \Rightarrow \left\|\frac{x+y}{2}\right\| \le 1 - \delta.$$

The modulus of uniform convexity $\delta_B(\varepsilon)$ is defined as the "best possible" δ i.e.

$$\delta_B(\varepsilon) = \inf\left\{1 - \left\|\frac{x+y}{2}\right\| \;\middle|\; \|x\| \le 1, \|y\| \le 1, \|x-y\| \ge \varepsilon\right\}.$$

Note that uniform convexity obviously passes from B to any subspace (resp. any quotient B/S) $S \subset B$ with $\delta_S(\varepsilon) \geq \delta_B(\varepsilon)$ (resp. $\delta_{B/S}(\varepsilon) \geq \delta_B(\varepsilon)$) for all $0 < \varepsilon \leq 2$.

It is easy to see that if $B = \mathbb{C}$ or if B is a Hilbert space of $\mathbb{R}$-dimension ≥ 2, we have $\delta_B(\varepsilon) = 1 - (1 - \varepsilon^2/4)^{1/2}$. Indeed, the parallelogram identity can be equivalently written as

$$\left\|\frac{x+y}{2}\right\|^2 + \left\|\frac{x-y}{2}\right\|^2 = \frac{\|x\|^2 + \|y\|^2}{2}$$

from which $\delta_B(\varepsilon) \geq 1 - (1 - \varepsilon^2/4)^{1/2}$ $(\geq \varepsilon^2/8)$ can be deduced and this is obviously optimal.

Since by Dvoretzky's theorem (see Th. 4.38 below) any infinite dimensional Banach space B contains ℓ_2^n's almost isometrically (in particular for $n = 2$), we must have $\delta_B(\varepsilon) \leq (1 - \varepsilon^2/4)^{1/2}$ hence $\delta_B(\varepsilon) \in O(\varepsilon^2)$ when $\varepsilon \to 0$. Actually, by [221], this already holds for any B with $\dim(B) > 1$. We will show in §4.3 that

$$\delta_{L_p}(\varepsilon) \sim \begin{cases} C_p\varepsilon^2 & \text{if } 1 < p \leq 2 \\ C_p\varepsilon^p & \text{if } 2 \leq p < \infty. \end{cases}$$

Moreover, it is easy to see that L_1 and ℓ_1 are not uniformly convex. Also (note that ℓ_1 isometrically embeds in L_∞ or ℓ_∞) L_∞ and ℓ_∞ are not uniformly convex.

The following result (due to David Milman) is classical.

Theorem 4.3. *Any uniformly convex Banach space is reflexive.*

Proof. Let U_B denote the unit ball of B. Fix $x^{**} \in B^{**}$ with $\|x^{**}\| = 1$. Let (x_i) be a generalized sequence in the unit ball of B converging to $x^{**} \in B^{**}$ for the topology $\sigma(B^{**}, B^*)$, i.e. such that $\langle x_i, \xi\rangle \to \langle x^{**}, \xi\rangle$ for any ξ in U_{B^*}. Clearly this implies $\|x_i\| \to \|x^{**}\|$ (indeed, for $\varepsilon > 0$ choose ξ such that $\langle x^{**}, \xi\rangle > 1 - \varepsilon$ and note $|\langle x_i, \xi\rangle| \leq \|x_i\|$), and similarly $\|2^{-1}(x_i + x_j)\| \to \|x^{**}\| = 1$ when $i, j \to \infty$. If B is assumed uniformly convex, this forces $\|x_i - x_j\| \to 0$ when $i, j \to \infty$ and hence by Cauchy's criterion x_i converges in norm to $x \in B$. Obviously we must have $x^{**} = x$ so we conclude $B^{**} = B$. □

We will use below the following results due to Figiel ([133, 134]). Although their proofs are elementary, the details are tedious so we skip them here.

Lemma 4.4. *Let B be uniformly convex.*

(i) *The function $\varepsilon \to \delta_B(\varepsilon)/\varepsilon$ is non-decreasing on* $[0, 2]$.

(ii) *For any measure space (Ω, μ), and any $1 < r < \infty$ the space $L_r(\mu; B)$ (in particular $L_2(\mu; B)$) is uniformly convex.*

Remark 4.5. With our definition of $\delta_B(\varepsilon)$, it is obvious that $\varepsilon \to \delta_B(\varepsilon)$ is non-decreasing. This is less obvious (but nevertheless true) for the function $\varepsilon \to \hat{\delta}_B(\varepsilon)$ defined by

$$\hat{\delta}_B(\varepsilon) = \inf\left\{1 - \left\|\frac{x+y}{2}\right\| \;\middle|\; \|x\| = \|y\| = 1,\ \|x - y\| = \varepsilon\right\}.$$

Indeed, it turns out that $\hat{\delta}_B(\varepsilon) = \delta_B(\varepsilon)$ if the (real) dimension of B is at least 2 (see e.g. [133]).

To illustrate the next statement by a concrete example, let us anticipate the forthcoming § 4.3 and consider the case of $B = L_q$. As we will show in § 4.3: If $2 \le q < \infty$ we have

$$\forall x, y \in L_q \qquad \left\|\frac{x+y}{2}\right\|_q^q + \left\|\frac{x-y}{2}\right\|_q^q \le \frac{\|x\|_q^q + \|y\|_q^q}{2},$$

which implies $\delta_{L_q}(\varepsilon) \ge 1 - (1 - (\varepsilon/2)^q)^{1/q} \sim q^{-1}(\varepsilon/2)^q$.
If $1 < q \le 2$, we will show :

$$\forall x, y \in L_q \qquad \left(\left\|\frac{x+y}{2}\right\|_q^2 + (q-1)\left\|\frac{x-y}{2}\right\|_q^2\right)^{1/2} \le \left(\frac{\|x\|_q^q + \|y\|_q^q}{2}\right)^{1/q},$$

from which we deduce $\delta_{L_q}(\varepsilon) \ge 1 - (1 - (q-1)(\varepsilon/2)^2)^{1/2} \ge (q-1)\varepsilon^2/8$.

Theorem 4.6. *Let $2 \le q < \infty$ and let $\alpha > 0$ and C be fixed positive constants. The following two properties of a Banach space B are equivalent:*

(i) *There is a norm $|\cdot|$ on B such that for all x, y in B we have $\alpha\|x\| \le |x| \le \|x\|$ and*

$$|\frac{x+y}{2}|^q + \|\frac{x-y}{2C}\|^q \le \frac{|x|^q + |y|^q}{2}. \tag{4.3}$$

(ii) *For all B-valued dyadic martingales $(M_n)_{n\ge 0}$ in $L_q(B)$ we have*

$$\alpha^q \mathbb{E}\|M_0\|^q + C^{-q}\sum\nolimits_1^\infty \mathbb{E}\|dM_n\|^q \le \sup\nolimits_{n\ge 0}\mathbb{E}\|M_n\|^q. \tag{4.4}$$

Moreover, this implies:

(iii) *All B-valued martingales $(M_n)_{n\ge 0}$ in $L_q(B)$ satisfy*

$$\alpha^q \mathbb{E}\|M_0\|^q + 2(2C)^{-q}\sum\nolimits_1^\infty \mathbb{E}\|dM_n\|^q \le \sup\nolimits_{n\ge 0}\mathbb{E}\|M_n\|^q.$$

Proof. (i) $\Rightarrow$ (ii) Consider a dyadic martingale on $\Omega = \{-1, 1\}^{\mathbb{N}_*}$ associated to $\mathcal{A}_n = \sigma(\varepsilon_1, \dots, \varepsilon_n)$, where $\varepsilon_n\colon\ \Omega \to \{-1, 1\}$ denotes the n-th coordinate. Then $\forall n \ge 1$ $dM_n = \varepsilon_n \Delta_{n-1}$ with Δ_{n-1} $(n-1)$-measurable. Let $x = M_{n-1}(\omega)$, $y = \Delta_{n-1}(\omega)$. Then (i) implies for any fixed ω

$$|M_{n-1}(\omega)|^q + C^{-q}\|dM_n(\omega)\|^q \le \int |M_{n-1}(\omega) + \varepsilon_n(\omega')\Delta_{n-1}(\omega)|^q dP(\omega').$$

Integrating this with respect to ω, we find (since ε_n and $\mathcal{A}_{n-1}$ are independent)

$$\mathbb{E}|M_{n-1}|^q + C^{-q}\mathbb{E}\|dM_n\|^q \le \mathbb{E}|M_n|^q,$$

which yields after a summation over $n \geq 1$

$$\mathbb{E}|M_0|^q + C^{-q} \sum_{n \geq 1} \mathbb{E}\|dM_n\|^q \leq \sup \mathbb{E}|M_n|^q.$$

Finally, replacing $|\ |$ by the equivalent norm $\|\ \|$, we obtain (ii).
(i) $\Rightarrow$ (iii) The proof is similar to the preceding. We take $x = M_{n-1}, y = M_n$. This yields after integration of (4.3)

$$\mathbb{E}|M_{n-1} + 2^{-1}dM_n|^q + C^{-q}\mathbb{E}\|2^{-1}dM_n\|^q \leq 2^{-1}(\mathbb{E}|M_{n-1}|^q + \mathbb{E}|M_n|^q) \tag{4.5}$$

but also we have trivially (Jensen)

$$\mathbb{E}|M_{n-1}|^q \leq \mathbb{E}|M_{n-1} + 2^{-1}dM_n|^q$$

hence plugging this into (4.5) we may subtract $2^{-1}\mathbb{E}|M_{n-1}|^q$ to both sides of the resulting inequality and after multiplication by 2 we find

$$\mathbb{E}|M_{n-1}|^q + 2C^{-q}\mathbb{E}\|2^{-1}dM_n\|^q \leq \mathbb{E}|M_n|^q$$

then the proof of (i) $\Rightarrow$ (iii) is completed exactly as above for (i) $\Rightarrow$ (ii).

(ii) $\Rightarrow$ (i) Assume (ii). We define the norm $|\ |$ as follows: for any x in B we set

$$|x|^q = \inf \left\{ \mathbb{E}\|M_N\|^q - C^{-q} \sum\nolimits_{n=1}^{N} \mathbb{E}\|dM_n\|^q \right\}$$

where the infimum runs over all N and all (finite) dyadic martingales $(M_0, M_1, \ldots, M_N)$ which start at x, i.e. such that $M_0 = x$.

By (4.4), we have for any x in B

$$\alpha^q \|x\|^q \leq |x|^q$$

and consideration of the trivial martingale $M_n \equiv x$ yields

$$|x|^q \leq \|x\|^q, \tag{4.6}$$

so that $|\ |$ is indeed equivalent to the original norm on B. Now consider x, y in B and fix $\varepsilon > 0$. Let M', M'' be finite martingales with $x = M'_0$ and $y = M''_0$ such that (note that we may clearly increase N, by adding null increments, in order to use the same N for both martingales)

$$\mathbb{E}\|M'_N\|^q - C^{-q} \sum\nolimits_1^N \mathbb{E}\|dM'_n\|^q < |x|^q + \varepsilon$$
$$\mathbb{E}\|M''_N\|^q - C^{-q} \sum\nolimits_1^N \mathbb{E}\|dM''_N\|^q < |y|^q + \varepsilon.$$

Then, let (M_n) be the martingale that starts at $(x+y)/2$, i.e. $M_0 \equiv (x+y)/2$, jumps with M_1 either to x or to y with equal probability $1/2$ and then continues along the paths of M' or M'' depending on $M_1 = x$ or $M_1 = y$. More precisely, we can write M_n as follows: (Since M'_k and M''_k depend only on $\varepsilon_1, \dots, \varepsilon_k$ we may denote them as $M'_k(\varepsilon_1, \dots, \varepsilon_k)$ and $M''_k(\varepsilon_1, \dots, \varepsilon_k)$.)

$$
\begin{aligned}
&M_0 \equiv (x+y)/2, \\
&M_1 = (x+y)/2 + \varepsilon_1(x-y)/2, \dots, \\
&M_n = ((1+\varepsilon_1)/2)M'_{n-1}(\varepsilon_2, \dots, \varepsilon_n) + ((1-\varepsilon_1)/2)M''_{n-1}(\varepsilon_2, \dots, \varepsilon_n).
\end{aligned}
$$

Finally, we clearly have

$$\mathbb{E}\|M_{N+1}\|^q = (\mathbb{E}\|M'_N\|^q + \mathbb{E}\|M''_N\|^q)/2$$

and

$$\sum_1^{N+1} \mathbb{E}\|dM_n\|^q = \|(x-y)/2\|^q + \left(\sum\nolimits_1^N \mathbb{E}\|dM'_n\|^q + \sum\nolimits_1^N \mathbb{E}\|dM''_N\|^q\right)/2$$

thus we find (recalling the original choice of M' and M'')

$$
\begin{aligned}
|(x+y)/2|^q &\le \mathbb{E}\|M_{N+1}\|^q - C^{-q}\sum\nolimits_1^{N+1} \mathbb{E}\|dM_n\|^q \\
&\le (|x|^q + |y|^q)/2 - C^{-q}\|(x-y)/2\|^q + \varepsilon
\end{aligned}
$$

so we obtain

$$|(x+y)/2|^q + C^{-q}\|(x-y)/2\|^q \le (|x|^q + |y|^q)/2, \tag{4.7}$$

and hence by (4.6)

$$|(x+y)/2|^q + C^{-q}|(x-y)/2|^q \le (|x|^q + |y|^q)/2.$$

It is not entirely evident that $|\cdot|$ is a norm, but (4.7) guarantees that for any pair x, y in B the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(t) = |x + ty|^q$ satisfies $f((t_1 + t_2)/2) \le (f(t_1) + f(t_2))/2$ for any $t_1, t_2 \in \mathbb{R}$, and the latter implies (see e.g. [53]) that f is a convex (and hence continuous) function on $\mathbb{R}$. Knowing this, it becomes obvious that $\{x \mid |x| \le 1\}$ is a convex set, so that $|\cdot|$ is indeed a norm on B.

This completes the proof of (ii) $\Rightarrow$ (i). $\square$

Note that when $\alpha = 1$ we have $\|x\| = |x|$ for all x, so that the original norm coincides with the "new" one and hence is uniformly convex. The next result corresponds to the case $\alpha < 1$.

Corollary 4.7. *Fix $2 \le q < \infty$. The following properties of a Banach space B are equivalent.*

(i) *There is an equivalent norm $|\cdot|$ on B such that* (4.1) *holds for some $\delta > 0$.*

(ii) *There is a constant C such that all B-valued martingales $(M_n)_{n\ge 0}$ in $L_q(B)$ satisfy (recall the convention $dM_0 = M_0$)*

$$\sum_{n\ge 0} \mathbb{E}\|dM_n\|^q \le C^q \sup_{n\ge 0} \mathbb{E}\|M_n\|^q. \tag{4.8}$$

(iii) *Same as (ii) for all dyadic martingales, i.e. all martingales based on the dyadic filtration of [0,1], or the corresponding one on* $\{-1,1\}^{\mathbb{N}}$.

We now turn to the main point, i.e. the proof that any uniformly convex B satisfies (4.8) for some p and C. We first place ourselves in a more "abstract" setting, replacing martingales by monotone basic sequences, defined as follows.

Definition 4.8. A finite sequence $\{x_1, \ldots, x_N\}$ of elements in a Banach space is called a monotone basic sequence if for any sequence of scalars $\lambda_1, \ldots, \lambda_N$ we have

$$\sup_{1 \le n \le N} \left\| \sum\nolimits_1^n \lambda_k x_k \right\| \le \left\| \sum\nolimits_1^N \lambda_k x_k \right\|.$$

An infinite sequence $(x_n)_{n \ge 1}$ is called a monotone basic sequence if $(x_1, \ldots, x_N)$ is one for any $1 \le N < \infty$.

Independently of James's work on basic sequences in super-reflexive spaces analogous results (such as (4.9) below) were proved in the USSR by the Gurarii brothers [155] for uniformly convex spaces.

Theorem 4.9. *Let B be a uniformly convex Banach space. Then for any monotone basic sequence $(x_1, \ldots, x_N)$ in B, the following implication holds*

$$\left\| \sum\nolimits_1^N x_k \right\| \le 1 \Rightarrow \|x_1\| + \sum\nolimits_2^N \delta_B(\|x_k\|) \le 1. \tag{4.9}$$

Consequently, there is a number $1 \le q < \infty$ and a constant C such that any monotone basic sequence $(x_1, \ldots, x_N)$ satisfies

$$\left(\sum \|x_k\|^q \right)^{1/q} \le C \left\| \sum\nolimits_1^N x_k \right\|. \tag{4.10}$$

Proof. Let $S_n = \sum_1^n x_k$. Assume $\|S_N\| \le 1$. Fix $2 \le k \le N$. Let $a_k = \|S_k\|$. Using $x = a_k^{-1} S_k$ and $y = a_k^{-1} S_{k-1}$ we find

$$a_k^{-1} \|S_{k-1} + x_k/2\| \le 1 - \delta_B(a_k^{-1} \|x_k\|)$$

and since (by monotone basicity) $a_{k-1} \le \|S_{k-1} + x_k/2\|$, we find

$$a_{k-1} \le a_k(1 - \delta_B(a_k^{-1} \|x_k\|))$$

or equivalently for all $k \ge 2$

$$a_k \delta_B(a_k^{-1} \|x_k\|) \le a_k - a_{k-1}.$$

But then, since $a_k^{-1} \ge 1$, by Lemma 4.4 (i)

$$\delta_B(\|x_k\|) \le a_k - a_{k-1}$$

from which (4.9) follows immediately. We deduce from (4.9) that for all $N \ge 2$

$$\left\| \sum\nolimits_1^N x_k \right\| \le 1 \Rightarrow \inf_{1 \le k \le N} \delta_B(\|x_k\|) \le (N-1)^{-1}.$$

Let $\varepsilon(N)$ be the largest $\varepsilon > 0$ such that $\delta_B(\varepsilon) \le (N-1)^{-1}$. Note that $\varepsilon(N) \to 0$ since $\delta_B(\varepsilon) > 0$ for all $\varepsilon > 0$ and δ_B is non-decreasing. Hence

$$\left\|\sum\nolimits_1^N x_k\right\| \le 1 \Rightarrow \inf \|x_k\| \le \varepsilon(N)$$

which we may rewrite by homogeneity

$$\inf_{1\le k\le 1} \|x_k\| \le \varepsilon(N) \left\|\sum\nolimits_1^N x_k\right\|. \tag{4.11}$$

We will now show that (4.11) implies the second assertion of Theorem 4.9. This follows by a very general principle based on the fact that (4.11) automatically holds for any sequence of N "blocks" built out of a longer monotone basic sequence. More precisely, let us denote by $b(N)$ the best constant b such that for any monotone basic sequence $(x_1, \dots, x_N)$ we have

$$\inf_{1\le k\le N} \|x_k\| \le b \left\|\sum\nolimits_1^N x_k\right\|.$$

It is easy to see that $b(N) \ge b(N+1)$ for all $N \ge 1$. Moreover, a moment of thought shows that b is "submultiplicative" i.e. for all integers N, K we have

$$b(NK) \le b(N)b(K).$$

(Hint: Given $y_1, \dots, y_{NK}$ consider $x_1 = y_1 + \cdots + y_K$, $x_2 = y_{K+1} + \cdots + y_{2K}$, $x_N = y_{(K-1)N+1} + \cdots + y_{NK}$.)

But now (4.11) ensures that $b(N) \le \varepsilon(N)$ and hence that $b(N) \to 0$ when $N \to \infty$. Let us then choose an integer m such that $b(m) < 1$ and let $0 < r < \infty$ be determined by $b(m) = m^{-1/r}$.

Then, by submultiplicativity, we have $b(m^k) \le (m^k)^{-1/r}$ for any $k \ge 1$. If n is arbitrary we choose k so that $m^k \le n < m^{k+1}$ and hence, since $b(\cdot)$ is non-increasing we find finally $b(n) \le m^{1/r} n^{-1/r}$ for all $n \ge 1$.

Let $x_1, \dots, x_N$ be a monotone basic sequence with $\left\|\sum_1^N x_k\right\| \le 1$. Let $(x_{\sigma(1)}, \dots, x_{\sigma(N)})$ be a permutation chosen so that $\|x_{\sigma(1)}\| \ge \cdots \ge \|x_{\sigma(N)}\|$. Note that of course this is a priori no longer a monotone basic sequence. Fix j. Let $1 \le m(1) < m(2) < \cdots < m(j) \le N$ be the places corresponding to $\{\sigma(1), \dots, \sigma(j)\}$ in $[1, \dots, N]$. Let $y_1 = \sum_1^{m(1)} x_k$, $y_2 = \sum\limits_{m(1)<k\le m(2)} x_k, \dots, y_j = \sum\limits_{m(j-1)<k\le N} x_k$.

We have then

$$\inf_{1\le t\le j} \|y_t\| \le b(j)$$

and moreover by the triangle inequality and the "monotony"

$$\|x_{m(1)}\| \le \|y_1\| + \left\|\sum\nolimits_1^{m(1)-1} x_k\right\| \le 2\|y_1\|$$

and similarly

$$\|x_{m(2)}\| \le 2\|y_2\|, \ldots, \|x_{m(j)}\| \le 2\|y_j\|$$

so that we find

$$\|x_{\sigma(j)}\| = \inf_{1\le t\le j} \|x_{m(t)}\| \le 2 \inf_{t\le j} \|y_t\| \le 2b(j).$$

We conclude $\|x_{\sigma(j)}\| \le 2m^{1/r} j^{-1/r}$ and hence for any $q > r$

$$\sum \|x_j\|^q = \sum \|x_{\sigma(j)}\|^q \le (2m^{1/r}) \cdot \sum j^{-q/r}.$$

Thus, for any $q > r$, setting $C = (2m^{1/r} \sum_1^\infty j^{-q/r})^{1/q}$, we obtain the announced result (4.10) . □

Corollary 4.10. *Let B be isomorphic to a uniformly convex Banach space. Fix $1 < s < \infty$. Then there is a number $2 \le q < \infty$ and a constant C such that any B-valued martingale (f_n) in $L_q(B)$ satisfies*

$$\left(\sum_0^\infty \|f_n - f_{n-1}\|^q_{L_s(B)}\right)^{1/q} \le C \sup \|f_n\|_{L_s(B)}. \tag{4.12}$$

Proof. If B is uniformly convex, so is $L_s(B)$ by Lemma 4.4. So this follows from the preceding Theorem. □

We will need a very simple "dualization" of the preceding inequality:

Proposition 4.11. *Let $(\mathcal{A}_n)_{n\ge0}$ be a filtration on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with $\mathcal{A} = \mathcal{A}_\infty$. Let $1 < s < \infty$ and $1 \le q' \le 2 \le q \le \infty$ with $\frac{1}{q} + \frac{1}{q'} = 1$. The following properties of a Banach space B are equivalent.*

(i) *There is a constant C such that for all B-valued martingales $(f_n)_{n\ge0}$ adapted to $(\mathcal{A})_{n\ge0}$ we have (recall $f_{-1} \equiv 0$ by convention)*

$$\left(\sum_0^\infty \|f_n - f_{n-1}\|^q_{L_s(B)}\right)^{1/q} \le C \sup \|f_n\|_{L_s(B)}.$$

(ii) *There is a constant C' such that for all B^*-valued martingales $(g_n)_{n\ge0}$ adapted to $(\mathcal{A}_n)_{n\ge0}$ we have*

$$\sup \|g_n\|_{L_{s'}(B^*)} \le C' \left(\sum_0^\infty \|g_n - g_{n-1}\|^{q'}_{L_{s'}(B^*)}\right)^{1/q'}.$$

Moreover the best constants C and C' satisfy $C/2 \le C' \le C$.

Proof. Assume (i). Fix n. Let $g_n \in L_{s'}(B^*)$. For any $\varepsilon > 0$ there is f_n with $\|f_n\|_{L_s(B)} = 1$ such that

$$\|g_n\|_{L_{s'}(B^*)} \le (1 + \varepsilon)|\langle f_n, g_n\rangle|,$$

but $\langle f_n, g_n\rangle = \sum_0^n \langle df_k, dg_k\rangle$ and hence

$$\begin{aligned}\|g_n\|_{L_{s'}(B^*)} &\le (1+\varepsilon)\left|\sum\nolimits_0^n \langle df_k, dg_k\rangle\right| \\ &\le (1+\varepsilon)\left(\sum\nolimits_0^n \|df_k\|^q_{L_s(B)}\right)^{1/q}\left(\sum\nolimits_0^n \|dg_k\|^{q'}_{L_{s'}(B^*)}\right)^{1/q'},\end{aligned}$$

so that by (i) we find

$$\|g_n\|_{L_{s'}(B^*)} \le (1+\varepsilon)C\left(\sum\nolimits_0^n \|dg_k\|^{q'}_{L_{s'}(B^*)}\right)^{1/q'}$$

and (ii) follows immediately with $C' \le C$. Conversely, assume (ii). Fix n and let $f_n \in L_s(B)$. For any $\varepsilon > 0$ there are $\varphi_0, \dots, \varphi_n$ in $L_{s'}(\Omega, \mathcal{A}, \mathbb{P}; B^*)$ with $(\sum_0^n \|\varphi_k\|^{q'}_{L_{s'}(B^*)})^{1/q'} \le 1+\varepsilon$ such that

$$\left|\sum\nolimits_0^n \langle df_k, \varphi_k\rangle\right| = \left(\sum\nolimits_0^n \|df_k\|^q_{L_s(B)}\right)^{1/q}.$$

Note that

$$\sum\nolimits_0^n \langle df_k, \varphi_k\rangle = \sum\nolimits_0^n \langle df_k, (\mathbb{E}_k - \mathbb{E}_{k-1})\varphi_k\rangle = \langle f_n, g_n\rangle$$

where $g_n = \sum_0^n (\mathbb{E}_k - \mathbb{E}_{k-1})(\varphi_k)$. In addition since $dg_k = \mathbb{E}_k\varphi_k - \mathbb{E}_{k-1}\varphi_k$ we have $\|dg_k\|_{L_{s'}(B^*)} \le 2\|\varphi_k\|_{L_{s'}(B^*)}$ and hence $(\sum \|dg_k\|^{q'}_{L_{s'}(B^*)})^{1/q'} \le 2$. Thus we obtain by (ii)

$$\left(\sum\nolimits_0^n \|df_k\|^q_{L_s(B)}\right)^{1/q} = |\langle f_n, g_n\rangle| \le \|f_n\|_{L_s(B)}\|g_n\|_{L_{s'}(B^*)} \le 2C'\|f_n\|_{L_s(B)}.$$

This shows that (ii) $\Rightarrow$ (i) with $C \le 2C'$. □

To prove Theorem 4.1 we apply Corollary 4.10 with B replaced by $L_2(B)$. If we wished, we could use $L_s(B)$ for some $1 < s < \infty$, but the reader should note that we have a priori no control over how q depends on s so we cannot just set $q = s$! Thus the main difficulty is to pass from (4.12) to (4.8). This is precisely what the next crucial result achieves, with a slight loss on the exponent.

Lemma 4.12. *Let $2 \le s < \infty$. Let B be a Banach space. Assume that for some constant χ, all B-valued martingales $(f_n)_{n\ge 0}$ satisfy*

$$\text{(4.13)}\quad \forall N \ge 0 \qquad (N+1)^{-1}\sum_{0\le n\le N} \|df_n\|_{L_2(B)} \le \chi(N+1)^{-1/s}\|f_N\|_{L_2(B)}.$$

Then for each $q > s$ there is a constant $C = C(q, s)$ such that all dyadic B-valued martingales, on $\{1,1\}^{\mathbb{N}}$ with the usual filtration $\mathcal{A}_n = \sigma(\varepsilon_j, j \le n)$, satisfy (4.8).

Proof. By the dualization given by Proposition 4.11, this Lemma is equivalent to the next one, which is proved below. □

Lemma 4.13. *Let $1 < r \le 2$. Let B be a Banach space. Assume that for some constant χ, all B-valued martingales $(f_n)_{n\ge 0}$ satisfy*

$$\forall N \ge 0 \qquad \|f_N\|_{L_2(B)} \le \chi (N+1)^{1/r} \sup_{n\le N} \|df_n\|_{L_2(B)}. \tag{4.14}$$

Then for each $1 < p < r$ there is a constant $C = C(p,r)$ such that all dyadic B-valued martingales satisfy

$$\sup_{n\ge 0} \mathbb{E}\|f_n\|^p \le (C)^p \left(\mathbb{E}\|f_0\|^p + \sum\nolimits_{n\ge 1} \mathbb{E}\|df_n\|^p\right).$$

Proof. **Step 1.** We will first show that for any $p < r$ there is a constant $C_1 = C_1(p,r)$ such that all (finite) dyadic B-valued martingales (f_n) satisfy

$$\left\|\sum\nolimits_0^\infty df_n\right\|_{L_2(B)} \le C_1 \left\|\left(\sum\nolimits_{n\ge 0} \|df_n\|^p\right)^{1/p}\right\|_\infty. \tag{4.15}$$

Clearly this reduces to finite martingales so we will assume that there is $N > 0$ so that $df_n \equiv 0$ for all $n > N$ and we let $f = \sum_0^\infty df_n = f_N$. By homogeneity we may assume $\|(\sum_{n\ge 0} \|df_n\|^p)^{1/p}\|_\infty = 1$. Then we let

$$I_k(\omega) = \{n \ge 0 \mid 2^{-(k+1)/p} < \|df_n(\omega)\| \le 2^{-k/p}\}.$$

Note that since $\|df_n(\omega)\|$ is $\mathcal{A}_{n-1}$ measurable, the set $\{\omega \mid n \in I_k(\omega)\}$ is in $\mathcal{A}_{n-1}$. Therefore we may enumerate the integers in $I_k(\omega)$ using stopping times: we define

$$T_0^{(k)}(\omega) = \inf\{n \in I_k(\omega)\} \quad \text{and for} \quad m = 1, 2, \ldots$$
$$T_m^{(k)}(\omega) = \inf\{n > T_{m-1}^{(k)}(\omega), n \in I_k(\omega)\},$$

with the convention $\inf \phi = N+1$ (we could choose equivalently $\inf \phi = \infty$). Let

$$\Delta_m^{(k)} = f_{T_m^{(k)}} - f_{T_m^{(k)}-1}.$$

Observe that if $T_m^{(k)}(\omega) = N+1$, then $\Delta_m^{(k)}(\omega) = 0$, so that $\sum_m \Delta_m^{(k)}(\omega) = \sum_{n\in I_k(\omega)} df_n(\omega)$. Moreover, we have

$$1 \ge \sum_{n\in I_k(\omega)} \|df_n(\omega)\|^p > 2^{-(k+1)}|I_k(\omega)|$$

and hence

$$|I_k(\omega)| < 2^{k+1}.$$

This implies $\Delta_m^{(k)} = 0$, $\forall m > 2^{k+1}$. Moerover, by definition of $I_k(\omega)$ we have $\|\Delta_m^{(k)}(\omega)\| \le 2^{-k/p}$ for a.a. ω. Therefore our hypothesis (4.14) implies that for each fixed k

$$\left\|\sum_{n\in I_k(\omega)} df_n(\omega)\right\|_{L_2(B)} = \left\|\sum \Delta_m^{(k)}\right\|_{L_2(B)} \le \chi 2^{(k+1)/r} \sup\nolimits_m \|\Delta_m^{(k)}\|_{L_2(B)}$$
$$\le \chi 2^{1/r} 2^{k(\frac{1}{r}-\frac{1}{p})}.$$

We remind the reader (see Proposition 1.8 and the exercises following it) that if T_1, T_2 are stopping times and f_n converges (say) in $L_1(B)$ we have

$$f_{T_1\wedge T_2} = \mathbb{E}^{\mathcal{A}_{T_1}} f_{T_2} = \mathbb{E}^{\mathcal{A}_{T_1}} f_{T_2},$$

therefore, if $T_1 \le T_2 \le T_3$ are stopping times, we have

$$\mathbb{E}^{\mathcal{A}_{T_1}}(f_{T_3} - f_{T_2}) = 0. \tag{4.16}$$

Since, by an earlier observation, the sets $\{\omega \mid n \in I_k(\omega)\}$ are all in $\mathcal{A}_{n-1}$, the stopping times $T_m^{(k)}$ are "predictable," i.e. $T_m^{(k)} - 1$ is also a stopping time, and hence, since $\Delta_m^{(k)} = (f_{T_m^{(k)}} - f_{T_m^{(k)}-1})$ if we set $\mathcal{B}_m = \mathcal{A}_{T_m^{(k)}}$, by (4.16) we have $\mathbb{E}^{\mathcal{B}_{m-1}}(\Delta_m^{(k)}) = 0$ (because $T_{m-1}^{(k)} \le T_m^{(k)} - 1 \le T_m^{(k)}$). Thus we obtain

$$\left\|\sum\nolimits_0^\infty df_n\right\|_{L_2(B)} \le \sum\nolimits_{k\ge 0}\left\|\sum_{n\in I_k(\omega)} df_n(\omega)\right\|_{L_2(B)} \le \chi 2^{1/r}\sum\nolimits_0^\infty 2^{k(\frac{1}{r}-\frac{1}{p})},$$

and Step 1 follows with $C_1(p,r) = \chi 2^{\frac{1}{r}}(1-2^{\frac{1}{r}-\frac{1}{p}})^{-1}$.

Step 2. For all $1 < p < r$ there is a constant $C_2 = C_2(p,r)$ such that all B-valued dyadic martingales $(f_n)_{n\ge 0}$ satisfy

$$(\sup_{t>0} t^{p/2}\mathbb{P}\{\sup\|f_n\| > t\})^{2/p} \le C_2\left(\sum\nolimits_0^\infty \mathbb{E}\|df_n\|^p\right)^{1/p}.$$

By Step 1 and Doob's inequality (see Theorem 1.9) we have

$$(\sup_{t>0} t^2\mathbb{P}\{\sup\|f_n\| > t\})^{1/2} \le 2C_1\left\|\left(\sum\nolimits_0^\infty \|df_n\|^p\right)^{1/p}\right\|_\infty. \tag{4.17}$$

Let $V_n = (\sum_0^n \|df_k\|^p)^{1/p}$ for all $n \ge 0$ and $V_\infty = (\sum_0^\infty \|df_n\|^p)^{1/p} = \sup_n V_n$. Fix $s > 0$. Let $T = \inf\{n \ge 0 \mid V_{n+1} > s\}$. Note that, since V_{n+1} is $\mathcal{A}_n$-measurable for all $n \ge 0$, T is a stopping time. We then repeat the trick used for Lemma 8.20: We have $1_{\{T>0\}}V_T \le s$. By (4.17) applied to the martingale $(1_{T>0}f_{n\wedge T})$ this implies

$$t^2\mathbb{P}\{\sup\|f_{n\wedge T}\| > t, T > 0\} \le (2C_1 s)^2. \tag{4.18}$$

Note $\{T < \infty\} = \{V_\infty > s\}$. Therefore for any $t > 0$

$$\begin{aligned}\mathbb{P}\{\sup\|f_n\| > t\} &\le \mathbb{P}\{T < \infty\} + \mathbb{P}\{T = \infty,\ \sup\|f_n\| > t\}\\ &\le \mathbb{P}\{V_\infty > s\} + \mathbb{P}\{T > 0,\ \sup\|f_{n\wedge T}\| > t\}\\ &\le \mathbb{P}\{V_\infty > s\} + (2C_1 s)^2/t^2\\ &\le s^{-p}\mathbb{E}V_\infty^p + (2C_1 s)^2/t^2.\end{aligned}$$

We may assume $\mathbb{E}V_\infty^p = 1$ by homogeneity. Choosing (say) $s = \sqrt{t}$ we obtain (we could use a better choice for s but it is of no consequence for the next step)

$$\forall t \geq 1 \qquad \mathbb{P}(\sup \|f_n\| > t\} \leq t^{-p/2} + (2C_1)^2 t^{-1} \leq (1 + (2C_1)^2) t^{-p/2}$$

and Step 2 follows with $C_2 = (1 + (2C_1)^2)^{2/p}$.

Step 3. For any $p < r$ there is a constant $C_3 = C_3(p, r)$ such that all B-valued dyadic martingales $(f_n)_{n \geq 0}$ satisfy

$$\sup_{t>0} t^p \mathbb{P}\{\sup \|f_n\| > t\} \leq (C_3)^p \sum\nolimits_0^\infty \mathbb{E}\|df_n\|^p.$$

We will use the reverse Hölder principle from Appendix 2 in Chapter 8 (this trick goes back to Burkholder [100]). Consider f in $L_p(\mathcal{A}_N; B)$ with $f_0 = 0$ so that f depends only on $(\varepsilon_1, \ldots, \varepsilon_N)$ and assume that $\mathbb{E}\sum_0^\infty \|df_n\|^p = 1$. We introduce a sequence of independent copies of f on $\{-1, 1\}^{\mathbb{N}}$ as follows: Let $\omega = (\varepsilon_n)_{n \in \mathbb{N}}$. We set

$$\begin{aligned}
&f^{(1)}(\omega) = f(\varepsilon_1, \ldots, \varepsilon_N)\\
&f^{(2)}(\omega) = f(\varepsilon_{N+1}, \ldots, \varepsilon_{2N})\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&f^{(m)}(\omega) = f(\varepsilon_{(m-1)N+1}, \ldots, \varepsilon_{mN}).
\end{aligned}$$

We then consider $g = m^{-1/p}(f^{(1)} + \cdots + f^{(m)})$. We have clearly $\mathbb{E}\sum_0^\infty \|dg_n\|^p = \mathbb{E}\sum_0^\infty \|df_n\|^p = 1$. Therefore by Step 2

$$\forall t > 0 \qquad t^{p/2} \mathbb{P}\{\sup \|g_n\| > t\} \leq (C_2)^{p/2}.$$

Note $\sup_{1 \leq k \leq m} \|m^{-1/p} f^{(k)}\| \leq 2 \sup_{n \geq 0} \|g_n\|$ and hence

$$\mathbb{P}(\sup_{1 \leq k \leq m} m^{-1/p} \|f^{(k)}\| > t/2) \leq (C_2 t^{-1})^{p/2}, \tag{4.19}$$

so we may invoke the reverse Hölder principle formulated in Proposition 8.53 to deduce from (4.19) that for some constant C_3 we have

$$\forall t > 0 \qquad \mathbb{P}\{\|f\| > t\} \leq (C_3)^p t^{-p}$$

and Step 3 follows by homogeneity.

We can now complete the proof by the Marcinkiewicz Theorem 8.51: indeed, for any $1 < p < r$, choose p_0, p_1 such that $1 < p_0 < p < p_1 < r$. Let $D = \{-1, 1\}^{\mathbb{N}}$ equipped with its usual probability ν. Let $D_n = \{-1, 1\}^n$ equipped with the uniform probability ν_n. Consider the space Λ_p of all sequences $(\varphi_n)_{n \geq 0}$ with $\varphi_0 \in L_p(D_0, \nu_0; B)$, $\varphi_n \in L_p(D_{n-1}, \nu_{n-1}; B)$ for all $n \geq 1$ such that $\sum \|\varphi_n\|_{L_p(B)}^p < \infty$. We set $\|(\varphi_n)\|_{\Lambda_p} = (\sum_0^\infty \|\varphi_n\|_{L_p(B)}^p)^{1/p}$. Note that we may clearly identify isometrically Λ_p with $L_p(\Omega, \mu; B)$ for a suitable measure space (Ω, μ) obtained as the disjoint union of (D_0, ν_0) and $(D_{n-1}, \nu_{n-1})_{n\geq 1}$. We consider the operator

$$T\colon\ L_p(\Omega, \mu; B) \to L_p(D, \nu; B)$$

defined by $T((\varphi_n)_{n\geq 0}) = \sum_0^\infty \varepsilon_n \varphi_n$. By Step 3 this operator is of weak type (p_0, p_0) and (p_1, p_1) for any p_0, p_1 as above. Therefore by the Marcinkiewicz theorem, T is of strong type (p, p) for any $p_0 < p < p_1$ and this is exactly the conclusion of the Lemma. □

Remark 4.14. Recall that the space BMO associated to $(\mathcal{A}_n)$ is defined in §8.9. It is easy to deduce from (4.15) that $\|(\sum \|df_n\|^p)^{1/p}\|_\infty < \infty$ implies $f = \sum df_n \in$ BMO and that if $f_0 = 0$ we have

$$\|f\|_{\text{BMO}} \leq C_1 \left\| \left(\sum \|df_n\|^p \right)^{1/p} \right\|_\infty .$$

Indeed, for any A atom of $\mathcal{A}_n$, let P_A be the probability defined on A by $\mathbb{P}_A(B) = \mathbb{P}(A)^{-1}\mathbb{P}(A \cap B)$, for any B in $\mathcal{A}_\infty$. Note that for any fixed $n \geq 1$, the sequence $(1_A(f_k - f_{n-1}))_{k\geq n}$ is a (dyadic) martingale on $(A, \mathbb{P}_A)$. Applying (4.15) to that martingale yields

$$\left(\frac{1}{\mathbb{P}(A)} \int_A \|f - f_{n-1}\|^2 d\mathbb{P} \right)^{1/2} \leq C_1 \left\| \left(\sum\nolimits_{k\geq n} \|df_k\|^p \right)^{1/p} \right\|_\infty$$

and hence

$$\|(\mathbb{E}_n \|f - f_{n-1}\|^2)^{1/2}\|_\infty \leq C_1 \left\| \left(\sum\nolimits_1^\infty \|df_n\|^p \right)^{1/p} \right\|_\infty .$$

Proof of Theorem 4.1. If B is uniformly convex, Corollary 4.10 (with $s = 2$) shows that B satisfies the assumption of Lemma 4.12. Therefore we conclude by Corollary 4.7.

□

Theorem 4.1 admits the following refinement:

Theorem 4.15. *Let $2 < q_0 < \infty$. If a uniformly convex Banach space B satisfies*

$$\delta_B(\varepsilon)\varepsilon^{-q_0} \to \infty \quad \textit{when} \quad \varepsilon \to 0$$

then there is an equivalent norm on B for which the associated modulus of convexity δ satisfies for some $q < q_0$

$$\inf_{0<\varepsilon\leq 2} \delta(\varepsilon)\varepsilon^{-q} > 0.$$

Proof. By results due to Figiel, we may replace B by $L_2(B)$. We can then argue exactly as in the preceding proof of Theorem 4.1. Here is a slightly more direct argument: Let $a(N)$ be the smallest constant C such that for all N-tuples of B-valued martingale differences $d_1, \cdots, d_N$ we have

$$N^{-1} \sum_{1 \le n \le N} \|d_n\|_{L_2(B)} \le C \|d_1 + \cdots + d_N\|_{L_2(B)}.$$

Then it is easy to check that $a(NK) \le a(N)a(K)$ for all $N, K \ge 1$. Applying (4.9) in $L_2(B)$ shows that $a(N)N^{-1/q_0} \to 0$ when $N \to \infty$. Then the submultiplicativity implies that there is $q_1 < q_0$ such that $a(N)N^{-1/q_1}$ is bounded. Thus, (4.13) holds with $s = q_1$, so the conclusion follows again, with $q_1 < q < q_0$, from Lemma 4.12 and Corollary 4.7. ☐

Definition 4.16. We will say that a Banach space B is q-uniformly convex if there is a constant $c > 0$ such that $\delta_B(\varepsilon) \ge c\varepsilon^q$ for all $0 < \varepsilon \le 2$.

With this terminology, let us recapitulate:
B is q-uniformly convex iff there is $C > 0$ such that $\forall x, y \in B$ $\|\frac{x+y}{2}\|^q + C^{-q}\|\frac{x-y}{2}\|^q \le \frac{\|x\|^q + \|y\|^q}{2}$ and the latter holds iff (4.4) holds with $\alpha = 1$.

Moreover, B is isomorphic to a q-uniformly convex space iff it satisfies (4.8) for some constant C.

Lastly, any uniflormly convex space is isomorphic to a q-uniformly convex one for some $q < \infty$ (see Theorem 4.1).

4.2 Uniform smoothness

Uniform smoothness is dual to uniform convexity: B is uniformly smooth (resp. uniformly convex) iff B^* is uniformly convex (resp. uniformly smooth). Therefore many of its properties can be deduced from the corresponding properties of uniform convexity. Nevertheless, the intrinsic geometric significance of uniform smoothness is of considerable interest in many questions involving e.g. differentiability of functions on B.

Definition 4.17. A Banach space B is called uniformly smooth if there is a function $t \to \rho(t)$ on $\mathbb{R}_+$ that is $o(t)$ when $t \to 0$ such that for any x, y in the unit sphere of B we have

$$\frac{\|x + ty\| + \|x - ty\|}{2} \le 1 + \rho(t).$$

The modulus of (uniform) smoothness $\rho_B(t)$ is defined as the "best possible" ρ, i.e.

$$\rho_B(t) = \sup\{2^{-1}(\|x + ty\| + \|x - ty\|) - 1 \mid x, y \in B,\ \|x\| = \|y\| = 1\}.$$

With this notation, B is uniformly smooth iff

$$\lim_{t \to 0} \rho_B(t)/t = 0.$$

For example, for a Hilbert space H, we have $\rho_H(t) = (1+t^2)^{1/2} - 1 \simeq t^2/2$. By Dvoretzky's theorem (see Th. 4.38), for any infinite dimensional space B, we must have $\rho_B(t) \geq (1+t^2)^{1/2} - 1$, since $H = \ell_2$ is f.r. in B.

The following formula due to Lindenstrauss [189] (see also [126]) illustrates the dual relationship between δ_B and ρ_{B^*}.

Lemma 4.18. *For any (real or complex) Banach space B*

$$\rho_{B^*}(t) = \sup\{t\varepsilon/2 - \delta_B(\varepsilon) \mid 0 < \varepsilon \leq 2\}. \tag{4.20}$$

Proof. Let $U_B = \{x \in B \mid \|x\| \leq 1\}$ and $S_B = \{x \in B \mid \|x\| = 1\}$. By definition we have in the real case

$$\begin{aligned}
\rho_{B^*}(t) &= \sup\{2^{-1}(\|\xi + t\eta\| + \|\xi - t\eta\|) - 1 \mid \xi, \eta \in S_{B^*}\} \\
&= \sup\{2^{-1}(\langle \xi + t\eta, x\rangle + \langle \xi - t\eta, y\rangle) - 1 \mid \xi, \eta \in S_{B^*}, x, y \in U_B\} \\
&= \sup\left\{\left\|\frac{x+y}{2}\right\| + t\left\|\frac{x-y}{2}\right\| - 1 \;\middle|\; x, y \in U_B\right\} \\
&= \sup_{0<\varepsilon\leq 2} \sup\left\{\left\|\frac{x+y}{2}\right\| + t\varepsilon/2 - 1 \;\middle|\; x, y \in U_B,\ \|x-y\| \geq \varepsilon\right\} \\
&= \sup_{0<\varepsilon\leq 2} \{t\varepsilon/2 - \delta_B(\varepsilon)\}.
\end{aligned}$$

In the complex case, just use $\|\xi \pm t\eta\| = \sup\{\Re(\langle \xi \pm t\eta, x\rangle) \mid x \in U_B\}$. □

It is natural to wonder whether conversely δ_{B^*} is in duality with ρ_B. Unfortunately, this is not true because, unlike ρ_B, the function δ_{B^*} is in general not convex (see [192]). Nevertheless, if we denote by $\widetilde{\delta}_B$ the largest convex function dominated by δ_B, we have a nice duality, and moreover, $\widetilde{\delta}_B$ and δ_B are essentially equivalent. We refer to [136, 133] for more on this.

Lemma 4.19. *For any (real or complex) Banach space B*

$$\widetilde{\delta}_{B^*}(\varepsilon) = \sup\{t\varepsilon/2 - \rho_B(t) \mid 0 < t < \infty\}. \tag{4.21}$$

Moreover for any $0 < \gamma < 1$ and $\varepsilon > 0$ we have

$$(\gamma^{-1} - 1)\delta_{B^*}(\gamma\varepsilon) \leq \widetilde{\delta}_{B^*}(\varepsilon) \leq \delta_{B^*}(\varepsilon).$$

Proof. The first formula is proved just like (4.20), and we find $\sup\{t\varepsilon/2 - \rho_B(t) \mid 0 < t < \infty\} = \sup_{t>0}(\inf_{0<s\leq 2}\{\delta_B(s) + t(\varepsilon - s)/2\}$ but note that being the supremum of affine functions the right hand side of (4.21) is a convex function, that majorizes any affine function f (say $f(\varepsilon) = a\varepsilon + b$) such that $f \leq \delta_{B^*}$ because it is easy to see that

$$\sup_{t>0}(\inf_{0<s\leq 2}\{f(s) + t(\varepsilon - s)/2\} = f(\varepsilon).$$

This establishes (4.21). The second assertion is more delicate, we refer the reader to [133]. □

As a consequence we have

Proposition 4.20. *A Banach space B is uniformly convex (resp. uniformly smooth) iff its dual B^* is uniformly smooth (resp. uniformly convex).*

Proof. If B is uniformly convex, the formula (4.20) clearly implies by elementary calculus that B^* is uniformly smooth (note that ρ_{B^*} is essentially the Legendre conjugate of δ_B). Conversely, if B is uniformly smooth, then Lemma 4.19 implies that B^* is uniformly convex. Note that by Theorem 4.3, B is reflexive if either B or B^* is uniformly convex. From this it is easy to complete the proof. □

In view of the preceding (almost perfect) duality, it is not surprising that the results of §6.1 have analogues for uniform smoothness, so we will content ourselves with a brief outline with mere indications of proofs.

Theorem 4.21. *Let $1 < p \le 2$ and let $\alpha > 0$ and $C > 0$ be fixed constants. The following two properties of a Banach space B are equivalent:*

(i) *There is a equivalent norm $|\ \ |$ on B such that for all x, y in B we have $\|x\| \le |x| \le \alpha^{-1}\|x\|$ and*

$$2^{-1}(|x+y|^p + |x-y|^p) \le |x|^p + C^p\|y\|^p. \tag{4.22}$$

(ii) *For any dyadic B-valued martingale $(M_n)_{n\ge 0}$ in $L_p(B)$ we have*

$$\sup \mathbb{E}\|M_n\|^p \le \alpha^{-p}\mathbb{E}\|M_0\|^p + C^p \sum\nolimits_1^\infty \mathbb{E}\|dM_n\|^p. \tag{4.23}$$

Moreover, this implies:

(iii) *All B-valued martingales in $L_p(B)$ satisfy*

$$\sup \mathbb{E}\|M_n\|^p \le \alpha^{-p}\mathbb{E}\|M_0\|^p + 2C^p \sum\nolimits_1^\infty \mathbb{E}\|dM_n\|^p.$$

Proof. (ii) ⇒ (i). Assume (ii). We define

$$|x|^p = \sup\left\{\mathbb{E}\|M_N\|^p - C^p \sum\nolimits_1^N \mathbb{E}\|dM_n\|^p\right\}$$

where the supremum runs over all $N \ge 1$ and all dyadic martingales $M_0, M_1, \dots, M_N$ such that $M_0 = x$. Note that we trivially have $|x| \ge \|x\|$ (by choosing $M_N = x$), and by (ii) we have $|x| \le \alpha^{-1}\|x\|$, so that $|\cdot|$ and $\|\cdot\|$ are equivalent. The same idea as in the previous section shows that

$$\forall x, y \in B \qquad 2^{-1}(|x|^p + |y|^p) \le |2^{-1}(x+y)|^p + \|2^{-1}(x-y)\|^p$$

or equivalently (replace (x, y) by $(x+y, x-y)$)

$$2^{-1}(|x+y|^p + |x-y|^p) \le |x|^p + C^p\|y\|^p \le |x|^p + C^p|y|^p$$

so we obtain (i). Lastly, in case $x \to |x|$ is not a norm we define

$$|x|_1 = \inf \left\{ \sum |x_k| \right\}$$

over all the decompositions $x = \sum x_k$ as a finite sum of elements of B. Note that for any $t > |x|_1$ we can write $x = \sum \lambda_k x_k$ with $\lambda_k \geq 0$, $\sum \lambda_k = 1$ and $|x_k| \leq t$. Using this, it is then easy to check that (4.22) remains true when $|\ \ |_1$ replaces $|\ \ |$, completing the proof that (ii) $\Rightarrow$ (i).

(i) $\Rightarrow$ (iii) and (i) $\Rightarrow$ (ii). For any $n \geq 1$ and ω we have

$$2^{-1}(|M_{n-1}(\omega)+dM_n(\omega)|^p+|M_{n-1}(\omega)-dM_n(\omega)|^p) \leq |M_{n-1}(\omega)|^p+C^p\|dM_n(\omega)\|^p$$

and hence after integration

$$2^{-1}(\mathbb{E}|M_n|^p + \mathbb{E}|M_{n-1} - dM_n|^p) \leq \mathbb{E}|M_{n-1}|^p + C^p\mathbb{E}\|dM_n\|^p$$

but since $\mathbb{E}|M_{n-1}|^p \leq \mathbb{E}|M_{n-1} - dM_n|^p$ we deduce

$$\mathbb{E}|M_n|^p \leq \mathbb{E}|M_{n-1}|^p + 2C^p\mathbb{E}\|dM_n\|^p$$

and hence (note the telescoping sum)

$$\sup \mathbb{E}|M_n|^p \leq \mathbb{E}|M_0|^p + 2C^p \sum\nolimits_1^\infty \mathbb{E}\|dM_n\|^p.$$

Since $|\cdot|$ is an equivalent norm, (iii) follows. To check (i) $\Rightarrow$ (ii) just observe that in the dyadic case $\mathbb{E}|M_{n-1} - dM_n|^p = \mathbb{E}|M_{n-1} + dM_n|^p = \mathbb{E}|M_n|^p$ (so the factor 2 disappears in the prceding argument). $\square$

Corollary 4.22. *Fix $1 < p \leq 2$. The following properties of a Banach space B are equivalent*

(i) *There is an equivalent norm $|\cdot|$ on B such that for some constant C we have*

$$\forall x, y \in B \qquad 2^{-1}(|x+y|^p + |x-y|^p) \leq |x|^p + C^p|y|^p. \tag{4.24}$$

(ii) *There is a constant C such that all B-valued martingales $(M_n)_{n\geq 0}$ in $L_p(B)$ satisfy (recall $dM_0 = M_0$ by convention)*

$$\sup \mathbb{E}\|M_n\|^p \leq C^p \sum\nolimits_0^\infty \mathbb{E}\|dM_n\|^p. \tag{4.25}$$

(iii) *Same as (ii) for all dyadic martingales.*

The next result is the dual analogue of Theorem 4.9, and although we prefer to give a direct argument, it can be proved by duality.

Theorem 4.23. *Assume B uniformly smooth. Then for any (finite or infinite) monotone basic sequence (x_n) in B, we have*

$$\sup_n \|x_n\| + 2\sum_1^\infty \rho_B(\|x_n\|) \le 1 \Rightarrow \sup_n \|x_1 + \cdots + x_n\| \le 2.$$

Consequently, there is a constant C and $1 < p \le 2$ such that for all N and all monotone basic sequences (x_n) we have

$$\|\sum_1^N x_n\| \le C(\sum \|x_n\|^p)^{1/p}.$$

Proof. Let $S_n = x_1 + \cdots + x_n$. We have

$$\|S_{n-1}\|^{-1}(\|S_{n-1} + x_n\| + \|S_{n-1} - x_n\|) - 2 \le 2\rho_B(\|x_n\|\|S_{n-1}\|^{-1})$$

and also $1 \le \|S_{n-1}\|^{-1}\|S_{n-1} - x_n\|$ by monotony. Put together, this yields

$$\|S_n\| \le \|S_{n-1}\| + 2\|S_{n-1}\|\rho_B(\|x_n\|\|S_{n-1}\|^{-1}).$$

Assume $\|S_{n-1}\| \ge 1$, then $t \to \rho_B(t)/t$ is non-decreasing (since ρ_B is convex) so that $\|S_{n-1}\|\rho_B(\|x_n\|\|S_{n-1}\|^{-1}) \le \rho_B(\|x_n\|)$ and we find

$$\|S_n\| \le \|S_{n-1}\| + 2\rho_B(\|x_n\|).$$

This yields (telescoping sum) that if $\|S_N\| \ge 1$ we have for all $n \ge N$

$$\|S_n\| \le \|S_N\| + 2\sum_{n>N} \rho_B(\|x_n\|).$$

Let N be the first integer (if any) such that $\|S_N\| \ge 1$. Then $\|S_N\| \le 1 + \|x_N\|$ so we obtain $\sup \|S_n\| \le 1 + \|x_N\| + 2\sum_{n>N} \rho_B(\|x_n\|) \le 2$. □

The analogue of Theorem 4.1 for smoothness is now immediate:

Theorem 4.24. *Any uniformly smooth Banach space B admits an equivalent norm $|\cdot|$ satisfying for some constant $C > 0$ and some $1 < p \le 2$*

$$\forall x, y \in B \qquad \frac{|x+y|^p + |x-y|^p}{2} \le |x|^p + C|y|^p. \tag{4.26}$$

Proof. Using Proposition 4.20, this can be easily deduced from Theorem 4.1 by duality. Alternatively, a direct proof can be obtained by combining Theorem 4.9 with Lemma 4.13 and Corollary 4.22. □

Theorem 4.24 admits the following refinement:

Theorem 4.25. *Let $1 < r < 2$. If a uniformly smooth Banach space B satisfies*

$$\rho_B(t)t^{-r} \to 0 \quad \textit{when} \quad t \to 0 \tag{4.27}$$

then there is an equivalent norm on B for which the associated modulus of smoothness ρ satisfies for some $p > r$

$$\sup_{t>0} \rho(t)t^{-p} < \infty.$$

Proof. By the Lindenstrauss duality formula (see (4.18) and Lemma 4.19), this can be immediately deduced from Theorem 4.15 by duality. □

Definition. We will say that B (its unit sphere or its norm) is smooth if for any x, y in B with $x \neq 0$ the function $t \to \|x + ty\|$ is differentiable at $t = 0$.

Remark. Fix $x, y \in B$. Let $f(t) = \|x + ty\|$ $(t \in \mathbb{R})$. Assume that

$$(f(t) + f(-t))/2 - 1 \to 0$$

when $t \to 0$. Then f is differentiable at 0. Indeed, since f is a convex function, it admits left and right derivatives everywhere, in particular at $t = 0$ where we denote them by $f'_-(0)$ and $f'_+(0)$ respectively, but our assumption implies $f'_-(0) = f'_+(0)$ so $f'(0)$ exists (and the converse is obvious). Let us denote $\xi_x(y) = f'(0)$. We now assume that B is smooth, i.e. $\xi_x(y)$ exists for any y in B. We will show that, if B is a *real* Banach space

$$\xi_x \in B^*, \quad \|\xi_x\|_{B^*} = 1 \quad \text{and} \quad \xi_x(x) = \|x\|. \tag{4.28}$$

Taking $y = x$, we immediately find $\xi_x(x) = \|x\|$. Note that $\xi_x(sy) = s\xi_x(y)$ for any $s \in \mathbb{R}$. Moreover, from $\|x + t(y_1 + y_2)/2\| \leq (\|x + ty_1\| + \|x + ty_2\|)/2$ we deduce easily that if $\xi_x(y)$ exists for any y, then we must have $\xi_x(y_1 + y_2) = \xi_x(y_1) + \xi_x(y_2)$, so that $y \to \xi_x(y)$ is a linear form on B. Moreover, from $|f(t) - \|x\|| \leq |t|\|y\|$ we deduce $|\xi_x(y)| \leq \|y\|$ so that (since $\xi_x(x) = \|x\|$) $\|\xi_x\|_{B^*} = 1$.

In addition, ξ_x is the unique $\xi \in B^*$ satisfying (4.28). Indeed, for any such ξ we have (when $|t| \to 0$)

$$\|x\| + t\xi(y) \leq \|x + ty\| = \|x\| + t\xi_x(y) + o(|t|)$$

and hence $\xi = \xi_x$. □

By the preceding remark, if B is uniformly smooth, a fortiori its unit sphere S_B is "smooth," and for any $x \neq 0$ in B there is a *unique* $\xi_x \in S_{B^*}$ satisfying (4.28). It is useful to observe that when B is uniformly smooth the map

$$x \mapsto \xi_x\colon\ B - \{0\} \to S_{B^*}$$

is uniformly continuous when restricted to closed bounded subsets of $B - \{0\}$. More precisely we have (here we reproduce a proof in [4]).

Proposition 4.26. *Let B be a uniformly smooth Banach space. Then*

$$\forall x, y \in B \qquad \|\xi_x - \xi_y\| \leq 2\rho_B(2\big\|x\|x\|^{-1} - y\|y\|^{-1}\big\|)/\big\|x\|x\|^{-1} - y\|y\|^{-1}\big\|. \tag{4.29}$$

In particular, if $\|x\| = \|y\| = 1$

$$\|\xi_x - \xi_y\| \leq 2\rho_B(2\|x - y\|)/\|x - y\|.$$

Proof. Recall that, by definition of ξ_a, for all a, b in B with $a \neq 0$

$$\lim_{|t|\to 0} t^{-1}(\|a+tb\| - \|a\|) = \langle \xi_a, b\rangle.$$

By convexity of $t \mapsto \psi(t) = \|a+tb\| - \|a\|$, the function $t \mapsto \psi(t)/t$ must be non-decreasing on $\mathbb{R}_+$, and hence

$$\langle \xi_a, b\rangle \leq \|a+b\| - \|a\|. \tag{4.30}$$

Since $\xi_x = \xi_{x\|x\|^{-1}}$ it suffices to prove (4.29) when $\|x\| = \|y\| = 1$. In that case, (4.29) becomes

$$\|x-y\|\|\xi_x - \xi_y\| \leq 2\rho_B(2\|x-y\|). \tag{4.31}$$

Let $z \in B$ be such that $\|z\| = \|x-y\|$. Assuming $\|x\| = \|y\| = 1$, we have by repeated use of (4.30) (note also $\langle \xi_x, x-y\rangle = 1 - \langle \xi_x, y\rangle \geq 0$)

$$\begin{aligned}
\langle \xi_y, z\rangle - \langle \xi_x, z\rangle &\leq \|y+z\| - 1 - \langle \xi_x, z\rangle \\
&\leq \|y+z\| - 1 + \langle \xi_x, x-y-z\rangle \\
&\leq \|y+z\| - 1 + \|2x-y-z\| - 1 \\
&= \|x+(y-x+z)\| + \|x-(y-x+z)\| - 2 \\
&\leq 2\rho_B(\|y-x+z\|) \leq 2\rho_B(2\|y-x\|).
\end{aligned}$$

The last step because $\|z\| = \|y-x\|$. Taking the supremum of the preceding over all z with $\|z\| = \|x-y\|$, we obtain (4.31). □

Corollary 4.27. *For $0 < \delta < R < \infty$, let $B(\delta, R) = \{x \in B \mid \delta \leq \|x\| \leq R\}$. The following properties of a Banach space B are equivalent.*

(i) *B is uniformly smooth.*

(ii) *B is smooth and $x \to \xi_x$ is uniformly continuous on the unit sphere S_B.*

(iii) *B is smooth and $x \to \xi_x$ is uniformly continuous on $B(\delta, R)$ for any $0 < \delta < 1 < R < \infty$.*

Proof. (i) ⇒ (ii) follows from Proposition 4.26 and (ii) ⇔ (iii) is easy using $\xi_x = \xi_{x\|x\|^{-1}}$. If (iii) holds, then assuming $\|x\| = \|y\| = 1$ and $|t| < \min(1-\delta, R-1)$ we have by the "calculus fundamental formula"

$$\|x+ty\| - \|x\| = \int_0^t \langle \xi_{x+sy}, y\rangle ds$$

and hence

$$2^{-1}(\|x+ty\| + \|x-ty\| - 2\|x\|) = \int_0^t \langle \xi_{x+sy} - \xi_{x-sy}, y\rangle ds/2.$$

Therefore, we find

$$\rho_B(t) \leq |t| \sup\{\|\xi_x - \xi_{x'}\| \mid x, x' \in B(\delta, R), \|x-x'\| \leq t\}/2,$$

from which (iii) ⇒ (i) is immediate. □

Corollary 4.28. *Let $1 < p \leq 2$. Assume that $\rho_B(t) \in O(t^p)$ when $t \to 0$. Then, for any $0 < \delta < R < \infty$, there is a constant $C = C_{\delta,R}$ such that*

$$\forall x, y \in B(\delta, R) \qquad \|\xi_x - \xi_y\| \leq C\|x - y\|^{p-1}.$$

In particular, if $p = 2$, the map $x \mapsto \xi_x$ is Lipschitzian on $B(\delta, R)$.

Proof. This is an immediate consequence of (4.29) by elementary calculus. □

We refer the reader to [14, 27] for supplementary information and more references.

The property appearing in Corollary 4.28 was already considered in early pioneering work by Fortet and Mourier on the strong law of large numbers for Banach space valued random variables, cf. [137]. As we will show in the next chapter (see Theorem 3.45), the validity of the strong law of large numbers for B-valued martingales is equivalent to the super-reflexivity of B.

Definition 4.29. We will say that a Banach space B is p-uniformly smooth if there is a constant $c > 0$ such that $\rho_B(t) \leq ct^p$ for all $t > 0$.

With this terminology, let us recapitulate:
B is p-uniformly smooth iff there is $C > 0$ such that $\forall x, y \in B$ $\|\frac{x+y}{2}\|^p + C\|\frac{x-y}{2}\|^p \geq \frac{\|x\|^p+\|y\|^p}{2}$ and the latter holds iff (4.23) holds with $\alpha = 1$.

Moreover, B is isomorphic to a p-uniformly smooth space iff it satisfies (4.25) for some constant C.

Lastly, any uniflormly smooth space is isomorphic to a p-uniformly smooth one for some $p > 1$ (see Theorem 4.24).

4.3 Uniform convexity and smoothness of L_p

We should first note of course that any Hilbert space H is both uniformly convex and uniformly smooth, by the "parallelogram identity"

$$\forall x, y \in H \qquad 2^{-1}(\|x + y\|^2 + \|x - y\|^2) = \|x\|^2 + \|y\|^2.$$

The latter implies

$$\delta_H(\varepsilon) = (1 - (1 - \varepsilon^2/4)^{1/2} \simeq \varepsilon^2/8 \quad \text{and} \quad \rho_H(t) = (1 + t^2)^{1/2} - 1 \simeq t^2/2$$

In this section, we denote simply by L_p the space $L_p(\Omega, \mathcal{A}, m)$ where $(\Omega, \mathcal{A}, m)$ is an arbitrary measure space. Our goal is to prove

Theorem 4.30. *(i) If $1 < p \leq 2$, we have: $\forall t > 0$ $\forall \varepsilon \in [0, 2]$*

$$\rho_{L_p}(t) \leq t^p/p \quad \textit{and} \quad \delta_{L_p}(\varepsilon) \geq (p - 1)\varepsilon^2/8$$

(ii) If $2 \leq p' < \infty$, we have $\forall t > 0$ $\forall \varepsilon \in [0, 2]$

$$\rho_{L_{p'}}(t) \leq (p' - 1)t^2/2 \qquad \delta_{L_{p'}}(\varepsilon) \geq (\varepsilon/2)^{p'}/p'.$$

Remark. The constants in the preceding estimates are sharp, i.e. they give the right order of magnitude when t or ε are small. For instance, if $1 < p \le 2$, we have $\rho_{L_p}(t) = t^p/p + o(t^p)$ when $t \to 0$, and similarly for the other estimates.

Part of the preceding statement is very easy to prove by interpolation:

Lemma 4.31. *Let $1 < p \le 2 \le p' < \infty$. We have then:*

$$\forall x, y \in L_p \qquad \left(\frac{\|x+y\|_p^p + \|x-y\|_p^p}{2}\right)^{1/p} \le (\|x\|_p^p + \|y\|_p^p)^{1/p}, \tag{4.32}$$

$$\forall x, y \in L_{p'} \qquad (\|x\|_{p'}^{p'} + \|y\|_{p'}^{p'})^{1/p'} \le \left(\frac{\|x+y\|_{p'}^{p'} + \|x-y\|_{p'}^{p'}}{2}\right)^{1/p'}. \tag{4.33}$$

Proof. Assume $L_p = L_p(\Omega, \mu)$. Let $D_1 = \{-1, 1\}$ equipped with $\nu_1 = (\delta_1 + \delta_{-1})/2$. For (4.32), we consider the operator

$$T\colon\ \ell_p^{(2)}(L_p) \to L_p(D_1, \nu_1; L_p)$$

defined by $T(x, y) = x + \varepsilon_1 y$ (here $\varepsilon_1(\omega) = \omega\ \forall \omega \in D_1$). Note that $\ell_p^{(2)}(L_p)$ (resp. $L_p(D_1, \nu_1; L_p)$) can obviously be identified with the L_p-space associated to the disjoint union of two copies of (Ω, μ) (resp. with $L_p(D_1 \times \Omega, \nu_1 \times \mu)$). Clearly T is a contraction both when $p = 1$ (triangle inequality) and when $p = 2$ (parallelogram inequality). Therefore by interpolation (cf. Corollary **??**) (4.32) is valid for any intermediate value: $1 < p < 2$. The proof of (4.33) can be done in a similar fashion by considering the operator

$$T^*\colon\ L_{p'}(D_1, \mu; L_{p'}) \to \ell_{p'}^{(2)}(L_{p'})$$

interpolating between $p' = 2$ and $p' = \infty$. Alternatively, one can simply observe that (4.33) follows from $\|T\| = \|T^*\|$ when $p^{-1} + {p'}^{-1} = 1$. $\square$

The other estimates follow from:

Lemma 4.32. *If $1 < p < 2$, then for all x, y in L_p*

$$(\|x\|_p^2 + (p-1)\|y\|_p^2)^{1/2} \le \left(\frac{\|x+y\|_p^p + \|x-y\|_p^p}{2}\right)^{1/p} \tag{4.34}$$

or equivalently

$$\left(\left\|\frac{x+y}{2}\right\|_p^2 + (p-1)\left\|\frac{x-y}{2}\right\|_p^2\right)^{1/2} \le \left(\frac{\|x\|_p^p + \|y\|_p^p}{2}\right)^{1/p}. \tag{4.35}$$

Proof. By the 2-point hypercontractive inequality (see Theorem 8.3 with $q = 2$) we know that for any fixed ω

$$(|x(\omega)|^2 + (p-1)|y(\omega)|^2)^{1/2} \le \left(\frac{|(x+y)(\omega)|^p + \|(x-y)(\omega)|^p}{2}\right)^{1/p}$$

taking the L_p-norm of both sides (and using the Hölder–Minkowski contractive inclusion $L_p(\ell_2) \subset \ell_2(L_p)$) we find (4.34) □

Proof of Theorem 4.30. (i) Assume $1 < p \le 2$. By (4.32) we have

$$2^{-1}(\|x+ty\|_p + \|x-ty\|_p) \le (1+t^p)^{1/p} \le 1 + t^p/p$$

and hence $\rho_{L_p}(t) \le t^p/p$. By (4.35), if $\|x\|_p, \|y\|_p \le 1$ and $\|x-y\|_p \ge \varepsilon$ then $\left\|\frac{x+y}{2}\right\|_p \le (1-(p-1)\varepsilon^2/4)^{1/2} \le 1-(p-1)\varepsilon^2/8$, and hence $\delta_{L_p}(\varepsilon) \ge (p-1)\varepsilon^2/8$.

(ii) Now assume $2 \le p' < \infty$. Replacing $(x+y, x-y)$ by (x, y) in (4.33) we obtain

$$\left\|\frac{x+y}{2}\right\|_{p'}^{p'} + \left\|\frac{x-y}{2}\right\|_{p'}^{p'} \le \frac{\|x\|_{p'}^{p'} + \|y\|_{p'}^{p'}}{2}$$

and hence we find

$$\begin{aligned}\delta_{L_{p'}}(\varepsilon) &\ge 1 - (1-(\varepsilon/2)^{p'})^{1/p'} \\ &\ge (\varepsilon/2)^{p'}/p'.\end{aligned}$$

By duality, (4.34) implies

$$(\frac{\|x+y\|_{p'}^{p'} + \|x-y\|_{p'}^{p'}}{2})^{1/p'} \le (\|x\|_{p'}^2 + (p-1)^{-1}\|y\|_{p'}^2)^{1/2} \tag{4.36}$$

and hence

$$\rho_{L_{p'}}(t) \le (1+(p-1)^{-1}t^2)^{1/2} - 1 \le t^2/(2(p-1)) = (p'-1)t^2/2.$$

□

4.4 Type, cotype and UMD

The notions of type/cotype provide a classification of Banach spaces that parallels in many ways the one given by uniform smoothness/uniform convexity. To give a more complete picture for the reader, we feel the need to describe the basic results of that theory, but since it is only loosely related to martingales, we limit ourselves to a survey without proofs (for more detailed information see [207, 230, 46, 206] and also [37, Chapter 9] and [38, Chapter 4]).

Recall our notation $D = \{-1, 1\}^{\mathbb{N}}$ equipped with the uniform probability measure ν. We denote by $\epsilon_n\colon\ D \to \{-1, 1\}$ the n-th coordinate on D, so that the sequence (ϵ_n) is an i.i.d. sequence of symmetric $\{-1, 1\}$-valued random variables. Let B be a Banach space. We will denote simply by $\|\ \ \|_{L_p(B)}$ the "norm" in the space $L_p(D, \nu; B)$, for $0 < p \le \infty$.

Definitions. i) Let $1 \le p \le 2$. A Banach space B is called of type p if there is a constant C such that, for all finite sequences (x_j) in B

$$\left\|\sum \epsilon_j x_j\right\|_{L_2(B)} \le C\left(\sum \|x_j\|^p\right)^{1/p}. \tag{4.37}$$

We denote by $T_p(B)$ the smallest constant C for which (4.37) holds.

ii) Let $2 \le q \le \infty$. A Banach space B is called of cotype q if there is a constant C such that for all finite sequences (x_j) in B

$$\left(\sum \|x_j\|^q\right)^{1/q} \le C\left\|\sum \epsilon_j x_j\right\|_{L_2(B)}. \tag{4.38}$$

We denote by $C_q(B)$ the smallest constant C for which (4.38) holds. Clearly, if $p_1 \le p_2$ then type $p_2 \Rightarrow$ type p_1 while cotype $p_1 \Rightarrow$ cotype p_2. Let us immediately observe that every Banach space is of type 1 and of cotype ∞ with constants equal to 1. In some cases this cannot be improved, for instance if $B = \ell_1$ it is easy to see that (4.37) holds for no $p > 1$. Similarly, if $B = \ell_\infty$ or c_0, then (4.38) holds for no $q < \infty$. We make this more precise in Remark 4.41 below. At the other end of the classification, if B is a Hilbert space then

$$\forall x_1, \dots, x_n \in B \qquad \left\|\sum \epsilon_j x_j\right\|_{L_2(B)} = \left(\sum \|x_j\|^2\right)^{1/2}.$$

Therefore a Hilbert space is of type 2 and cotype 2 (with constants 1). More generally, any space B that is isomorphic to a Hilbert space is of type 2 and cotype 2. It is a striking result of Kwapień [184] that the converse is true: if B is of type 2 and cotype 2, then B must be isomorphic to a Hilbert space.

Actually, by Kahane's inequality (Theoem 8.1), the choice of the norm in $L_2(D, \nu; B)$ plays an inessential rôle in the above definitions. In the case $B = \mathbb{R}$, Kahane's reduces to Khintchine's inequality (8.7). These inequalities make it very easy to analyze the type and cotype of the L_p-spaces:

Proposition 4.33. *If $1 \le p \le 2$, every L_p-space is of type p and of cotype 2. If $2 \le p < \infty$, any L_p-space is of type 2 and of cotype p.*

These are essentially best possible. The space L_∞ contains isometrically any separable Banach space, in particular the already mentioned ℓ_1 and c_0. Therfore, L_∞ is of type 1 and cotype ∞ and nothing more.

Using Kahane's inequality, one can easily generalize the preceding observation.

Proposition 4.34. *Let B be a Banach space of type p and of cotype q. Let (Ω, m) be any measure space. Then $L_r(\Omega, m; B)$ is of type $r \wedge p$ and of cotype $r \vee q$.*

Similar ideas lead to the following result that shows how to use type and cotype to study sums of independent random variables.

Proposition 4.35. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let (Y_n) be a sequence of independent mean zero random variables with values in a Banach space B. Assume that B is of type p and cotype q, and that the series $\sum Y_n$ is a.s. convergent. Then for $0 < r < \infty$, we have*

$$\alpha \mathbb{E}\left(\sum \|Y_n\|^q\right)^{r/q} \le \mathbb{E}\left\|\sum Y_n\right\|^r \le \beta \mathbb{E}\left(\sum \|Y_n\|^p\right)^{r/p}$$

where α and β are positive constants depending only on r, q and B.

Proof. Assume first that each Y_n is symmetric. Consider the sequence $(\epsilon_n Y_n)_{n \ge 1}$ defined on $(D \times \Omega, \nu \times \mathbb{P})$. This sequence has the same distribution as $(Y_n)_{n \ge 1}$. It is therefore easy to deduce Proposition 4.35 in that case from (4.37), (4.38) and Kahane's inequality. The general case follows by an easy symmetrization argument. □

In particular, taking $r = p$ (resp. $r = q$) in Proposition 4.35 we find

$$\mathbb{E}\left\|\sum Y_n\right\|^p \le \beta \sum \mathbb{E}\|Y_n\|^p \tag{4.39}$$

$$\left[\text{resp.} \quad \alpha \sum \mathbb{E}\|Y_n\|^q \le \mathbb{E}\left\|\sum Y_n\right\|^q\right]. \tag{4.40}$$

We now compare type and cotype with the notions introduced in Definitions 4.16 and 4.29.

Proposition 4.36. *Let B be a Banach space. If B is isomorphic to a p-uniformly smooth (resp. q-uniformly convex) space then B is of type p (resp. cotype q).*

Proof. This is an immediate consequence of (4.25) (resp. (4.8)) applied to the martingale $M_n = \sum_1^n \varepsilon_j x_j$. □

The converse to the preceding Proposition is not true in general. This is obvious if we consider only cotype: Indeed L_1 or ℓ_1 is of cotype 2 but being non-reflexive has no equivalent uniformly convex (or smooth) norm.
For type, this is much less obvious, but we will present in Chapter 7 examples of non-reflexive spaces of type 2 and cotype $q > 2$ (see Corollary 7.20). Again, being non-reflexive, these necessarily admit no equivalent uniformly convex (or smooth) norm.

The situation changes dramatically for the class of UMD spaces. In the latter class, the notions we are comparing actually coincide:

Proposition 4.37. *Assume B UMD. Then B is p-uniformly smooth (resp. q-uniformly convex) iff B is of type p (resp. cotype q).*

Proof. Assume B UMD$_p$ so that (8.18) holds. Then if B is of type p we find that (8.18) implies (4.25). The proof for cotype is similar. □

The notions of type and cotype have appeared in various problems involving the analysis of vector valued functions or random variables. One of the great advantages of the classification of Banach spaces in terms of type and cotype is the existence of a rather satisfactory "geometric" characterization of these notions. We first explain the characterization of spaces which have a nontrivial type or a nontrivial cotype. The reader should compare this with the characterizations of super-reflexivity in the next chapter, for instance (i) in Theorem 4.40 below is reminiscent of the equivalence between J-convexity and the existence of $p > 1$ such that (4.25) holds.

Definition. Let $1 \le p \le \infty$. Fix $\lambda > 1$. We say that B contains ℓ_p^n's λ-uniformly if, for all n, there exist $x_1, \cdots, x_n$ in B such that

$$\forall (\alpha_j) \in \mathbb{R}^n \quad \left(\sum |\alpha_j|^p\right)^{1/p} \le \left\|\sum_1^n \alpha_j x_j\right\| \le \lambda \left(\sum |\alpha_j|^p\right)^{1/p}. \tag{4.41}$$

For future reference we recall here a fundamental result (see [135])

Theorem 4.38 (Dvoretzky's Theorem)**.** *For any $\varepsilon > 0$, any infinite dimensional Banach space contains ℓ_2^n's $(1+\varepsilon)$-uniformly.*

Remark 4.39. Krivine proved [181] that if a Banach space B contains ℓ_p^n's $(1+\varepsilon)$-uniformly for some $\varepsilon > 0$ then it also contains them $(1+\varepsilon)$-uniformly for all $\varepsilon > 0$. The cases $p = 1$ and $p = \infty$ (see Theorem 3.5 for that one), go back to James [162]. Therefore, from now on we simply say in that case that B contains ℓ_p^n's uniformly.

Theorem 4.40 ([207])**.** *Let B be a Banach space.*

i) *B is of type p for some $p > 1$ iff B does not contain ℓ_1^n's uniformly.*

ii) *B is of cotype q for some $q < \infty$ iff B does not contain ℓ_∞^n's uniformly.*

Remark 4.41. In such results, the "only if" part is trivial. Indeed, assume (4.41). Then we have

$$n^{1/p} \le \left\|\sum \epsilon_j x_j\right\|_{L_2(B)} \le \lambda n^{1/p},$$

and

$$n^{1/r} \le \left(\sum \|x_j\|^r\right)^{1/r} \le \lambda n^{1/r}.$$

Therefore B cannot be of type $r > p$ or of cotype $r < p$. In particular if $p = 1$ (resp. $r = \infty$) B cannot have a nontrivial type (resp. cotype).

Actually Theorem 4.40 can be extended as follows: Let $1 \le p_0 < 2 < q_0 < \infty$. A space B is of type p for some $p > p_0$ iff B does not contain $\ell_{p_0}^n$'s uniformly. The type and cotype indices are defined as follows:

$$p(B) = \sup\{p \mid B \text{ is of type } p\} \tag{4.42}$$

$$q(B) = \inf\{q \mid B \text{ is of cotype } q\}. \tag{4.43}$$

Corollary 4.42. *If $p(B) > 1$ then $q(B) < \infty$.*
Moreover, $p(B) > 1$ iff $p(B^*) > 1$.

Proof. These statements follow easily from Theorem 4.40. Indeed, if we note that ℓ_1^n embeds isometrically (in the real case) into $\ell_\infty^{2^n}$, we immediately see that B contains ℓ_1^n's uniformly as soon as it contains ℓ_∞^n's uniformly. This shows that $p(B) > 1$ implies $q(B) < \infty$. Similarly, it is easy to see that B contains ℓ_1^n's uniformly iff its dual B^* also does. We leave this as an exercise to the reader (use the fact that it is the same to embed ℓ_1^n in a quotient of B^* or in B^* itself). □

Remark. In addition, it is rather easy to show that B is of type p (resp. cotype q) iff its bidual B^{**} has the same property.

The main theorem relating the type and cotype of B to the geometry of B is

Theorem 4.43 ([207, 181])**.** *Let B be an infinite dimensional Banach space. Then for each $\epsilon > 0, B$ contains ℓ_p^n's $(1+\epsilon)$-uniformly both for $p = p(B)$ and $p = q(B)$.*

By Theorem 4.43 and Remark 4.41, we have

$$p(B) = \inf\{p \mid B \text{ contains } \ell_p^n\text{'s uniformly}\} \tag{4.44}$$

$$q(B) = \sup\{p \mid B \text{ contains } \ell_p^n\text{'s uniformly}\}. \tag{4.45}$$

For classical concrete spaces, the type and cotype has been completely elucidated. For instance, the case of Banach lattices is completely clear, cf. [204]. Here are the main results in that case (which includes Orlicz spaces, Lorentz spaces, etc.). Let us consider a Banach lattice B which is a sublattice of the lattice of all measurable functions on a measure space (Ω, m). Then if $x_1, \ldots, x_n$ are elements of B and if $0 < p < \infty$, the function $(\sum |x_j|^p)^{1/p}$ is well defined as a measurable function and is also in B (by the lattice property).

Maurey proved a Banach lattice generalization of Khintchine's inequality which reduces the study of type and cotype for lattices to some very simple "deterministic" inequalities:

Theorem 4.44 ([204])**.** *Let B be a Banach lattice as above. Assume $q(B) < \infty$. Then there is a constant β depending only on B such that for all $x_1, \ldots, x_n$ in B we have*

$$\frac{1}{\sqrt{2}} \left\| \left(\sum |x_j|^2\right)^{1/2} \right\| \le \left\| \sum \epsilon_j x_j \right\|_{L_2(B)} \le \beta \left\| \left(\sum |x_j|^2\right)^{1/2} \right\|. \tag{4.46}$$

Note: The left side of (4.46) holds in any Banach lattice B; it follows from Khintchine's inequality, see (8.74).

It follows immediately that B (as above) is of type p (resp. cotype q) iff there is a constant C such that any finite sequence (x_j) in B satisfies

$$\left\| \left(\sum |x_j|^2\right)^{1/2} \right\| \le C \left(\sum \|x_j\|^p\right)^{1/p}$$

$$\text{resp.} \qquad \left(\sum \|x_j\|^q\right)^{1/q} \le C \left\| \left(\sum |x_j|^2\right)^{1/2} \right\|.$$

In the case $p < 2$ (or $q > 2$), one can even obtain a much simpler result as shown by the following:

Theorem 4.45 ([204])**.** *Let B be a Banach lattice as above.*

(i) *Let $2 < q < \infty$. Then B is of cotype q iff there is a constant C such that any sequence (x_j) of* disjointly supported *elements of B satisfies*

$$\left(\sum \|x_j\|^q\right)^{1/q} \le C \left\| \sum x_j \right\|.$$

(ii) *Assume $q(B) < \infty$. Let $1 < p < 2$. Then B is of type p iff there is a constant C such that any sequence (x_j) of disjointly supported elements satisfies*

$$\left\| \sum x_j \right\| \le C \left(\sum \|x_j\|^p\right)^{1/p}.$$

Remark. For $q = 2$ (or $p = 2$) the preceding statement is false, the Lorentz spaces $L^{2,1}$ (or $L^{2,q}$ for $2 < q < \infty$) provide counterexamples.

Note that for a disjointly supported sequence (x_j) we have

$$\left\| \sum |x_j| \right\| = \| \sup |x_j| \, \| = \left\| \sum x_j \right\|.$$

Remark. In the particular case of Banach lattices, type and cotype are closely connected to the moduli of uniform smoothness or uniform convexity. This is investigated in great detail in the paper [134].

We should mention that there are several relatively natural spaces for which the type or cotype is not well understood. For instance, by [255] the projective tensor product $\ell_2 \widehat{\otimes} \ell_2$ is of cotype 2, but is unknown whether $\ell_2 \widehat{\otimes} \ell_2 \widehat{\otimes} \ell_2$ is also of cotype 2.

In the rest of this section we briefly review the notion of K-convexity that is the key to the duality between type and cotype. More precisely, let B be a Banach space. We will see below (Proposition 4.46) that if B is of type p, then B^* is of cotype p' with $\frac{1}{p} + \frac{1}{p'} = 1$, the converse fails in general, but it is true if B is a K-convex space. The real meaning of K-convexity was elucidated in [228], where it is proved that a Banach space B is K-convex if (and only if) B does not contain ℓ_1^n's uniformly. The spaces that do not contain ℓ_1^n's uniformly are sometimes called B-convex; so that with this terminology B- and K-convexity are equivalent properties.

We now define K-convexity. We need some notation. We denote by I_B the identity operator on a Banach space B. Let us denote by R_1 the orthogonal projection from $L_2(D,\nu)$ onto the closed span of the sequence $\{\epsilon_n \mid n \in \mathbb{N}\}$. A Banach space B is called K-convex if the operator $R_1 \otimes I_B$ (defined a priori only on $L_2(D,\nu) \otimes B$) extends to a bounded operator from $L_2(D,\nu;B)$ into itself. We will denote by $K(B)$ the norm on $R_1 \otimes I_B$ considered as an operator acting on $L_2(D,\nu;B)$. Clearly $R_1 \otimes I_B$ is bounded on $L_2(B)$ iff $R_1 \otimes I_{B^*}$ is bounded on $L_2(B^*)$. Let us first treat a simple example, the case when $B = \ell_1^k$ with $k = 2^n$. Then, we may isometrically identify B with $L_1(D_n)$ where $D_n = \{-1,+1\}^n$, equipped with its normalized Haar measure. Let us denote by b_j the j-th coordinate on $\{-1,+1\}^n$ considered as an element of $L_1(D_n)$. Consider then the B-valued function $F\colon\ D \to B$ defined by $F(\omega) = \prod_{j=1}^{n}(1+\epsilon_j(\omega)b_j)$. We have $\|F(\omega)\|_B = 1$ hence $\|F\|_{L_2(B)} = 1$. But on the other hand, we have clearly

$$((R_1 \otimes I_B)F)(\omega) = \sum_{j=1}^{n} \epsilon_j(\omega) b_j, \tag{4.47}$$

so that

$$\|(R_1 \otimes I_B)(F)\|_{L_2(B)} = \mathbb{E}\left|\sum_1^n \epsilon_j\right| \geq A_1 n^{1/2} \tag{4.48}$$

for some positive numerical constant A_1. Returning the definition of $K(B)$, we find

$$K(\ell_1^{2^n}) \geq A_1 n^{1/2}.$$

In particular, $K(\ell_1^n)$ is unbounded when $n \to \infty$. From this (and the observation that if S is a closed subspace of B than $K(S) \leq K(B)$) we deduce immediately.

Proposition 4.46. *A K-convex Banach space cannot contain ℓ_1^n's uniformly.*

We now turn to the duality between type and cotype. We first state some simple observations.

Proposition 4.47. *Let B be a Banach space. Let $1 \leq p \leq 2 \leq p' \leq \infty$ be such that $\frac{1}{p} + \frac{1}{p'} = 1$.*

(i) *If B is of type p, then B^* is of cotype p'.*

(ii) *If B is K-convex, and if B^* is of cotype p' then B is of type p.*

To clarify the proof we state the following

Lemma 4.48. *Consider $x_1, \ldots, x_n$ in an arbitrary Banach space B. Define*

$$|||(x_j)||| = \sup\left\{ |\sum_1^n \langle x_j, x_j^*\rangle| \mid x_j^* \in B^* \quad \left\|\sum_1^n \epsilon_j x_j^*\right\|_{L_2(B^*)} \leq 1 \right\}. \tag{4.49}$$

Then

$$|||(x_j)||| = \inf\left\{\left\|\sum_1^n \epsilon_j x_j + \Phi\right\|_{L_2(B)}\right\} \tag{4.50}$$

where the infimum is over all Φ *in* $L_2(B)$ *such that* $\mathbb{E}(\epsilon_j\Phi) = 0$ *for all* $j = 1, 2, \ldots, n$ *(or equivalently over all* Φ *in* $L_2 \otimes B$ *such that* $(R_1 \otimes I_B)(\Phi) = 0$*).*

Proof of Lemma 4.48. We consider the natural duality between $L_2(B)$ and $L_2(B^*)$. Let $S \subset L_2(B^*)$ be the subspace

$$S = \left\{\sum_1^n \epsilon_j x_j^* \mid x_j^* \in B^*\right\}.$$

The norm that appears on the right side of (4.50) is the norm of the space $\mathcal{X} = L_2(B)/S^\perp$. Clearly $\mathcal{X}^* = S^{\perp\perp} = S$. Therefore, the identity (4.50) is nothing but the familiar equality

$$\forall z \in \mathcal{X} \qquad \sup\{|\langle z, z^*\rangle| \mid z^* \in \mathcal{X}^*,\ \|z^*\| \le 1\} = \|x\|. \qquad \square$$

Proof of Proposition 4.47. We leave part (i) as an exercise for the reader. Let us prove (ii). Assume B^* of cotype p' so that $\exists C\ \forall n\ \forall x_j^* \in B^*$

$$\left(\sum \|x_j^*\|^{p'}\right)^{1/p'} \le C\left\|\sum \epsilon_j x_j^*\right\|_{L_2(B^*)}.$$

This implies for all x_j in B

$$|||(x_j)||| \le C\left(\sum \|x_j\|^p\right)^{1/p}.$$

Assume $\sum \|x_j\|^p < 1$. By (4.50) there is a Φ in $L_2(B)$ such that $\mathbb{E}(\epsilon_j\Phi) = 0$ for all j and such that

$$\left\|\sum \epsilon_j x_j + \Phi\right\|_{L_2(B)} < C.$$

We have

$$\sum \epsilon_j x_j = (R_1 \otimes I_B)\left(\sum \epsilon_j x_j + \Phi\right)$$

hence

$$\left\|\sum \epsilon_j x_j\right\|_{L_2(B)} \le K(B)C.$$

By homogeneity, this proves that B is of type p with constant not more than $K(B)C$. $\square$

We come now to the main result of this section which is the converse to Proposition 4.46.

Theorem 4.49. *A Banach space* B *is* K*-convex if (and only if) it does not contain* ℓ_1^n*'s uniformly.*

The projection R_1 can be replaced by all kinds of projections which behave similarly in the preceding statement. For instance, let (g_n) be an i.i.d. sequence of normal Gaussian r.v.'s on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and let G_1 be the orthogonal projection from $L_2(\Omega, \mathcal{A}, \mathbb{P})$ onto the closed span of $\{g_n \mid n \in \mathbb{N}\}$. Then a space B is K-convex iff $G_1 \otimes I_B$ is a bounded operator from $L_2(\Omega, \mathcal{A}, \mathbb{P}; B)$ into itself.

We can proceed similarly in the context of Proposition 4.35, by introducing a projection Q_1 as follows. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. We write simply L_2 for $L_2(\Omega, \mathcal{A}, \mathbb{P})$. Let $(\mathcal{C}_n)_{n\geq 1}$ be a sequence of *independent* σ-subalgebras of $\mathcal{A}$. Let S_0 be the (one dimensional) subspace of L_2 formed by the constant functions. Let S_1 be the subspace formed by all the functions of the form

$$\sum\nolimits_{n=1}^{\infty} y_n$$

with $y_n \in L_2(\mathcal{C}_n)$ for all n, $\mathbb{E}y_n = 0$ and $\sum \mathbb{E}|y_n|^2 < \infty$. We denote by Q_1 the orthogonal projection from L_2 onto S_1. One can then show (see Theorem 4.50 below) that if B is K-convex then $Q_1 \otimes I_B$ is bounded on $L_2(B)$. Note that, in the case $(\Omega, \mathbb{P}) = (D, \nu)$, if we take for $\mathcal{C}_n$ the σ-algebra generated by ϵ_n then Q_1 coincides with R_1.

Let us return to our probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We may as well assume that $\bigcup\limits_n \mathcal{C}_n$ generates the σ-algebra $\mathcal{A}$. Actually we can define a sequence of projections $(Q_k)_{k\geq 0}$ as follows. Let us denote by F_k the closed subspace of L_2 spanned by all the functions f for which there are $n_1 < n_2 < \cdots < n_k$ such that f is measurable with respect to the σ-algebra generated by $\mathcal{C}_{n_1} \cup \cdots \cup \mathcal{C}_{n_k}$.

{Consider the following special case: let (θ_n) be a sequence of independent r.v.'s and let $\mathcal{C}_n$ be the σ-algebra generated by θ_n. Then F_k is the subspace of all the functions in L_2 which depend on at most k of the functions $\{\theta_n \mid n \geq 1\}$.}

Note that $F_k \subset F_{k+1}$ and $\cup F_k$ is dense in L_2. Let then $S_k = F_k \cap F_{k-1}^{\perp}$, and let Q_k be the orthogonal projection from L_2 onto S_k. When $k = 0$, we denote by Q_0 the orthogonal projection onto the subspace of constant functions.

{Note: In the special case considered above, let us denote by λ_n the law of θ_n. Then S_k is the subspace spanned by all the functions of the form $F(\theta_{n_1}, \ldots, \theta_{n_k})$ such that

$$\int F(x_1, \ldots, x_j, \ldots, x_k) d\lambda_{n_j}(x_j) = 0$$

for all $j = 1, 2, \ldots, k$.}

We can now formulate a strengthening of Theorem 4.49.

Theorem 4.50. *Let $(Q_k)_{k\geq 0}$ be as above. If a Banach space B does not contain ℓ_1^n's uniformly then $Q_k \otimes I_B$ defines a bounded operator on $L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$ for $1 < p < \infty$ and any $k \geq 0$. Moreover there is a constant $C = C(p, B)$ such that the norm of $Q_k \otimes I_B$ on $L_p(B)$ satisfies*

$$\|Q_k \otimes I_B\colon\ L_p(B) \to L_p(B)\| \leq C^k \quad \textit{for all} \quad k \geq 0.$$

Clearly Theorem 4.49 is a consequence of Theorem 4.50. The proofs of these results are intimately connected with the theory of *holomorphic semi-groups.*

Since this would take us too far from our main theme, we refer the reader to [228, 229] (or to [206]) for complete proofs and details.

4.5 Square function inequalities in q-uniformly convex and p-uniformly smooth spaces

Let $(f_n)_{n\geq 0}$ be a B-valued martingale in $L_1(B)$. We denote (with the convention $df_0 = f_0$)

$$S_p(f)(\omega) = \left(\sum\nolimits_0^\infty \|df_n(\omega)\|_B^p\right)^{1/p} \quad \text{and} \quad f^*(\omega) = \sup\nolimits_{n\geq 0} \|f_n(\omega)\|_B.$$

Recall that by Doob's inequality (1.20) we have

$$\sup\nolimits_n \|f_n\|_{L_r(B)} \leq \|f^*\|_r \leq r' \sup\nolimits_n \|f_n\|_{L_r(B)}. \tag{4.51}$$

When $p = 2$ and B is either $\mathbb{R}, \mathbb{C}$ or a Hilbert space, we recover the classical square function, see §8.1. In that case, we already know that, for any $1 \leq r < \infty$, $\|S_2(f)\|_r$ and $\|f^*\|_r$ are equivalent, see (8.29) and (8.36). Our main result in this section is an analogue of this for $S_p(f)$ (resp. $S_q(f)$) in case B is p-uniformly smooth (resp. q-uniformly convex). Unfortunately however, we cannot take $p = q$ in general (unless B is a Hilbert space) and hence the analogous inequalities are only one sided, as in the next two statements.

Theorem 4.51. *Let B be a Banach space. Fix $2 \leq q < \infty$. The properties in Corollary 4.7 are equivalent to:*

(iv) *For any $1 \leq r < \infty$, there is a constant $C = C(q, r)$ such that all B-valued martingales $(f_n)_{n\geq 0}$ in $L_r(B)$ satisfy*

$$\|S_q(f)\|_r \leq C\|f^*\|_r. \tag{4.52}$$

Theorem 4.52. *Let B be a Banach space. Fix $1 < p \leq 2$. The properties in Corollary 4.22 are equivalent to:*

(iv) *For any $1 \leq r < \infty$ there is a constant $C' = C'(p, r)$ such that all B-valued martingales $(f_n)_{n\geq 0}$ in $L_r(B)$ satisfy*

$$\|f^*\|_r \leq C'\|S_p(f)\|_r. \tag{4.53}$$

To clarify the duality between (4.52) and (4.53) the following Lemma will be used.

Lemma 4.53. *Let $1 < r, p < \infty$, let $(\mathcal{A}_n)_{n\geq 0}$ be any filtration and set as usual $\mathbb{E}_n = \mathbb{E}^{\mathcal{A}_n}$. Then for any sequence $(\varphi_n)_{n\geq 0}$ in $L_r(\Omega, \mathcal{A}, \mathbb{P})$ we have for any $1 < p < \infty$*

$$\left\|\left(\sum |\mathbb{E}_n\varphi_n|^p\right)^{1/p}\right\|_r \leq C(1/p) \left\|\left(\sum |\varphi_n|^p\right)^{1/p}\right\|_r, \tag{4.54}$$

where $C(1/p) = r(r-1)^{\frac{1}{p}-1}$.

Proof. By Doob's inequality (1.14) (resp. the dual Doob inequality (1.15)) (4.54) holds for $p = \infty$ (resp. $p = 1$) with $C(0) = r' = r(r-1)^{-1}$ (resp. $C(1) = r$). Therefore by the complex interpolation of "mixed normed spaces" (see Theorem **??**), (4.54) holds for a general $1 < p < \infty$ with $C(1/p) = C(0)^{1-\theta}C(1)^{\theta}$ where $\theta = 1/p$. This yields $C(1/p) = (r')^{1/p'}(r)^{1/p} = r(r-1)^{-1/p'}$. □

Proposition 4.54. *Fix $1 < r, r' < \infty$ and $1 < p, p' < \infty$ with $\frac{1}{r} + \frac{1}{r'} = 1$ and $\frac{1}{p} + \frac{1}{p'} = 1$. For a Banach space B and a given filtration $(\mathcal{A}_n)_{n\geq 0}$ on $(\Omega, \mathcal{A}, \mathbb{P})$ the following are equivalent:*

(i) *There is a constant C such that all B-valued martingales $(f_n)_{n\geq 0}$ in $L_r(B)$, adapted to $(\mathcal{A}_n)_{n\geq 0}$, satisfy*

$$\sup \|f_n\|_r \leq C\|S_p(f)\|_r. \tag{4.55}$$

(ii) *There is a constant C' such that all B^*-valued martingales $(g_n)_{n\geq 0}$ in $L_{r'}(B^*)$, adapted to $(\mathcal{A}_n)_{n\geq 0}$, satisfy*

$$\|S_{p'}(g)\|_{r'} \leq C' \sup \|g_n\|_{r'}. \tag{4.56}$$

Moreover, we may exchange the rôles of B and B^* if we wish.

Proof. The proof that (ii) ⇒ (i) is very easy: assuming (ii), for $\varepsilon > 0$, choose g in the unit ball of $L_{r'}(B^*)$ so that $\|f_n\|_{L_r(B)} \leq (1+\varepsilon)\langle g, f_n\rangle$. Let $g_n = \mathbb{E}_n g$. Then note $\langle g, f_n\rangle = \langle g_n, f_n\rangle = \sum_0^n \langle dg_k, df_k\rangle$, it follows $\|f_n\|_{L_r(B)} < (1+\varepsilon)\mathbb{E}|\sum_0^n \langle dg_k, df_k\rangle| \leq (1+\varepsilon)\mathbb{E}(S_{p'}(g)S_p(f)) \leq (1+\varepsilon)\|S_{p'}(g)\|_{r'}\|S_p(f)\|_r \leq (1+\varepsilon)C'\|S_p(f)\|_r$ so we obtain (i) with $C \leq C'$. Conversely, assume (i). To prove (4.56) we may assume $(g_n)_{n\geq 0}$ is a finite martingale so that $g_k = g_n$ for all $k \geq n$. Fix $\varepsilon > 0$. Let $\varphi_0, \ldots, \varphi_n \in L_r(B)$ be such that

$$\left\| \left(\sum\nolimits_0^n \|\varphi_k\|^p\right)^{1/p} \right\|_r \leq 1 \quad \text{and} \quad \mathbb{E}\sum\nolimits_0^n \langle \varphi_k, dg_k\rangle \geq (1+\varepsilon)\|S_{p'}(g)\|_{r'}. \tag{4.57}$$

Note that

$$\mathbb{E}\sum\nolimits_0^n \langle \varphi_k, dg_k\rangle = \mathbb{E}\sum\nolimits_0^n \langle df_k, dg_k\rangle = \mathbb{E}\langle f_n, g_n\rangle \leq \|f_n\|_{L_r(B)}\|g_n\|_{L_{r'}(B^*)}, \tag{4.58}$$

where $df_k = (\mathbb{E}_k - \mathbb{E}_{k-1})\varphi_k$ and $f_n = \sum_0^n df_k$. Moreover, by the triangle inequality and (4.54), we have

$$\|S_p(f)\|_r \leq 2C(1/p) \left\| \left(\sum \|\varphi_k\|^p\right)^{1/p} \right\|_r \leq 2C(1/p). \tag{4.59}$$

Thus we obtain by (4.57), (4.58), (4.55) and (4.59)

$$(1+\varepsilon)\|S_{p'}(g)\|_{r'} \leq \|f_n\|_{L_r(B)}\|g_n\|_{L_{r'}(B^*)} \leq 2CC(1/p)\|g_n\|_{L_{r'}(B^*)},$$

so we obtain (ii) with $C' \leq 2CC(1/p)$. □

It will be convenient to break the proofs of Theorems 4.51 and 4.52 in two. The first parts are formulated in the next two Lemmas.

Lemma 4.55. *Let us denote by* (iv)$_r$ *the assertion* (iv) *in Theorem 4.51 for a fixed value* $1 \le r < \infty$. *Then* (iv)$_s \Rightarrow$ (iv)$_r$ *for any* $1 \le r < s$.

Proof. We will use the "extrapolation method" described in Lemma 8.20 and the B. Davis decomposition in Lemma 8.22. By Lemma 8.22, we have $f_n = g_n + h_n$ with $h_0 = 0$, $\|dg_n\|_B \le 6f^*_{n-1}$ for all $n \ge 1$ and $\left\| \sum_0^\infty \|dh_n\| \right\|_r \le 6r\|f^*\|_r$ for any $1 \le r < \infty$. We set $v_n(\omega) = (\sum_0^n \|dg_k(\omega)\|_B^q)^{1/q}$ and $w_n(\omega) = \|g_n(\omega)\|_B$. Applying (iv)$_s$ to the martingale $(1_{\{T>0\}} g_{n\wedge T})_{n\ge 0}$, we find

$$\|1_{\{T>0\}} v_T\|_s \le C(q,s)\|1_{\{T>0\}} w_T\|_s. \tag{4.60}$$

Fix r such that $1 \le r < s$. By Lemma 8.20

$$\|S_q(g)\|_r \le ((s/(s-r))^{1/r} + 1)\|g^*\|_r + 6\|f^*\|_r. \tag{4.61}$$

But since g is essentially a "perturbation of f by h" we have $S_q(f) \le S_q(g) + \sum \|dh_n\|$ and $g^* \le f^* + \sum \|dh_n\|$, and hence $\|S_q(f)\|_r \le \|S_q(g)\|_r + 6r\|f^*\|_r$ and $\|g^*\|_r \le (1+6r)\|f^*\|_r$, so that (4.61) yields (4.52) with $C(q,r) \le (s((s-r))^{1/r}(1+6r) + 6 + 6r$. □

Lemma 4.56. *Assume B p-uniformly smooth (actually we use only type p). Then there is a constant t_p such that for any $1 \le r < \infty$ and any martingale $(f_n)_{n\ge 0}$ in $L_r(B)$, there is a choice of sign $\xi_n \pm 1$ such that the transformed martingale $\tilde f_n = \sum_0^n \xi_k df_k$ satisfies*

$$\|\tilde f^*\|_r \le t_p \|S_p(f)\|_r. \tag{4.62}$$

Proof. Since B is p-uniformly smooth, a fortiori by Proposition 4.36, it is of type p, i.e. there is a constant C such that for any finite sequence (x_j) in B we have

$$\left\| \sum \varepsilon_j x_j \right\|_{L_r(\nu;B)} \le C \left(\sum \|x_j\|^p \right)^{1/p}.$$

By (1.22) we have

$$\left\| \sup_n \left\| \sum_0^n \varepsilon_j x_j \right\| \right\|_r^r \le 2C^r \left(\sum \|x_j\|^p \right)^{r/p}.$$

Replacing x_j by $df_j(\omega)$ and integrating in ω we find $I \le 2^{1/r} C \|S_p(f)\|_r$, where

$$I = \left(\iint \sup_n \left\| \sum_0^n \varepsilon_j df_j \right\|^r \, d\nu d\mathbb{P} \right)^{1/r}.$$

Thus to conclude it suffices to choose $\xi_j = \varepsilon_j(\omega_0)$ so that $\|\tilde f^*\|_r \le I$ (the latter because the infimum over ω is not more than the average). □

Lemma 4.57. *Let us denote by* (iv)^r *the assertion* (iv) *in Theorem 4.52 for a fixed value of* $1 \le r < \infty$. *Then* (iv)$^s \Rightarrow$ (iv)^r *for all* $1 \le r < s$.

Proof. The idea is the same as for Lemma 4.55 but there is an extra difficulty which is overcome by using Lemma 4.56 above. As earlier, we use the decomposition in Lemma 8.22: we have $f = g + h$ with $h_0 = 0$, $\|dg_n\| \le 6f^*_{n-1}$ for all $n \ge 1$ and $\|\sum \|dh_n\|_B\|_r \le 6r\|f^*\|_r$ for all $1 \le r < \infty$. Let $\xi_n = \pm 1$ be an arbitrary choice of signs. Again we denote

$$\tilde{g}_n = \sum\nolimits_0^n \xi_k dg_k \quad \text{and} \quad \tilde{f}_n = \sum\nolimits_0^n \xi_k df_k.$$

We set $v_n(\omega) = \tilde{g}^*_n(\omega)$ and $w_n(\omega) = (\sum_0^n \|dg_n(\omega)\|^p)^{1/p}$. Assuming (iv)^s, we find for any stopping time T

$$\|1_{\{T>0\}} v_T\|_s \le C'(p,s)\|1_{\{T>0\}} w_T\|_s.$$

Fix r such that $1 \le r \le s$. By (8.32) (applied with $\psi_n = 6f^*_n$) there is a constant C (depending on r and s) such that

$$\|\tilde{g}^*\|_r \le C\|S_p(g)\|_r + 6\|f^*\|_r. \tag{4.63}$$

Since $\|\tilde{f}^*\|_r \le \|\tilde{g}^*\|_r + \big\|\sum \|dh_n\|\big\|_r$ and $\|S_p(g)\|_r \le \|S_p(f)\|_r + \|\sum \|dh_n\|\big\|_r$ we deduce from (4.63)

$$\|\tilde{f}^*\|_r \le C\|S_p(f)\|_r + (6r(C+1)+1)\|f^*\|_r.$$

Since this holds for any choice of signs $\xi_n = \pm 1$ we may exchange the roles of f and $\tilde{f}$ (note that $\tilde{\tilde{f}} = f$!) and we find

$$\|f^*\|_r \le C\|S_p(f)\|_r + (6r(C+1)+1)\|\tilde{f}^*\|_r.$$

If we now choose ξ_n according to Lemma 4.56 we obtain (4.53) with $C' \le C + (6r(C+1)+1)t_p$. □

Proof of Theorem 4.51. The case $1 \le r \le q$ is covered by Lemma 4.55. Recall $\|f^*\|_r \le r' \sup_n \|f_n\|_r$ by Doob's inequality. Let $p = q'$ so that $p^{-1} + q^{-1} = 1$. If $q \le r' < \infty$ then $1 < r \le p$, and, by Proposition 4.54, (4.55) holds at least with $r = p$. By Lemma 4.57, (4.55) holds for all $1 < r < p$, therefore by Proposition 4.54 again, (4.56) (and a fortiori also (4.52)) holds for all r' with $q = p' < r' < \infty$. □

Proof of Theorem 4.52. The argument is the same as for Theorem 4.51: The case $1 \le r \le p$ is covered by Lemma 4.57 and the case $p < r < \infty$ (i.e. $1 < r' < p'$) can be deduced from Lemma 4.55 (applied to $q = p'$) by duality using Proposition 4.54. □

Remark 4.58. The preceding two proofs actually show that if (4.52) (resp. (4.53)) holds for some $1 < r < \infty$ (for some constant C) then it also holds for all $1 < r < \infty$ (with a different constant). In other words, the assertions denoted above by (iv)$_r$ (resp.(iv)^r) are actually independent of $1 < r < \infty$.

4.6 Strong p-variation, uniform convexity and smoothness

We will now extend the method presented in Chapter 6 to the Banach space valued case. The extension to the Hilbert space valued case is straightforward, but the martingale inequalities (4.8) (resp. (4.25)) satisfied by q-uniformly convex (resp. p-uniformly smooth) spaces allow us to go much further:

Theorem 4.59. *Let $1 < p_1 \leq 2 \leq q_0 < \infty$*

(i) *Assume that B is isomorphic to a p_1-uniformly smooth space. Then for all $1 < p < p_1$ there is a constant $C = C(p, p_1)$ such that all B-valued martingales $f = (f_n)_{n \geq 0}$ in $L_p(B)$ satisfy*

$$\mathbb{E}V_p(f)^p \leq C\mathbb{E}\sum\nolimits_0^\infty \|df_n\|_B^p.$$

(ii) *Assume that B is isomorphic to a q_0-uniformly convex space. Then for all $q > q_0$ there is a constant $C = C(q, q_0)$ such that all B-valued martingales in $L_q(B)$ satisfy*

$$\mathbb{E}V_q(f)^q \leq C \sup\nolimits_n \mathbb{E}\|f_n\|^q.$$

The proof is based on the following key fact:

Lemma 4.60. *Let $1 < r < \infty$ and let $0 < \theta < 1$ be such that $1 - \theta = \frac{1}{r}$. Let $(f_n)_{n \geq 0}$ be a B-valued martingale converging in $L_r(B)$. Assume that for all increasing sequences of stopping times $0 \leq T_0 \leq T_1 \leq T_2 \leq \ldots$ we have*

$$\mathbb{E}\|f_{T_0}\|^r + \sum\nolimits_{k \geq 1} \mathbb{E}\|f_{T_k} - f_{T_{k-1}}\|^r \leq 1.$$

Then

$$\|\{f_n\}\|_{(L_1(v_1(B)), L_\infty(\ell_\infty(B)))_{\theta,\infty}} \leq 2.$$

Proof. This can be proved by an obvious adaptation of the argument for Lemma 6.3. One just chooses $T_k = \inf\{n > T_{k-1} \mid \|f_n - f_{T_{k-1}}\| > t^{\theta - 1}\}$. $\square$

We will use repeatedly the identity (see Theorem 5.7)

$$L_p((B_0, B_1)_{\theta,p}) = (L_1(B_0), L_\infty(B_1))_{\theta,p} \tag{4.64}$$

valid for any $0 < \theta < 1$ provided p is linked to θ by $1 - \theta = 1/p$.

Proof of Theorem 4.59. Here again we can adapt the proof of Theorem 6.2.
(i) Assume B p_1-uniformly smooth. Then by (4.8) applied (with p replaced by p_1) to the martingale $M_n = f_{T_k \wedge n} - f_{T_{k-1} \wedge n}$ (here k is fixed) we have (for some constant C_1)

$$\mathbb{E}\|f_{T_k} - f_{T_{k-1}}\|^{p_1} \leq C_1 \sum_{T_{k-1} < n \leq T_k} \mathbb{E}\|f_n - f_{n-1}\|^{p_1}$$

and hence

$$\sum\nolimits_0^\infty \mathbb{E}\|f_{T_k} - f_{T_{k-1}}\|^{p_1} \le C_0 \sum\nolimits_0^\infty \mathbb{E}\|df_n\|^{p_1}.$$

Let $\|f\|_{D(p_1)} = \left(\sum_0^\infty \mathbb{E}\|df_n\|^{p_1}\right)^{1/p_1}$. Let θ_1 be such that $1 - \theta_1 = 1/p_1$. On one hand, by the preceding Lemma we have a bounded inclusion

$$D(p_1) \subset (L_1(v_1(B)), L_\infty(\ell_\infty(B)))_{\theta_1,\infty} \tag{4.65}$$

and on the other hand we have trivially (actually this is an equality)

$$D(1) \subset L_1(v_1(B)). \tag{4.66}$$

By the same argument as in Chapter 6 we know (see (6.14)) that $D(p_1) = (D(1), D(\infty))_{\theta_1,p_1}$ where we set $D(\infty) = \ell_\infty(L_\infty(B))$. Therefore, by the reiteration Theorem, (4.65) and (4.66) imply that, for any θ with $0 < \theta < \theta_1$ and any $1 \le p \le \infty$, we have

$$(D(1), D(\infty))_{\theta,p} \subset (L_1(v_1(B)), \ell_\infty(L_\infty(B)))_{\theta,p}.$$

We now choose p so that $1 - \theta = 1/p$. This gives us $(D(1), D(\infty))_{\theta,p} = D(p)$ and also by (4.64) (see (6.3))

$$(L_1(v_1(B)), \ell_\infty(L_\infty(B)))_{\theta,p} = L_p((v_1(B), \ell_\infty(B))_{\theta,p})$$

but by Lemma 6.1

$$(v_1(B), \ell_\infty(B))_{\theta,p} \subset v_p(B).$$

Thus we obtain that the inclusion

$$D(p) \subset L_p(v_p(B))$$

is bounded, and this is precisely (i).

To prove (ii) assume B q_0-uniformly convex. Let θ_0 be such that $1 - \theta_0 = 1/q_0$. By (4.25) applied to the martingale $(f_{T_k})_{k\ge0}$ we have for some constant C_0

$$\sum \mathbb{E}\|f_{T_k} - f_{T_{k-1}}\|^{q_0} \le C_0 \sup \mathbb{E}\|f_n\|^{q_0}.$$

Let $L_{q_0}(B) \to \ell_\infty(L_{q_0}(B))$ be defined by $T(M) = (\mathbb{E}_n M - \mathbb{E}_{n-1} M)_{n\ge0}$. By the preceding Lemma we have (boundedly)

$$T(L_{q_0}(B)) \subset (L_1(v_1), L_\infty(\ell_\infty(B)))_{\theta_0,\infty} \tag{4.67}$$

and trivially

$$T(L_\infty(B)) \subset \ell_\infty(L_\infty(B)). \tag{4.68}$$

Observe that $L_{q_0}(B) = (L_1(B), L_\infty(B))_{\theta_0,q_0}$ (see (4.64)). Therefore, by the reiteration Theorem, (4.67) and (4.68) imply

$$T((L_1(B), L_\infty(B))_{\theta,q}) \subset (L_1(v_1), \ell_\infty(L_\infty(B)))_{\theta,q}$$

for any θ with $\theta_0 < \theta < 1$ and any $1 \le q \le \infty$. If we choose q so that $1-\theta = 1/q$ we find by (4.64)

$$T(L_q(B)) \subset L_q((v_1(B), \ell_\infty(B))_{\theta,q})$$

and again $(v_1(B), \ell_\infty(B))_{\theta,q} \subset v_q(B)$ so we obtain

$$T(L_q(B)) \subset L_q(v_q(B))$$

which is exactly (ii). □

Remark. In the situation of Theorem 4.59, fix $1 < p < p_1$ (resp. $q > q_0$). Then for each $1 \le r \le p$ (resp. $1 \le r \le q$) there is a constant C such that

$$\|V_p(f)\|_r \le C \left\| \left(\sum \|df_n\|_B^p \right)^{1/p} \right\|_r$$

(resp. $\|V_q(f)\|_r \le C \| \sup \|f_n\|_B \|_r$ and $\|V_q(f)\|_{1,\infty} \le C \sup \|f_n\|_{L_1(B)}$). Indeed, this can be proved exactly as above in the proofs of part (ii) in Theorems 6.2 and 6.5.

4.7 Notes and Remarks

The source of this chapter is mainly [227], but the latter paper was inspired by Enflo's fundamental results on super-reflexivity that are described in detail in the next chapter. Enflo's main result from [131] was that "super-reflexive" implies "isomorphic to uniformly convex," thus completing a program initiated by R.C. James ([162, 163]), that we describe in the notes and remarks of the next chapter. While Enflo and James work with "trees," in [227] the relevance of martingales was recognized and a new proof, was given of Enflo's theorem with an improvement: the modulus of convexity can always be found of power type, or equivalently we can always find a renorming satisfying (4.1).

In our presentation, we find it preferable to separate the two steps: in this chapter we show that any uniformly convex is isomorphic to a space with a modulus "of power type" (i.e. satisfying (4.1)) and only in the next one do we show Enflo's result that "super-reflexive" implies "isomorphic to uniformly convex".

In both chapters, we replace the Banach space B by $L_q(B)$ with $1 < q < \infty$ and we treat martingale difference sequences simply as monotone basic sequences in $L_q(B)$. The corresponding inequalities for basic sequences in uniformly convex (resp. smooth) spaces are due to the Gurarii brothers [155] (resp. to Lindenstrauss [189]).

We learnt about the work of Fortet–Mourier through unpublished work by J. Hoffmann–Jørgensen. The presentation in §4.2 was strongly influenced by [4], to which we refer the reader interested in non-linear aspects of Banach space theory. The estimates for the modulus of convexity (and of smoothness) of L_p in §4.3 are due to O. Hanner [158]. See also [71] for more recent results including the non-commutative case. The results of §4.6 come essentially from [227], while those of §4.5 come from [236].

Chapter 5

The Real Interpolation method

We assume in this chapter that the reader has some familiarity with the complex method of interpolation, or at least with the famous Riesz interpolation theorem. We will mainly use the real method of interpolation in our later exposition to analyze the type and cotype of the spaces of sequences that are in the interpolated space between "bounded variation" and "bounded".

Roughly the common interpolation methods produce a family of "interpolated" Banach spaces $(B_\theta)_{\theta\in[0,1]}$ starting from a pair (B_0, B_1). We will need to assume that the initial pair (B_0, B_1) is "compatible". This means that we are given a topological vector space V and continuous injections

$$j_0\colon\ B_0 \to V \quad \text{and} \quad j_1\colon\ B_1 \to V.$$

This very rudimentary structure is just what is needed to define the intersection $B_0 \cap B_1$ and the sum $B_0 + B_1$.
The space $B_0 \cap B_1$ is defined as $j_0(B_0) \cap j_1(B_1)$ equipped with the norm

$$\|x\| = \max\{\|j_0^{-1}(x)\|_{B_0}, \|j_1^{-1}(x)\|_{B_1}\}.$$

The space $B_0 + B_1$ is defined as the setwise sum $j_0(B_0) + j_1(B_1)$ equipped with the norm

$$\|x\|_{B_0+B_1} = \inf\{\|x_0\|_{B_0} + \|x_1\|_{B_1} \mid x = j_0(x_0) + j_1(x_1)\}.$$

It is an easy exercise to check that $B_0 \cap B_1$ and $B_0 + B_1$ are Banach spaces. It is a well established tradition to identify B_0 and B_1 with $j_0(B_0)$ and $j_1(B_1)$, so that j_0 and j_1 become the inclusion mappings $B_0 \subset V$ and $B_1 \subset V$. We then have $\forall i = 0, 1$

$$B_0 \cap B_1 \subset B_i \subset B_0 + B_1$$

and these inclusions have norm ≤ 1. Note that if we wish we may now replace V by $B_0 + B_1$, so that we may as well assume that V is a Banach space.

5.1 The real interpolation method

Let (B_0, B_1) be a compatible couple. For any $\epsilon > 0$ and any x in $B_0 + B_1$, we define

$$K_t(x; B_0, B_1) = \inf\{\|b_0\|_{B_0} + t\|b_1\|_{B_1} \mid x = b_0 + b_1\}.$$

We will often abbreviate and write simply $K_t(x)$ instead of $K_t(x; B_0, B_1)$ when the context leaves no room for ambiguity. Let $0 < \theta < 1$ and $1 \le p \le \infty$. We define

$$(B_0, B_1)_{\theta,q} = \left\{ x \in B_0 + B_1 \;\Big|\; \int_0^\infty (t^{-\theta} K_t(x))^q \frac{dt}{t} < \infty \right\}$$

and we equip it with the norm

$$\|x\|_{(B_0,B_1)_{\theta,q}} = \left(\int_0^\infty (t^{-\theta} K_t(x))^q \frac{dt}{t} \right)^{1/q}.$$

Of course, when $q = \infty$, this should be understood as meaning $\sup\limits_t t^{-\theta} K_t(x)$.

Note

$$K_t(x; B_0, B_1) = t K_{t^{-1}}(x; B_1, B_0)$$

and hence

$$(5.1) \qquad (B_0, B_1)_{\theta,q} = (B_1, B_0)_{1-\theta,q} \quad \text{isometrically.}$$

Remark 5.1. Obviously we have inclusions (with norms at most 1)

$$B_0 \cap B_1 \subset (B_0, B_1)_{\theta,q} \subset B_0 + B_1.$$

Moreover, it is easy to show that if $q < \infty$ $B_0 \cap B_1$ is dense in $(B_0, B_1)_{\theta,q}$.

Note that $t \to K_t(x)$ is by definition the infimum of a family of affine functions, hence it is concave on $\mathbb{R}_+$, nonnegative and nondecreasing.

Since K_t is nondecreasing

$$\theta^{-1} t^{-\theta} K_t(x) = K_t(x) \int_t^\infty s^{-\theta} ds/s \le \int_t^\infty s^{-\theta} K_s(x) ds/s,$$

and hence

$$(5.2) \qquad (B_0, B_1)_{\theta,1} \subset (B_0, B_1)_{\theta,\infty}.$$

More generally, for any $q_0 \le q_1$ we have

$$(5.3) \qquad (B_0, B_1)_{\theta,q_0} \subset (B_0, B_1)_{\theta,q_1}.$$

If we assume $B_0 \subset B_1$, then it is easy to check that, when $0 < \theta_0 < \theta_1 < 1$, for arbitrary $1 \le q_0, q_1 \le \infty$, we have bounded inclusions

$$(5.4) \qquad B_0 \subset (B_0, B_1)_{\theta_0,q_0} \subset (B_0, B_1)_{\theta_1,q_1} \subset B_1.$$

Just as for the complex case, the fundamental interpolation property holds:

Theorem 5.2. *Let (B_0, B_1) and (C_0, C_1) be two compatible couples. Let $T_0\colon\ B_0 \to C_0$ and $T_1\colon\ B_1 \to C_1$ be bounded operators that are "essentially the same". Then the resulting operator $T\colon\ B_0 + B_1 \to C_0 + C_1$ maps $(B_0, B_1)_{\theta,q}$ to $(C_0, C_1)_{\theta,q}$ for any $0 < \theta < 1$ and $1 \le p \le \infty$, and moreover, if we denote its restriction by $T_{\theta,q}\colon\ (B_0, B_1)_{\theta,q} \to (C_0, C_1)_{\theta,q}$, have*

$$\|T_{\theta,q}\| \le \|T_0\|^{1-\theta}\|T_1\|^{\theta}.$$

Proof. We obviously can write for any x in $B_0 + B_1$, say $x = x_0 + x_1$ with $x_j \in B_j$

$$Tx = T_0x_0 + T_1x_1$$

and hence

$$\|T_0x_0\| + t\|T_0\|\|T_1\|^{-1}\|T_1x_1\| \le \|T_0\|(\|x_0\|_{B_0} + t\|x_1\|_{B_1})$$

so that

$$K_{t\|T_0\|\|T_1\|^{-1}}(Tx) \le \|T_0\|K_t(x). \tag{5.5}$$

Let $\lambda = \|T_0\|\|T_1\|^{-1}$. Since $\frac{dt}{t}$ is a Haar measure over the multiplicative group $(0, \infty)$, we have by (5.5)

$$\begin{aligned}\|t^{-\theta}K_t(Tx)\|_{L^p(\frac{dt}{t})} &= \|(t\lambda)^{-\theta}K_{t\lambda}(Tx)\|_{L^p(\frac{dt}{t})}\\ &\le \lambda^{-\theta}\|t^{-\theta}\|T_0\|K_t(x)\|_{L^p(\frac{dt}{t})}\\ &\le \|T_0\|\lambda^{-\theta}\|t^{-\theta}K_t(x)\|_{L^p(\frac{dt}{t})},\end{aligned}$$

and hence

$$\|T_{\theta,q}\| \le \|T_0\|\lambda^{-\theta} = \|T_0\|^{1-\theta}\|T_1\|^{\theta}.$$

□

The fundamental example is the case of L_p-spaces: Let $(\Omega, \mathcal{A}, m)$ be a measure space, and let $f\colon\ \Omega \to \mathbb{R}$ be a measurable function. We define its decreasing rearrangement $f^*\colon\ (0, \infty) \to \mathbb{R}_+$ by setting

$$f^*(t) = \inf\{c > 0 \mid m(\{|f| > c\}) \le t\}.$$

Then $f^* \ge 0$ is non-increasing, right continuous and such that

$$|\{f^* > c\}| = m(\{|f| > c\}). \tag{5.6}$$

The latter equality shows that f^* and $|f|$ have the same distribution relative respectively to Lebesgue measure on $(0, \infty)$ and m. Recall that

$$\int |f|^p dm = \int_0^\infty pc^{p-1}m(\{|f| > c\})\ dc.$$

As an immediate consequence of (5.6) we have in particular

$$\forall p > 0 \qquad \int_0^\infty f^*(t)^p dt = \int |f|^p \, dm.$$

More generally, the Lorentz spaces $L_{p,q}(\Omega, m)$ (or simply $L_{p,q}$) are defined ($0 < p, q < \infty$) as formed of the functions f such that

$$\int_0^\infty (t^{1/p} f^*(t))^q \frac{dt}{t} < \infty$$

equipped with the quasi-norm

$$\|f\|_{p,q} = \left(\int_0^\infty (t^{1/p} f^*(t))^q \frac{dt}{t} \right)^{1/q}. \tag{5.7}$$

Note that $L_{p,p} = L_p$ isometrically. When $q = \infty$, the above should be understood as

$$\|f\|_{p,\infty} = \sup_{t>0} t^{1/p} f^*(t) = (\sup_{c>0} c^p m(\{|f| > c\}))^{1/p}.$$

The space $L_{p,\infty}$ is usually called "weak L_p".

Theorem 5.3. *Let $L_p = L_p(\Omega, \mathcal{A}, m)$ on an arbitrary measure space. Consider $f \in L_1 + L_\infty$ ($0 < p < \infty$). Then*

$$K_t(f; L_1, L_\infty) = \int_0^t f^*(s) ds.$$

Consequently, for any $1 \le q \le \infty$

$$(L_1, L_\infty)_{\theta,q} = L_{p,q}$$

where $\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{\infty}$, with equivalent quasi-norms. This shows that (5.7) is equivalent to a norm. In particular $(L_1, L_\infty)_{\theta,p} = L_p$.

Proof. Let $w(x) = f(x)|f(x)|^{-1}$ (sign of $f(x)$). Fix $t > 0$. Let

$$\begin{aligned} f_0 &= 1_{\{|f|>f^*(t)\}}(f - f^*(t)w) = w 1_{\{|f|>f^*(t)\}}(|f| - f^*(t)) \\ f_1 &= f - f_0. \end{aligned}$$

Note $|f_1| = |f| \wedge f^*(t)$. Then let $\Omega_t = \{|f| > f^*(t)\}$. We have

$$\begin{aligned} K_t(f; L_1, L_\infty) &\le \|f_0\|_1 + t\|f_1\|_\infty \\ &\le \int_{\Omega_t} [|f| - f^*(t)] \, dm + t f^*(t) \\ &= \int_0^{m(\Omega_t)} (f^*(s) - f^*(t)) \, ds + t f^*(t) \end{aligned}$$

and hence since $m(\Omega_t) \le t$ and $f^*(s) = f^*(t)$ on $[m(\Omega_t), t]$, we obtain

$$K_t(f) \le \int_0^t f^*(s)\ ds.$$

Conversely, assume that $f = f_0 + f_1,\ f_0 \in L_1,\ f_1 \in L_\infty$. Clearly $m(\{|f| > c_0 + c_1\}) \le m(\{|f_0| > c_0\}) + m(\{|f_1| > c_1\})$, and hence for any $0 < \varepsilon < 1$

$$f^*(s) \le f_0^*((1-\varepsilon)s) + f_1^*(\varepsilon s),$$

so that

$$\begin{aligned}\int_0^t f^*(s)ds &\le \int_0^t f_0^*((1-\varepsilon)s)\ ds + \int_0^t f_1^*(\varepsilon s)ds\\ &\le \int_0^t f_0^*((1-\varepsilon)s)ds + tf_1^*(0)\\ &\le (1-\varepsilon)^{-1}\|f_0^*\|_1 + t\|f_1^*\|_\infty = (1-\varepsilon)^{-1}\|f_0\|_1 + t\|f_1\|_\infty.\end{aligned}$$

Taking the limit when $\varepsilon \to 0$ and the infimum over f_0, f_1 yields

$$\int_0^t f^*(s)\ ds \le K_t(f; L_1, L_\infty).$$

To complete the proof it suffices to prove

$$\|f\|_{p,q} \le \|f\|_{(L_1,L_\infty)_{\theta,q}} \le \theta^{-1}\|f\|_{p,q}, \tag{5.8}$$

where $1-\theta = p^{-1}$ (and hence $\theta^{-1} = p'$).
Let $f^{**}(t) = t^{-1}\int_0^t f^*(s)ds$. Note that $f^*(t) \le f^{**}(t)$, and hence

$$\|f\|_{p,q} = \|t^{1-\theta}f^*\|_{L_q(\frac{dt}{t})} \le \|t^{1-\theta}f^{**}\|_{L_q(t^{-1}dt)}.$$

For the converse direction, we write $\int_0^t f^*(s)ds = \int_0^1 f^*(st)t\ ds$, so that $f^{**}(t) = \int_0^1 f^*(st)ds$. Then by Jensen's inequality (since $q \ge 1$) we have

$$\begin{aligned}\|t^{1-\theta}f^{**}\|_{L_q(t^{-1}dt)} &\le \int_0^1 \|t^{1-\theta}f^*(st)\|_{L_q(t^{-1}dt)}ds\\ &= \|t^{1-\theta}f^*(t)\|_{L_q(t^{-1}dt)}\int_0^1 s^{\theta-1}ds = \theta^{-1}\|f\|_{p,q},\end{aligned}$$

which proves (5.8).

□

Remark. With the preceding notation, since (5.7) is equivalent to a norm, for some constant $C = C(p,q)$ we have:

$$\|f^*\|_{p,q} \le \|f^{**}\|_{p,q} \le C\|f^*\|_{p,q}\ , \tag{5.9}$$

where $\|\ \|_{p,q}$ denotes the quasi-norm in $L_{p,q}((0,\infty), dt)$. Indeed, this can be verified by Jensen's inequality as we just did to prove (5.8).

Lemma 5.4. *If μ is non-atomic, for any f in $L_1 + L_\infty$ we have*

$$\int_0^t f^*(s)ds = \sup\left\{ \int_E |f|d\mu \mid \mu(E) = t \right\}.$$

Proof. It is easy to check that $(1_E f)^*(s) = 0$ for all $s > \mu(E)$ and also that $(1_E f)^*(s) \leq f^*(s)$ for all $s > 0$. Therefore,

$$\int_E |f|d\mu = \int_0^\infty (1_E f)^*(s)ds \leq \int_0^t f^*(s)ds.$$

This yields

$$\sup_{\mu(E)=t} \int_E |f|d\mu \leq \int_0^t f^*(s)ds. \tag{5.10}$$

If $\mu\{|f| = f^*(t)\} = 0$, the converse inequality is easy: we have $|\{f^* > f^*(t)\}| = \mu(\{|f| > f^*(t)\}) = t$ and hence the choice of $E = \{|f| > f^*(t)\}$ shows that (5.10) is an equality. If $\mu\{|f| = f^*(t)\} > 0$, a little more care is needed. We have $\mu(\{|f| > f^*(t)\}) \leq t \leq \mu(\{|f| \geq f^*(t)\})$. We will use the assumption that μ is non-atomic to select a set E such that $\{|f| > f^*(t)\} \subset E \subset \{|f| \geq f^*(t)\}$, with $\mu(E) = t$. Let $t' = \mu\{|f| > f^*(t)\}$. Since $|f|$ and f^* have the same distribution and $\{f^* > f^*(t)\} = [0, t')$, we have then

$$\int_{\{|f|>f^*(t)\}} |f|d\mu = \int_0^{t'} f^* ds$$

and hence

$$\int_E |f|d\mu \leq \int_{\{|f|>f^*(t)\}} |f|d\mu + (t-t')f^*(t) = \int_0^{t'} f^* ds + (t-t')f^*(t) = \int_0^t f^*(s)ds.$$

□

More generally, using a suitable version of the reiteration theorem for the real method we have

Theorem 5.5. *Consider $0 < p_0, q_0, p_1, q_1 \leq \infty$. Assume $p_0 \neq p_1$. Then, for any $0 < \theta < 1$ and $0 < q \leq \infty$,*

$$(L_{p_0,q_0}, L_{p_1,q_1})_{\theta,q} = L_{p_\theta,q}$$

with equivalent norms, where

$$1/p_\theta = (1-\theta)/p_0 + \theta/p_1.$$

In particular, $(L_{p_0}, L_{p_1})_{\theta,q} = L_{p_\theta,q}$, and the latter space coincides with L_{p_θ} if $q = p_\theta$. Moreover, if $p_0 = p_1 = p$ then we have

$$(L_{p,q_0}, L_{p,q_1})_\theta = L_{p,q_\theta}$$

with equivalent norms where $1/q_\theta = (1-\theta)/q_0 + \theta/q_1$.

Remark. Historically, the preceding result was inspired by, and appears as an abstract version of the Marcinkiewicz interpolation Theorem (see Theorem 8.51 above). It implies it as an easy corollary: if an operator is bounded both from L_{p_0} to $L_{p_0,\infty}$ and from L_{p_1} to $L_{p_1,\infty}$ $(p_0 \neq p_1)$, then it is bounded from $(L_{p_0}, L_{p_1})_{\theta,p}$ to $(L_{p_0,\infty}, L_{p_1,\infty})_{\theta,p}$, and hence, choosing $p = p_\theta$, we conclude, by the preceding Theorem that it is bounded from L_{p_θ} to itself.

Remark 5.6. Let (B_0, B_1) be a compatible pair. Let $0 < \theta < 1$, $1 \le q \le \infty$ and let p be determined by $p^{-1} = 1 - \theta$. Consider $x \in B_0 + B_1$ and let $f_x(t) = t^{-1} K_t(x; B_0, B_1)$. It is not difficult to check that there is a positive constant $C = C(p, q)$ such that for any x in $B_0 + B_1$ we have

$$C^{-1}\|f_x\|_{L_{p,q}([0,\infty))} \le \|x\|_{(B_0,B_1)_{\theta,q}} \le C\|f_x\|_{L_{p,q}([0,\infty))}.$$

Indeed, assume for simplicity that $t \mapsto K_t(x)$ is differentiable with derivative $K_t'(x)$. Then since $t \mapsto K_t(x)$ is concave, $K_t'(x)$ is non-increasing, so this equivalence follows from (5.9) with $t \mapsto K_t'(x)$ in place of f^*.

We will use several times the following result.

Theorem 5.7. *Let $1 \le p < \infty$ and $0 < \theta < 1$. Let (B_0, B_1) be a compatible pair and let (Ω, μ) be any measure space.*

(i) Then

$$(L_p(\mu; B_0), L_p(\mu; B_1))_{\theta,p} = L_p(\mu; (B_0, B_1)_{\theta,p}) \tag{5.11}$$

with equivalent norms.

(ii) More generally, if $1 \le p_0 \neq p_1 \le \infty$ are such that $\frac{1-\theta}{p_0} + \frac{\theta}{p_1} = \frac{1}{p}$, then

$$(L_{p_0}(\mu; B_0), L_{p_1}(\mu; B_1))_{\theta,p} = L_p(\mu; (B_0, B_1)_{\theta,p}) \tag{5.12}$$

with equivalent norms.

Proof. The proof of (i) is rather easy. For simplicity we write $L_p(B)$ instead of $L_p(\mu; B)$. Let $f \in L_p(B_0) + L_p(B_1)$. We will show

$$\begin{aligned} 2^{-1/p'} K_t(f; L_p(B_0), L_p(B_1)) &\le \left(\int K_t(f(\omega); B_0, B_1)^p d\mu(\omega) \right)^{1/p} \\ &\le K_t(f; L_p(B_0), L_p(B_1)). \end{aligned} \tag{5.13}$$

Indeed, if $f = f_0 + f_1$ with $f_j \in L_p(B_j)$ $(j = 0, 1)$ then

$$K_t(f(\omega); B_0, B_1) \le \|f_0(\omega)\|_{B_0} + t\|f_1(\omega)\|_{B_1}$$

from which the second inequality in (5.13) is immediate. To prove the first inequality, fix $\varepsilon > 0$, and let $f(\omega) = f_0(\omega) + f_1(\omega)$ be such that f_0, f_1 are Bochner measurable and such that $\|f_0(\omega)\|_{B_0} + t\|f_1(\omega)\|_{B_1} \le (1 + \varepsilon) K_t(f(\omega); B_0, B_1)$. We have then

$$\big\| \|f_0(\cdot)\|_{B_0} + t\|f_1(\cdot)\|_{B_1} \big\|_p \le (1 + \varepsilon) \|K_t(f(\cdot); B_0, B_1)\|_p,$$

and then using

$$\begin{aligned}\|f_0\|_{L_p(B_0)} + t\|f_1\|_{L_p(B_1)} &\le 2^{1/p'}\|(\|f_0(\cdot)\|_{B_0}^p + \|tf_1(\cdot)\|_{B_1}^p)^{1/p}\|_p \\ &\le 2^{1/p'}\big\|(\|f_0(\cdot)\|_{B_0} + t\|f_1(\cdot)\|_{B_1})\big\|_p\end{aligned}$$

the first inequality in (5.13) follows immediately. Clearly, (5.13) implies (5.11) by integration. □

Remark. More generally, the same argument yields that for any $q \ge p$ (resp. $q \le p$) we have a bounded inclusion

$$(L_p(B_0), L_p(B_1))_{\theta,q} \supset L_p((B_0, B_1)_{\theta,q})$$
$$(\text{resp. } (L_p(B_0), L_p(B_1))_{\theta,q} \subset L_p((B_0, B_1)_{\theta,q})).$$

This follows again by integration but using the fact ("Hölder–Minkowski") that $L_q(\frac{dt}{t}, L_p) \supset L_p(L_q(\frac{dt}{t}))$ (resp. $L_q(\frac{dt}{t}, L_p) \subset L_p(L_q(\frac{dt}{t}))$). See [119] for more on this.

Remark 5.8. When $p_1 = \infty$, (5.12) becomes

$$(L_{p_0}(B_0), L_\infty(B_1))_{\theta,p} = L_p((B_0, B_1)_{\theta,p}). \tag{5.14}$$

Recall however that $L_\infty(B_1)$ is defined as the space of essentially bounded *Bochner measurable* B_1-valued functions. This is rather restrictive in certain "concrete" situations. To extend the scope of (5.14) we record here a simple observation: Assume $L_\infty(B_1)$ isometrically embedded in an priori larger space $\mathcal{L}$ of B_1-valued functions (or classes of functions), for instance $\mathcal{L} = \Lambda_\infty(B_1)$. Intuitively, $\mathcal{L}$ is formed of bounded B_1-valued functions but measurable in a broader sense, and we assume that $L_\infty(B_1) \subset \mathcal{L}$ is formed of those elements in $\mathcal{L}$ that are Bochner measurable. Assume $B_0 \subset B_1$. Then, for any x that is a Bochner measurable B_1-valued function we have

$$\forall t > 0 \qquad K_t(x; L_{p_0}(B_0), L_\infty(B_1)) = K_t(x; L_{p_0}(B_0), \mathcal{L}).$$

Indeed, if $x = x_0 + x_1$ with $x_0 \in L_{p_0}(B_0)$ and $x_1 \in \mathcal{L}$, then a fortiori $x_0 \in L_{p_0}(B_1)$, so that $x_1 = x - x_0$ is Bochner-measurable as a B_1-valued function and hence automatically in $L_\infty(B_1)$. Consequently, the norms of such an x in the (θ, q) interpolated spaces is the same for the two pairs $(L_{p_0}(B_0), L_\infty(B_1))$ and $(L_{p_0}(B_0), \mathcal{L})$. We will use this for the following example: $B_1 = L_\infty(\ell_\infty(B))$ and $\mathcal{L} = \ell_\infty(L_\infty(B))$. Note that with our (Bochner sense) definition of $L_\infty(B)$, when we take as measure space $\mathbb{N}$ equipped with the counting measure, the space $L_\infty(B)$ is in general smaller than $\ell_\infty(B)$, but the latter coincides in that case with $\Lambda_\infty(B)$.

We will use later (especially in Chapter 6) the real interpolation analogue of the reiteration theorem, as follows (cf. [5, p. 50]).

Theorem 5.9. *Let (B_0, B_1) be a compatible couple of Banach spaces. Let $0 < \theta_0 \neq \theta_1 < 1$ and $1 \leq q_0, q_1 \leq \infty$. Consider the couple X_0, X_1 where*

$$X_j = (B_0, B_1)_{\theta_j, q_j} \qquad j = 0, 1.$$

Then for any $0 < \theta < 1$ and $1 \leq q \leq \infty$ we have

$$(X_0, X_1)_{\theta,q} = (B_0, B_1)_{\tau,q}$$

(with equivalent norms) where $0 < \tau < 1$ is determined by $(1-\theta)\theta_0 + \theta\theta_1 = \tau$.

Remark 5.10. The reiteration theorem can be viewed as an "abstract" version of the Marcinkiewicz theorem: Indeed this implies that if $1 \leq p_0 < p_1 < \infty$ and $1 \leq q \leq \infty$ we have

$$(L_{p_0\infty}, L_{p_1\infty})_{\theta,q} = (L_{p_0,1}, L_{p_1 1})_{\theta,q} = (L_{p_0}, L_{p_1})_{\theta,q}.$$

Therefore, if $T\colon\ L_{p_j} \to L_{p_j\infty}$ is bounded for $j = 0, 1$, it must be also bounded from $(L_{p_0}, L_{p_1})_{\theta,q}$ to itself for any $0 < \theta < 1$ and any $1 \leq q \leq \infty$. Choosing q so that $\frac{1}{q} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$ we find that T is bounded on L_q for any $p_0 < q < p_1$.

Remark. The "Holmstedt formula" expresses the K-functional for the reiterated pair (X_0, X_1) in terms of the K-functional for the original pair (B_0, B_1). See [5] for details on this very useful formula.

Note that it is crucial that $\theta_0 \neq \theta_1$ in order to obtain that $(X_0, X_1)_{\theta,q}$ does not depend on q_0 or q_1. In case $\theta_0 = \theta_1$, the result is as follows (see [5, p. 51 and p. 112].

Theorem 5.11. *With the notation of Theorem 5.9, assume now that $0 < \theta_0 = \theta_1 < 1$. In that case, we have*

$$(X_0, X_1)_{\theta,q} = B_{\tau,q}$$

(with equivalent norms) where τ is as before, but q is now restricted to satisfy $\frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}$.

The extremal endspaces $(B_0, B_1)_{\theta,1}$ and $(B_0, B_1)_{\theta,\infty}$ (recall (5.3)) play a very important rôle in real interpolation. The next Lemma helps to recognize when a space is intermediate between them.

Lemma 5.12. *Let B be any Banach space. Consider an operator $T\colon\ B_0 \cap B_1 \to B$. Fix $0 < \theta < 1$. If, for any $x \in B_0 \cap B_1$, we have*

$$\|Tx\|_B \leq \|x\|_{B_0}^{1-\theta} \|x\|_{B_1}^{\theta},$$

then T extends to a bounded operator from $(B_0, B_1)_{\theta,1}$ to B with

$$\|T\colon\ (B_0, B_1)_{\theta,1} \to B\| \leq C,$$

where C is a constant depending only on θ.

The duality for the real method is given by the following.

Theorem 5.13. *Let (B_0, B_1) be a compatible couple of Banach spaces. Assume that $B_0 \cap B_1$ is dense in both B_0 and B_1, so that the pair (B_0^*, B_1^*) is naturally compatible. Then, for any $0 < \theta < 1$ and $1 \le q < \infty$, setting as usual $q' = q/(q-1)$, we have*

$$(B_0, B_1)_{\theta,q}^* = (B_0^*, B_1^*)_{\theta,q'},$$

with equivalent norms.

The next result gives the main general known connection between the two methods.

Theorem 5.14. *Let (B_0, B_1) be a compatible couple of complex Banach spaces. Then, for any $0 < \theta < 1$, the following bounded inclusions hold*

$$(B_0, B_1)_{\theta,1} \subset (B_0, B_1)_\theta \subset (B_0, B_1)_{\theta,\infty}.$$

5.2 Dual and self-dual interpolation pairs

Let B be a reflexive Banach space. Assume given a continuous injection

$$T\colon\ B \to B^*$$

that is self-dual, i.e. such that

$$\forall x, y \in B \qquad T(x)(y) = T(y)(x). \tag{5.15}$$

For any x in B, we set

$$\|x\|_0 = \|x\|_B \quad \text{and} \quad \|x\|_1 = \|Tx\|_{B^*},$$

and we denote by B_1 the completion of the normed space $(B, \|\cdot\|_1)$. We have a canonical inclusion $B_0 \subset B_1$ that allows us to view (B_0, B_1) as a compatible pair of Banach spaces.

Note that, since B is reflexive and $T = T^*$ by (5.15), T injective implies that T has dense range. We have an isometric isomorphism

$$\Phi\colon\ B_1 \to B^*$$

defined by first setting $\Phi(x) = Tx$ for $x \in B$, observing that this is isometric (with B equipped with $\|\cdot\|_1$) and then noting that B and $T(B)$ are dense respectively in B_1 and B^*.

Theorem 5.15. *In the above situation, let $0 < \theta < 1$ and $1 \le q < \infty$.*

(i) *In the complex case, we have isometrically*

$$(B_0, B_1)_\theta^* \simeq (B_0, B_1)_{1-\theta}.$$

In particular, $(B_0, B_1)_{1/2}$ is isometric to its dual.

(ii) *In the real case, we have isomorphically*

$$(B_0, B_1)^*_{\theta,q} \simeq (B_0, B_1)_{1-\theta,q'}.$$

In particular $B_{1/2,2}$ is isomorphic to its dual.

Proof. The key is simply to observe that the pair (B_0^*, B_1^*) can be identified with (B_1, B_0). Indeed, consider $\varphi \in (B_0 \cap B_1)^* = B^*$. Obviously

$$\|\varphi\|_{B_0^*} = \|\varphi\|_{B^*}. \tag{5.16}$$

Also $\varphi \in B_1^*$ with norm c iff we have $|\varphi(y)| \leq c\|y\|_{B_1} = c\|T(y)\|_{B^*}$ for any y in B, or equivalently by Hahn–Banach, iff there is b in B with $\|b\| \leq c$ such that $\varphi(y) = T(y)(b)$ for any y in B. In other words, we have $\varphi = T^*b$ with $\|b\| \leq c$. Since we assume (5.15) we have $T = T^*$ and hence we find $B_1^* = T(B)$ with

$$\|\varphi\|_{B_1^*} = \|T^{-1}\varphi\|_B. \tag{5.17}$$

From (5.16) (resp. (5.17)) we see that the mapping $T\colon\ B \to B^*$ extends (simultaneously) to an isometric isomorphism from B_1 to B_0^* (resp. $B_0 \to B_1^*$). From this it is clear that T defines an isometric isomorphism from $(B_0, B_1)_\theta$ to $(B_1^*, B_0^*)_\theta$ and by Theorem **??** we have $(B_1^*, B_0^*)_\theta = (B_1, B_0)^*_\theta = (B_0, B_1)^*_{1-\theta}$. This completes the proof of (i). In the real case, the proof is the same but now we use Theorem 5.13. □

Remark 5.16. A slightly different but equivalent viewpoint consists in using the map $T\colon\ B \to B^*$ to define the compatibility of the couple (B, B^*). Let (β_0, β_1) be the resulting interpolation pair. It is easy to check that T extends (by density) to an isometric isomorphism from $(B_0, B_1)_\theta$ to $(\beta_0, \beta_1)_\theta$ and also from $(B_0, B_1)_{\theta,q}$ to $(\beta_0, \beta_1)_{\theta,q}$ for all $0 < \theta < 1$ and $1 \leq q \leq \infty$.

Remark 5.17. Let $v\colon\ B^* \to \alpha_1$ be an isometric (resp. isomorphic) isomorphism from B^* onto another Banach space α_1. Note that if we replace T by vT then the pair (B_0, B_1) is unchanged (resp. except for an equivalent norm on B_1). Therefore, the resulting complex (real) interpolation spaces are identical (resp. isomorphic). Moreover, since the symmetry of T was not used there, the preceding remark remains valid: If we set $\alpha_0 = B$ and use $vT\colon\ \alpha_0 \to \alpha_1$ to define compatibility, then vT extends to an isometric isomorphism from $(B_0, B_1)_\theta$ to $(\alpha_0, \alpha_1)_\theta$ and also from $(B_0, B_1)_{\theta,q}$ to $(\alpha_0, \alpha_1)_{\theta,q}$ for all $0 < \theta < 1$ and $1 \leq q \leq \infty$.

The Hilbert space self-duality is the classical illustration of the preceding principle:

Proposition 5.18. *Let B and T be as above.*

(i) *In the real case, assume $T(x)(x) \geq 0$ for all x in B. Then $B_{1/2,2}$ is isomorphic to a Hilbert space and, when restricted to B, its norm is equivalent to $x \mapsto T(x)(x)^{1/2}$.*

(ii) *In the complex case, assume there is an isometric antilinear involution* $J\colon\ B^* \to B^*$ *such that* $T(x)(Jx) \geq 0$ *for all* x *in* B. *Then* $B_{1/2}$ *is isometric to a Hilbert space and*

$$\forall x \in B \qquad \|x\|_{B_{1/2}} = (T(x)(Jx))^{1/2}.$$

Proof. Recall that since T is injective and symmetric its range is dense. We will prove the complex case. We have a bilinear map $\Phi\colon\ (x,y) \to T(x)(y)$ that is of norm ≤ 1 both on $B_0 \times B_1$ and on $B_1 \times B_0$. By the fundamental interpolation property (see Theorem **??**), we have $\|\Phi\colon\ B_\theta \times B_{1-\theta} \to \mathbb{C}\| \leq 1$ and in particular

$$\forall x, y \in B \qquad |T(x)(y)| \leq \|x\|_{B_{1/2}} \|y\|_{B_{1/2}}$$

and if $y = Jx$

$$|T(x)(Jx)| \leq \|x\|^2_{B_{1/2}}.$$

Thus we find

$$\|x\|_H \leq \|x\|_{B_{1/2}} \tag{5.18}$$

where, by definition, we set $\|x\|_H = T(x)(Jx)$. Let H be the completion of $(B, \|\cdot\|_H)$. By the duality (5.18) implies

$$\|x\|_{B^*_{1/2}} \leq \|x\|_H \tag{5.19}$$

and since, by (i) in Theorem 5.15, $B^*_{1/2} = B_{1/2}$ we conclude from (5.18) and (5.19) that equality holds in (5.18). This shows that $B_{1/2} = H$. The proof of the real case is entirely similar. □

For example, the preceding statement applies to the pair

$$(\ell_1^n, \ell_\infty^n)$$

with $T\colon\ \ell_1^n \to \ell_\infty^n$ the identity map. We recover the identity $(\ell_1^n, \ell_\infty^n)_{1/2} = \ell_2^n$ and (uniformly over n)

$$(\ell_1^n, \ell_\infty^n)_{1/2,2} \simeq \ell_2^n.$$

Note however that Theorem 5.15 is quite interesting also when there is *no Hilbert space* in the picture (we will use the next example in Chapter 7):

Example 5.19. Let v_1^n denote $\mathbb{K}^n$ equipped with the norm

$$\|x\|_{v_1^n} = |x_1| + |x_2 - x_1| + \cdots + |x_n - x_{n-1}|.$$

We consider the interpolation spaces $(v_1^n, \ell_\infty^n)_{\theta,q}$ and $(v_1^n, \ell_\infty^n)_\theta$ $(0 < \theta < 1, 1 \leq q \leq \infty)$. Let $p = (1-\theta)^{-1}$. We denote

$$\mathcal{W}^n_{p,q} = (v_1^n, \ell_\infty^n)_{p,q} \quad \text{and} \quad \mathcal{W}^n_p = \mathcal{W}_{p,p}.$$

Note that $\|x\|_{v_1^n} = \|Tx\|_{\ell_1^n}$ where

$$Tx = (x_n - x_{n-1}, x_{n-1} - x_{n-2}, \ldots, x_2 - x_1, x_1). \tag{5.20}$$

Note that T satisfies (5.15) with respect to the canonical duality on $\mathbb{K}^n$ (equivalently the matrix of T is symmetric). Note that by (5.1) we have isometrically

$$\mathcal{W}^n_{p,q} = (\ell^n_\infty, v^n_1)_{1-\theta,q}.$$

Therefore (exchanging the rôles of v_1^n and ℓ^n_∞ for convenience) Theorem 5.15 yields:

Corollary 5.20. *In the complex case we have isometrically* $(v_1^n, \ell^n_\infty)^*_\theta = (v_1^n, \ell^n_\infty)_{1-\theta}$ *and in particular* $(v_1^n, \ell^n_\infty)_{1/2}$ *is isometric to its dual, via the mapping* $T\colon\ (v_1^n, \ell^n_\infty)_{1-\theta} \to (v_1^n, \ell^n_\infty)^*_\theta$ *defined in* (5.20).

Let $(e_1, \ldots, e_n)$ denote the canonical basis in $\mathbb{K}^n$ and let $(e_1^*, \ldots, e_n^*)$ be the biorthogonal functionals in $(\mathbb{K}^n)^*$.

Corollary 5.21. *For all* $1 < p < \infty$ *and* $1 \le q \le \infty$ *there is a constant* C *(independent of* n*) such that*

$$\|T\colon\ \mathcal{W}^n_{p',q'} \to (\mathcal{W}^n_{p,q})^*\| \le C \quad \text{and} \quad \|T^{-1}\colon\ (\mathcal{W}^*_{p,q})^n \to \mathcal{W}^n_{p',q'}\| \le C.$$

Moreover, if we let $\sigma_j = \sum_1^j e_k$ $(1 \le j \le n)$ *then for all* x *in* $\mathbb{K}^n$ *we have*

$$\frac{1}{2C}\left\|\sum\nolimits_1^n x_j e_j^*\right\|_{(\mathcal{W}^n_{p,q})^*} \le \left\|\sum\nolimits_1^n x_j\sigma_j\right\|_{\mathcal{W}^n_{p',q'}} \le 2C\left\|\sum\nolimits_1^n x_j e_j^*\right\|_{(\mathcal{W}^n_{p,q})^*} \tag{5.21}$$

Proof. The first part is but a particular case of Theorem 5.15. Let $x = Ty$, $y \in \mathbb{K}^n$. We have

$$\frac{1}{C}\left\|\sum y_j e_j\right\|_{\mathcal{W}^n_{p',q'}} \le \left\|\sum x_j e_j^*\right\|_{(\mathcal{W}^n_{p,q})^*} \le C\left\|\sum y_j e_j\right\|_{\mathcal{W}^n_{p',q'}}. \tag{5.22}$$

Let $V\colon\ \mathbb{K}^n \to \mathbb{K}^n$ be defined by $V(z_1, \ldots, z_n) = (z_n, \ldots, z_1)$. Note that $\|Vz\|_{v_1^n} \le 2\|z\|_{v_1^n}$, V is an isometry on ℓ^n_∞ and $V = V^{-1}$, therefore we have

$$\forall z \in \mathbb{K}^n \qquad 2^{-1}\|z\|_{\mathcal{W}^n_{p',q'}} \le \|Vz\|_{\mathcal{W}^n_{p',q'}} \le 2\|z\|_{\mathcal{W}^n_{p',q'}}.$$

But then we have

$$y = T^{-1}x = (x_n, x_n + x_{n-1}, \ldots, x_n + \cdots + x_1) = V\sum\nolimits_1^n x_j\sigma_j.$$

Therefore

$$2^{-1}\left\|\sum\nolimits_1^n x_j\sigma_j\right\|_{\mathcal{W}^n_{p',q'}} \le \|y\|_{\mathcal{W}^n_{p',q'}} \le 2\left\|\sum\nolimits_1^n x_j\sigma_j\right\|_{\mathcal{W}^n_{p',q'}}$$

and (5.21) follows from (5.22). $\square$

5.3 Notes and Remarks

The material in this chapter is classical. The basic reference is [5] and also [35]. For real interpolation, see [10] and also [6].

The complex interpolation method was introduced independently by A. Calderón and J.L. Lions around 1960. While Lions wrote nothing but the Comptes Rendus note [193], Calderón published a very detailed, very thorough account of all aspects of his theory. His memoir [111] remains must reading for anyone interested in the subject.

In turn, J.L. Lions concentrated his efforts on the real interpolation in collaboration with J. Peetre, see notably [194]. Later on, Peetre introduced the K- and J-method that replaced advantageously the Lions–Peetre methods and have been tremendously successful in analysis and approximation theory.

The self-duality results in §5.2 go back to the early days of interpolation, both real and complex. However, the original versions required extra assumptions such as e.g. reflexivity, that were lifted later on. See [256] and the references there for the state of the art in that direction. Corollaries 5.20 and 5.21 go back to some 1974 discussions with Bernard Maurey.

Chapter 6

The strong p-variation of scalar valued martingales

This chapter is based on [236].

Let $0 < p < \infty$ and let $x = (x_n)$ be a sequence in a Banach space B. The strong p-variation of $x = (x_n)$, denoted by $V_p(x)$, is defined as follows

$$V_p(x) = \sup \left(\|x_0\|^p + \sum_{j \geq 1} \|x_{n(j)} - x_{n(j-1)}\|^p \right)^{1/p}$$

where the supremum runs over all increasing sequences of integers $0 = n(0) < n(1) < n(2) < \dots$. We denote by $v_p(B)$ the space of all sequences $x = (x_n)$ such that $V_p(x) < \infty$. When $B = \mathbb{R}$, we set $v_p = v_p(\mathbb{R})$.
Note that for all $0 < p < q < \infty$ we have

$$V_q(x) \leq V_p(x). \tag{6.1}$$

Clearly, when $p \geq 1$, the spaces $v_p(B)$ and v_p are Banach spaces. The extreme cases $p = \infty$ and $p = 1$ are especially simple. Indeed, the analogue of $V_p(x)$ for $p = \infty$ is equivalent to $\sup_{n \geq 0} \|x_n\|$, so it is natural to set $v_\infty(B) = \ell_\infty(B)$. As for $p = 1$, the triangle inequality shows that

$$V_1(x) = \|x_0\| + \|x_1 - x_0\| + \|x_2 - x_1\| + \cdots$$

so that $v_1(B)$ is just the space of sequences in B with bounded variation.

We will make crucial use of real interpolation. Consider a measure space $(\Omega, \mathcal{A}, \mu)$ and a Banach space B. For simplicity of notation we set $L_p(B) = L_p(\Omega, \mu, \mathcal{A}; B)$. Let (B_0, B_1) be an interpolation pair of Banach spaces. Consider the interpolation pair $(L_{p_0}(B_0), L_{p_1}(B_1))$ where $1 \leq p_0 \neq p_1 \leq \infty$. Let $0 < \theta < 1$ and $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$. By (5.12) we have

$$L_p((B_0, B_1)_{\theta,p}) = (L_{p_0}(B_0), L_{p_1}(B_1))_{\theta,p}, \tag{6.2}$$

with equivalent norms.

In this chapter the couple $(v_1(B), \ell_\infty(B))$ plays the central rôle. Let $1 < p < \infty$ and $\theta = 1 - 1/p$. We denote

$$\mathcal{W}_{p,q}(B) = (v_1(B), \ell_\infty(B))_{\theta,q} \quad (0 < \theta < 1,\ 1 \le q \le \infty).$$

We also set

$$\mathcal{W}_p(B) = \mathcal{W}_{p,p}(B).$$

We now apply (6.2) to the couple $(v_1(B), \ell_\infty(B))$. This gives us assuming $p = (1-\theta)^{-1}$ (i.e. $\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{\infty}$):

$$L_p(\mathcal{W}_p(B)) = (L_1(v_1(B)), L_\infty(\ell_\infty(B)))_{\theta,p} \tag{6.3}$$

with equivalent norms. The connection between $\mathcal{W}_p(B)$ and the strong p-variation lies in the following.

Lemma 6.1. *If $1 < p < \infty$ and $1 - \theta = \frac{1}{p}$, then $\mathcal{W}_p(B) \subset v_p(B)$ and this inclusion has norm bounded by a constant $K(p)$ depending only on p.*

Proof. This is easy to prove. Indeed for any fixed sequence $0 = n(0) < n(1) < \dots$ we introduce the operator $T\colon\ v_1(B) \to \ell_1(B)$ defined by

$$T(x) = (x_0, x_{n(1)} - x_0, \dots, x_{n(k)} - x_{n(k-1)}, \dots).$$

This has clearly norm ≤ 1. On the other hand, considered as operator from $\ell_\infty(B)$ into $\ell_\infty(B)$, T has norm ≤ 2. Therefore it follows from the interpolation theorem (cf. Theorem 5.2 above), that T has norm ≤ 2 as an operator from $\mathcal{W}_p(B)$ into $(\ell_1(B), \ell_\infty(B))_{\theta,p}$. By Theorem 5.7, this space can be identified with $\ell_p(B)$ with an equivalent norm. This yields (for some constant $K(p)$)

$$\left(\|x_0\|^p + \sum \|x_{n(k)} - x_{n(k-1)}\|^p\right)^{1/p} \le K(p)\|x\|_{\mathcal{W}_p(B)},$$

and the announced result clearly follows from this. □

In this chapter we study the strong p-variation of scalar martingales. We will return to the B-valued case in a later chapter. Our main result is the following

Theorem 6.2. *Assume $1 \le p < 2$.*

(i) *There is a constant C_p such that every martingale $M = (M_n)_{n\ge0}$ in L_p satisfies (with the convention $M_{-1} \equiv 0$)*

$$\mathbb{E}V_p(M)^p \le (C_p)^p \sum\nolimits_{n\ge0} \mathbb{E}|M_n - M_{n-1}|^p.$$

(ii) *More generally, if $1 \le r \le p$, there is a constant C_{pr} such that every martingale $M = (M_n)_{n\ge0}$ in L_r satisfies*

$$\|V_p(M)\|_r \le C_{pr} \left\| \left(\sum_{n\ge0} |M_n - M_{n-1}|^p \right)^{1/p} \right\|_r .$$

Throughout the sequel, we will set by convention $M_{-1} = 0$ whenever $M = (M_n)_{n\geq 0}$ is a martingale. All the r.v.'s are assumed to be defined on a given probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We will need the following key lemma.

Lemma 6.3. *For any martingale M in L_2, we have*

$$\|M\|_{(L_1(v_1), L_\infty(\ell_\infty))_{\frac{1}{2}\infty}} \leq 2 \left(\sum_{n\geq 0} \mathbb{E}|M_n - M_{n-1}|^2 \right)^{1/2}.$$

Note that by orthogonality we have

$$\sum_{n\geq 0} \mathbb{E}|M_n - M_{n-1}|^2 = \sup_{n\geq 0} \mathbb{E}|M_n|^2. \tag{6.4}$$

Proof of Lemma 6.3. Given a sequence of r.v.'s $X = (X_n)_{n\geq 0}$, we denote simply by $K_t(X)$ the K_t-norm of X with respect to the couple $(L_1(v_1),\ L_\infty(\ell_\infty))$. Explicitly, assuming that (X_n) converges a.s., we have

$$K_t(X) = \inf \left\{ \|X_0^0\|_1 + \sum_{n\geq 1} \|X_n^0 - X_{n-1}^0\|_1 + t \sup_n \|X_n^1\|_\infty \right\} \tag{6.5}$$

where the infimum runs over sequences of r.v.'s X^0 and X^1 such that $X_n = X_n^0 + X_n^1$ for all $n \geq 0$. Note that the assumed a.s. convergence allows us to invoke Remark 5.8 with $\Lambda_1 = \ell_\infty(L_\infty)$.

Let (M_n) be a martingale, relative to an increasing sequence of σ-algebras $(\mathcal{A}_n)_{n\geq 0}$, and let $0 \leq T_0 \leq T_1 \leq \ldots$ be a sequence of stopping times (relative to $(\mathcal{A}_n)_{n\geq 0}$) with values in $\mathbb{N} \cup \{\infty\}$. We assume that (M_n) is bounded in L_2, hence M_n converges a.s. (and in L_2) to a limit denoted by M_∞ which is in L_2. Moreover, we have $M_n = \mathbb{E}(M_\infty|\mathcal{A}_n)$ and $M_T = \mathbb{E}(M_\infty|\mathcal{A}_T)$ for any stopping time T with values in $\mathbb{N} \cup \{\infty\}$. Therefore, the sequence $(M_{T_k})_{k\geq 0}$ is a martingale, and (6.4) implies

$$\begin{aligned} \mathbb{E}\left(|M_{T_0}|^2 + \sum_{k\geq 1} |M_{T_k} - M_{T_{k-1}}|^2 \right) &\leq \sup \mathbb{E}|M_{T_k}|^2 \\ &\leq \mathbb{E}|M_\infty|^2 = \sum_{n\geq 0} \mathbb{E}|M_n - M_{n-1}|^2, \end{aligned} \tag{6.6}$$

To prove Lemma 6.3, we may assume for simplicity that $\|M_\infty\|_2 \leq 1$. Then we define by induction starting with $T_0 = \inf\{n \geq 0, |M_n| > t^{-1/2}\}$,

$$\begin{aligned} &T_1 = \inf\{n > T_0, |M_n - M_{T_0}| > t^{-1/2}\} \\ &\vdots \\ &T_k = \inf\{n > T_{k-1}, |M_n - M_{T_{k-1}}| > t^{-1/2}\} \end{aligned}$$

and so on. As usual, we make the convention $\inf \emptyset = +\infty$, i.e., we set $T_k = +\infty$ on the set where

$$\sup_{n>T_{k-1}} |M_n - M_{T_{k-1}}| \le t^{-1/2}.$$

Clearly $\{T_k\}$ is increasing sequence of stopping times so that (6.6) holds. We note that if $T_0(\omega) < \infty$ then $|M_{T_0(\omega)}(\omega)| \ge t^{-1/2}$ and

$$\text{if} \quad T_k(\omega) < \infty \quad \text{and} \quad k \ge 1 \quad \text{then} \quad (M_{T_k} - M_{T_{k-1}})(\omega) \ge t^{-1/2}. \tag{6.7}$$

Moreover, we have for all $k \ge 0$

$$\sup_{T_k \le n < T_{k+1}} |M_n - M_{T_k}| \le t^{-1/2} \text{ a.s.} \quad \text{and} \quad \sup_{n<T_0} |M_n| \le t^{-1/2} \text{ a.s.} \tag{6.8}$$

Hence, we can write $M_n = X_n^0 + X_n^1$, with X^0, X^1 defined as follows

$$X_n^0 = \sum\nolimits_{k\ge 0} 1_{\{T_k \le n < T_{k+1}\}} M_{T_k}$$
$$X_n^1 = \sum\nolimits_{k\ge 0} 1_{\{T_k \le n < T_{k+1}\}} (M_n - M_{T_k}) + 1_{\{n<T_0\}} M_n.$$

By (6.8), on one hand we have

$$\|\sup |X_n^1|\,\|_\infty \le t^{-1/2}. \tag{6.9}$$

On the other hand, let $\varDelta_0 = |M_{T_0}|$ and $\varDelta_k = |M_{T_k} - M_{T_{k-1}}|$ for $k \ge 1$.
We have

$$|X_0^0| + \sum\nolimits_{n\ge 1} |X_n^0 - X_{n-1}^0| = 1_{\{T_0<\infty\}} \varDelta_0 + \sum\nolimits_{k\ge 1} \varDelta_k 1_{\{T_k<\infty\}}. \tag{6.10}$$

This can be estimated as follows. We have by (6.7)

$$t^{-1/2} \left(1_{\{T_0<\infty\}} + \sum\nolimits_{k\ge 1} 1_{\{T_k<\infty\}}\right) \le \varDelta_0 1_{\{T_0<\infty\}} + \sum\nolimits_{k\ge 1} \varDelta_k 1_{\{T_k<\infty\}}. \tag{6.11}$$

Let $N = 1_{\{T_0<\infty\}} + \sum\limits_{k\ge 1} 1_{\{T_k<\infty\}}$. By Cauchy–Schwarz, (6.11) implies

$$N t^{-1/2} \le N^{1/2} \left(|\varDelta_0|^2 + \sum\nolimits_{k\ge 1} |\varDelta_k|^2\right)^{1/2}. \tag{6.12}$$

Clearly N is finite a.s. (since M_n converges a.s.), (6.12) implies

$$N^{1/2} \le t^{1/2} \left(|\varDelta_0|^2 + \sum |\varDelta_k|^2\right)^{1/2}$$

and hence by (6.6)

$$(\mathbb{E}N)^{1/2} \le t^{1/2} \|M_\infty\|_2 \le t^{1/2}. \tag{6.13}$$

Now going back to (6.10) we find again by Cauchy–Schwarz and (6.13)

$$\mathbb{E}\left(|X_0^0|+\sum\nolimits_{n\geq 1}|X_n^0-X_{n-1}^0|\right)\leq(\mathbb{E}N)^{1/2}\left\|\left(|\Delta_0|^2+\sum|\Delta_k|^2\right)^{1/2}\right\|_2$$
$$\leq t^{1/2}.$$

By (6.9) and (6.5), this yields $K_t(M)\leq 2t^{1/2}$ so that

$$\|M\|_{(L_1(v_1),L_\infty(\ell_\infty))_{1/2,\infty}}\leq 2.$$

By homogeneity, this completes the proof of Lemma 6.3. □

Proof of Theorem 6.2. Let $(\mathcal{A}_n)_{n\geq 0}$ be a fixed increasing sequence of σ-sub-algebras of $\mathcal{A}$. All martingales below will be with respect to $(\mathcal{A}_n)_{n\geq 0}$. For $1\leq p\leq\infty$, we will denote by D_p the subspace of $\ell_p(L_p)$ formed of all the sequences $\varphi=(\varphi_n)_{n\geq 0}$ such that φ_n is $\mathcal{A}_n$-measurable for all $n\geq 0$ and $\mathbb{E}(\varphi_n|\mathcal{A}_{n-1})=0$ for all $n\geq 1$. We first claim that if $1\leq p_0,\,p_1\leq\infty$ and if $\frac{1}{p}=\frac{1-\theta}{p_0}+\frac{\theta}{p_1}$ then

$$D_p=(D_{p_0},D_{p_1})_{\theta,p}. \tag{6.14}$$

This follows from an argument well known to interpolation theorists. Indeed, to check this, we first note that by (6.2) we have

$$(L_{p_0}(\ell_{p_0}),L_{p_1}(\ell_1))_{\theta,p}=L_p(\ell_p), \tag{6.15}$$

with equivalent norms.

We may clearly identify isometrically $L_p(\ell_p)$ and $\ell_p(L_p)$. There is a projection $P\colon\ L_p(\ell_p)\to D_p$ defined by

$$\forall X=(X_n)_{n\geq 0}\in L_p(\ell_p)\quad P(X)=(\varphi_n)_{n\geq 0}$$

with

$$\varphi_0=\mathbb{E}(X_0|\mathcal{A}_0)\quad\text{and}\quad\varphi_n=\mathbb{E}(X_n|\mathcal{A}_n)-\mathbb{E}(X_n|\mathcal{A}_{n-1}).$$

Clearly, P is a bounded projection onto D_p and

$$\|P(X)\|_{D_p}\leq 2\|X\|_{L_p(\ell_p)},$$

and consequently $\|P(X)\|_{(D_{p_0},D_{p_1})_{\theta,p}}\leq 2\|X\|_{(L_{p_0}(\ell_{p_0}),L_{p_1}(\ell_{p_1}))_{\theta,p}}$. By (6.15), this implies that for some constant $C=C(p_0,p_1,\theta)$

$$\|P(X)\|_{(D_{p_0},D_{p_1})_{\theta,p}}\leq C\|X\|_{L_1(\ell_p)}.$$

Applying this for X in D_p, we find

$$\|X\|_{(D_{p_0},D_{p_1})_{\theta,p}}\leq C\|X\|_{D_p}. \tag{6.16}$$

On the other hand, we have trivially

$$\|X\|_{L_{p_i}(\ell_{p_i})}\leq\|X\|_{D_{p_i}}\qquad\text{for } i=0,1$$

and hence by interpolation

$$\|X\|_{L_p(\ell_p)} \le C'\|X\|_{(D_{p_0},D_{p_1})_{\theta,p}} \tag{6.17}$$

for some constant $C' = C'(p_0, p_1, \theta)$.

Combining (6.16) and (6.17), we find the above claim (6.14).

We can now complete the proof of Theorem 6.2 (i).

Let us denote by T the operator which associates to any φ in D_1 the martingale $(M_n)_{n\ge 0}$ defined by $M_n = \sum\limits_{i\le n} \varphi_i$. Clearly $\|T(\varphi)\|_{L_1(v_1)} \le \|\varphi\|_{D_1}$. On the other hand, Lemma 6.3 implies that T is bounded from D_2 into $B_1 = (L_1(v_1), L_\infty(\ell_\infty))_{\frac{1}{2}\infty}$, with norm ≤ 2. Therefore if $1 < p < 2$ the interpolation Theorem 5.2 implies that T is bounded from $(D_1, D_2)_{\theta,p}$ into $(L_1(v_1), B_1)_{\theta,p}$. By the reiteration Theorem 5.9 we have $(L_1(v_1), B_1)_{\theta,p} = (L_1(v_1), L_\infty(\ell_\infty))_{\delta,p}$ with $\delta = \theta/2$. Now if θ is chosen so that $\frac{1}{p} = 1 - \delta$, we have by (6.3) and Lemma 6.1

$$(L_1(v_1), L_\infty(\ell_\infty))_{\delta,p} = L_p(\mathcal{W}_p) \subset L_p(v_p).$$

On the other hand, by (6.14) we have (since $\frac{1-\theta}{1} + \frac{\theta}{2} = \frac{1}{p}$) $(D_1, D_2)_{\theta,p} = D_p$. Recapitulating, we find a constant $C = C(p)$ depending only on $1 \le p < 2$ such that for all φ in D_p we have

$$\|T(\varphi)\|_{L_p(v_p)} \le C\|\varphi\|_{D_p}.$$

This establishes the first part of Theorem 6.2.

The second part follows from the standard arguments used to prove the Burkholder–Davis–Gundy inequalities. We use the general method described in Lemma 8.20. Let $g(\omega) = (g_n(\omega))_{n\ge 0}$ be a martingale in L_r. We set $v_\infty(\omega) = V_p(g(\omega))$ and for any $N \ge 1$ we denote by $v_N(\omega)$ the strong p-variation of $\{g_0(\omega), \dots, g_N(\omega)\}$ i.e. the strong p-variation of the *restriction* of our martingale to $[0, 1, \dots, N]$. Equivalently,$v_N(\omega)$ is the strong p-variation of $(g_{n\wedge N}(\omega))_{n\ge 0}$. We set $w_N = (\sum_0^N |dg_k|^p)^{1/p}$. Applying (i) to the martingale $(1_{\{T>0\}} g_{n\wedge T})_{n\ge 0}$, we find that (8.30) holds for any stopping time T. If we assume $|dg_{n+1}| \le \psi_n$ for all $n \ge 0$ with (ψ_n) adapted then Lemma 8.20 yields that for any $0 < r < p$ we have for some constant $C_1 = C_1(p, r)$

$$\|v_\infty\|_r \le C_1(\|w_\infty\|_r + \|\psi^*\|_r). \tag{6.18}$$

We now invoke Lemma 8.22 (B. Davis decomposition) with r replacing p. This gives us a decomposition $M_n = h_n + g_n$ with $h_0 = 0$, $|dg_n| \le 6M^*_{n-1}$ and $\|\sum |dh_n|\|_r \le 6r\|M^*\|_r$. By (6.18) we have

$$\|V_p(g)\|_r \le C'\left(\left\|\left(\sum |dg_n|^p\right)^{1/p}\right\|_r + 6\|M^*\|_r\right),$$

and since $V_p(M) \le V_p(g) + \sum |dh_n|$, $(\sum |dg_n|^p)^{1/p} \le (\sum |dM_n|^p)^{1/p} + \sum |dh_n|$ and also $\|\sum |dh_n|\|_r \le 6r\|M^*\|_r$, this implies that for some constant $C_2 = C_2(p, r)$ we have

$$\|V_p(M)\|_r \le C_2\left(\left\|\left(\sum |dM_n|^p\right)^{1/p}\right\|_r + \|M^*\|_r\right).$$

By the classical Burkholder–Gundy–Davis martingale inequalities (see (8.29) and (8.36)) and Doob's inequality (cf. (1.12)) we have for some constant $C_3 = C_3(r)$

$$\|M^*\|_r \leq C_3 \left\| \left(\sum |dM_n|^2 \right)^{1/2} \right\|_r$$

and hence, since $(\sum |dM_n|^2)^{1/2} \leq (\sum |dM_n|^p)^{1/p}$, we obtain

$$\|V_p(M)\|_r \leq C_2(1 + C_3) \left\| \left(\sum |dM_n|^p \right)^{1/p} \right\|_r . \qquad \square$$

□

The next result is an immediate consequence of Theorem 6.2.

Corollary 6.4. *Let $1 \leq p < 2$. Let $M = (M_t)_{t\geq 0}$ be a martingale in L_p. Assume that the paths of M are right continuous and admit left limits and that the continuous part of M is 0. Let*

$$V_p(M) = \sup_{0=t_0 \leq t_1 \leq \ldots} \left(|M_o|^p + \sum_{i \geq 1} |M_{t_i} - M_{t_{i-1}}|^p \right)^{1/p}$$

and

$$S_p(M) = \left(\sum_{t \in [0,\infty[} |M_t - M_{t^-}|^p \right)^{1/p} .$$

Then, for all $1 \leq r < \infty$, we have for any martingale M in L_r

$$\|V_p(M)\|_r \leq C_{pr} \|S_p(M)\|_r. \tag{6.19}$$

Remark. There are also inequalities similar to Theorem 6.2 (ii) or (6.19) with a "moderate" Orlicz function space instead of L_r, cf. [101, 108].

Our method gives (with almost no extra effort) a new proof of the following results of Lépingle [187].

Theorem 6.5. *Assume $2 < p < \infty$ and $1 \leq r \leq p$. Then there is a constant C_{pr} such that every martingale $M = (M_n)_{n\geq 0}$ in L_r satisfies*

$$\|V_p(M)\|_r \leq C_{pr} \| \sup_n |M_n| \, \|_r.$$

Moreover, there is a constant C'_p such that every martingale $M = (M_n)_{n\geq 0}$ in L_1 satisfies

$$\|V_p(M)\|_{1,\infty} \leq C'_p \sup_n \|M_n\|_1. \tag{6.20}$$

Proof. We first consider the particular case $r = p$. With the above notation, consider the operator $S\colon\ L_\infty \to L_\infty(\ell_\infty)$ defined for φ in L_∞ by

$$S(\varphi) = (\mathbb{E}(\varphi|\mathcal{A}_n))_{n\geq 0}.$$

Clearly $\|S\| \leq 1$. Let $B_0 = (L_1(v_1), L_\infty(\ell_\infty))_{1/2,\infty}$. By Lemma 6.3, S is bounded from L_2 into B_0. By the interpolation Theorem S must be bounded from $(L_2, L_\infty)_{\theta,p}$ into $(B_0, L_\infty(\ell_\infty))_{\theta,p}$ $(0 < \theta < 1, 1 \leq p \leq \infty)$. Now assume that $\frac{1}{p} = \frac{1-\theta}{2} + \frac{\theta}{\infty}$. Then, by (6.2), $(L_2, L_\infty)_{\theta,p} = L_p$. Moreover, by the reiteration Theorem 5.9

$$(B_0, L_\infty(\ell_\infty))_{\theta,p} = (L_1(v_1), L_\infty(\ell_\infty))_{\omega p}$$

for $\omega = \frac{1-\theta}{2} + \theta = \frac{1+\theta}{2}$. Note that $\frac{1-\omega}{1} + \frac{\omega}{\infty} = \frac{1}{p}$, hence by (6.2), the last equality implies $(B_0, L_\infty(\ell_\infty))_{\theta,p} = L_p(\mathcal{W}_p)$. Recapitulating, we find that S is bounded from L_p into $L_p(\mathcal{W}_p)$ with norm $\leq C_1(p)$ for some constant $C_1(p)$ depending only on p. Let $\varphi \in L_p$ and let $M_n = \mathbb{E}(\varphi|\mathcal{A}_n)$. Applying Lemma 6.1 again we conclude that

$$\|M\|_{L_p(v_p)} \leq K(p)\|M\|_{L_p(\mathcal{W}_p)} \leq K(p)C_1(p)\|\varphi\|_p.$$

This proves Theorem 6.5 in the case $r = p$.

We now turn to the case $1 \leq r \leq p$. We will argue as above for Theorem 6.2. Consider a martingale (M_n) in L_r. We apply the B. Davis decomposition (Lemma 8.22) in L_r, i.e. we have $M = g+h$ with $h_0 = 0$, $|dg_n| \leq 6M^*_{n-1}$ and $\|\sum |dh_n|\|_r \leq 6r\|M^*\|_r$. We define v_n and v_∞ as in the proof of Theorem 6.2, but we set

$$w_n = \sup\nolimits_{k\leq n} |g_n| \quad \text{and} \quad w_\infty = \sup\nolimits_n |g_n|.$$

Then applying the first part of the proof (i.e. the case $r = p$) to the martingale $(1_{\{T>0\}} g_{n\wedge T})$, we find that (8.30) holds for any stopping time T. Therefore, by Lemma 8.20 (note that in our case $\psi^* = 6M^*$) there is a constant $C_4 = C_4(p, r)$ such that

$$\|v_\infty\|_r \leq C_4(\|w_\infty\|_r + \|M^*\|_r),$$

or equivalently

$$\|V_p(g)\|_r \leq C_4(\|g^*\|_r + \|M^*\|_r).$$

But since $V_p(M) \leq V_p(g) + \sum |dh_n|$, $g^* \leq M^* + \sum |dh_n|$ and $\|\sum |dh_n|\|_r \leq 6r\|M^*\|_r$ we obtain finally

$$\|V_p(M)\|_r \leq (6r + 6rC_4 + C_4)\|M^*\|_r. \tag{6.21}$$

The weak type 1-1 inequality (6.20) is now an easy application of Theorem 8.13 (Gundy's decomposition). We leave the details as an exercise. $\square$

Remark. Of course, there is also a version of Theorem 6.5 in the case of a continuous parameter martingale $(M_t)_{t>0}$.

One can easily derive from Theorems 6.2 and 6.5 (by a classical stopping time argument) the following analogous "almost sure" statements.

Proposition 6.6. *Let $M = (M_n)$ be a martingale with $\mathbb{E} \sup_{n\geq 1} |M_n - M_{n-1}| < \infty$. If $1 \leq p < 2$, then $\{V_p(M) < \infty\} \overset{a.s.}{=} \{\sum_{n\geq 1} |M_n - M_{n-1}|^p < \infty\}$. Moreover if $2 < p < \infty$, then $\{V_p(M) < \infty\} \overset{a.s.}{=} \{\sup |M_n| < \infty\}$.*

Proof. Let $B_p = \{V_p(M) < \infty\}$, $A_p = \{\sum_0^\infty |dM_n|^p < \infty\}$, and $A_p(t) = \{\sum_0^\infty |dM_n|^p \leq t\}$. To prove the first assertion it suffices to show that $A_p(t) \subset B_p$ for all $0 < t < \infty$. Fix $0 < t < \infty$. Let $T = \inf\{n \mid \sum_0^n |dM_n|^p > t\}$. We may assume $M_0 = 0$ and hence $T > 0$ and $A_p(t) = \{T = \infty\}$. Let $f_n = M_{n\wedge T}$. Then $\sum |df_n|^p = \sum_{n\leq T} |dM_n|^p \leq t + \sup |dM_n|^p$. Therefore $(\sum |df_n|^p)^{1/p} \in L_1$. By (ii) (case $r = 1$) in Theorem 6.2 $V_p((f_n)) \in L_1$ and hence $V_p((f_n)) < \infty$ a.s., but on $\{T = \infty\}$ we have $(f_n) = (M_n)$ so $A_p(t) \subset B_p$. To prove the second assertion in case $2 < p < \infty$, we set $T = \inf\{n \mid |M_n| > t\}$ and again $f_n = M_{n\wedge T}$. Then $\{T = \infty\} = \{\sup |f_n| \leq t\}$, $\sup |f_n| \leq t + \sup |dM_n|$ and hence $\sup |f_n| \in L_1$. By Theorem 6.5 (case $r = 1$), $V_p((f_n)) \in L_1$ and hence $V_p((f_n)) < \infty$ a.s. and since $(f_n) = (M_n)$ on $\{T = \infty\}$ we conclude that $V_p(M) < \infty$ a.s. on the set $\{\sup |f_n| \leq t\}$. This proves the second assertion. □

6.1 Notes and Remarks

This chapter closely follows [236]. Theorem 6.5 was obtained first by Lépingle using the Skorokhod embedding of martingales into Brownian motion. Our proof is very different. Indeed we prove both Theorem 6.5 and Theorem 6.2 using the same idea, combining Lemma 6.3 and reiteration.

There is an extensive literature on the strong p-variation both in probability theory, function theory and harmonic analysis. We will only give below a few sample references.

Prior to [236], analogous questions had been considered mainly for sequences or processes with independent increments (cf. e.g. [92, 209, 210]). For a more recent approach to Corollary 6.4, see [226]. See [197] for a study of the strong p-variation of (strong) Markov processes.

See [251] for a more recent result on the strong p-variation of α-stable processes for $0 < \alpha \leq 2$ and $p > \alpha$.

See [18] for more information on the relations between p-variation, differentiability and empirical processes.

See [87, 174, 127] for inequalities analogous to those of this chapter in ergodic theory.

Note that our subsequent Chapter 7 contains a detailed study of the interpolation spaces $(v_1, \ell_\infty)_{\theta,q}$ that is quite useful to understand the spaces of functions with strong p-variation finite.

Chapter 7

Interpolation between strong p-variation spaces

In chapter 6 we already used interpolation to establish several martingale inequalities involving the space v_p of scalar sequences with finite "strong p-variation". When $p = 1$ the latter is just the space of sequences with bounded variation. In this chapter we will focus on the real interpolation space, defined, for $0 < \theta < 1$, $p = (1-\theta)^{-1}$ and $1 \le q \le \infty$, by

$$\mathcal{W}_{p,q} = (v_1, \ell_\infty)_{\theta,q}. \tag{7.1}$$

In words, we are interested in interpolating between the properties bounded variation and boundedness. Clearly

$$v_1 \subset \mathcal{W}_{p,q} \subset \ell_\infty.$$

Remark 7.1. Note that by Theorem 3.10 any intermediate Banach space between v_1 and ℓ_∞ is necessarily *non-reflexive*. Actually, the argument for (iv) $\Rightarrow$ (i) in Theorem 3.10 shows that the inclusion map $v_1 \to \ell_\infty$ is not weakly compact.

7.1 Strong p-variation: The spaces v_p and $\mathcal{W}_p$

If we replace v_1 by ℓ_1 in (7.1) we obtain the Lorentz space $\ell_{p,q}$ and in particular if $q = p$ we find the space ℓ_p. We will show that although (7.1) is non-reflexive it behaves in many ways like the spaces ℓ_{pq} and like ℓ_p when $q = p$.
When $q = p$ we write simply

$$\mathcal{W}_p = \mathcal{W}_{p,p}. \tag{7.2}$$

More generally, for any auxiliary Banach space B we define, again with $p = (1-\theta)^{-1}$

$$\mathcal{W}_{p,q}(B) = (v_1(B), \ell_\infty(B))_{\theta,q}, \tag{7.3}$$

$$\mathcal{W}_p(B) = \mathcal{W}_{p,p}(B) = (v_1(B), \ell_\infty(B))_{\theta,p}. \tag{7.4}$$

A simple application of interpolation yields:

Lemma 7.2. *For any increasing sequence* $0 = n(0) < n(1) < n(2) < \dots$ *of integers and any* x *in* $\mathcal{W}_{p,q}(B)$

$$\|(x_0, x_{n(1)} - x_0, \cdots, x_{n(k)} - x_{n(k-1)}, \cdots)\|_{\ell_{p,q}(B)} \le 2^{1-1/p}\|x\|_{\mathcal{W}_{p,q}(B)}. \tag{7.5}$$

Proof. Indeed, we just apply the fundamental interpolation property (Theorem 5.2) to the operator T defined by $T(x) = (x_0, x_{n(1)} - x_0, \cdots, x_{n(k)} - x_{n(k-1)}, \cdots)$. This is clearly bounded simultaneously from $v_1(B)$ to $\ell_1(B)$ (with norm ≤ 1) and from $\ell_\infty(B)$ to $\ell_\infty(B)$ (with norm ≤ 2), and hence from $\mathcal{W}_{p,q}(B) = (v_1(B), \ell_\infty(B))_{\theta,q}$ to $\ell_{p,q}(B) = (\ell_1(B), \ell_\infty(B))_{\theta,q}$ (with norm $\le 2^{1-1/p}$). □

By general interpolation theory (see (5.3) and (5.4)), for all $1 < p < r$ and arbitrary $1 \le q_0, q_1 \le \infty$ we have bounded inclusions

$$\mathcal{W}_{p,q_0}(B) \subset \mathcal{W}_{r,q_1}(B). \tag{7.6}$$

This also holds in case $p = r$, but then only if $q_1 \ge q_0$.

We denote as usual by c_0 (resp. $c_0(B)$) the subspace of ℓ_∞ (resp. $\ell_\infty(B)$) formed of all sequences that tend to zero. Similarly we will denote by v_1^0 (resp. $v_1^0(B)$) the subspace of v_1 (resp. $v_1(B)$) formed of all sequences that tend to zero. Recall that $\mathbb{K}$ denotes the sclars i.e. $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. Note that, by subtracting its limit from a sequence in v_1 or in $v_1(B)$ we find

$$v_1 \simeq \mathbb{K} \oplus v_1^0 \quad \text{and} \quad v_1(B) \simeq B \oplus v_1^0(B). \tag{7.7}$$

The pair (v_1, ℓ_∞) has a self-dual character that will be crucial in the sequel. Let us describe this duality. For $x \in v_1$, $y \in \ell_\infty$ we set

$$\langle x, y\rangle = x_0 y_0 + \sum\nolimits_1^\infty (x_n - x_{n-1}) y_n. \tag{7.8}$$

Note that $|\langle x, y\rangle| \le ||x||_{v_1}\|y\|_\infty$. Moreover, with this duality we have

$$(v_1^0)^* \simeq \ell_\infty \quad \text{and} \quad (c_0)^* = v_1.$$

More generally, we have

$$v_1^0(B)^* \simeq \ell_\infty(B^*) \quad \text{and} \quad c_0(B)^* = v_1(B^*) \tag{7.9}$$

with respect to the duality defined either for $x \in v_1(B^*)$ and $y \in c_0(B)$, or for $x \in \ell_\infty(B^*)$ and $y \in v_1^0(B)$, by

$$\langle x, y\rangle = \langle x_0, y_0\rangle + \sum\nolimits_1^\infty \langle x_n - x_{n-1}, y_n\rangle \tag{7.10}$$

$$= \lim_{n\to\infty} (\langle x_0, y_0 - y_1\rangle + \cdots + \langle x_{n-1}, y_{n-1} - y_n\rangle + \langle x_n, y_n\rangle) \tag{7.11}$$

$$= \sum\nolimits_1^\infty \langle x_{n-1}, y_{n-1} - y_n\rangle. \tag{7.12}$$

More precisely, we have

$$2^{-1}\|x\|_{\ell_\infty(B^*)} \le \|x\|_{(v_1^0(B))^*} \le \|x\|_{\ell_\infty(B^*)} \text{ and } \|x\|_{(c_0(B))^*} = \|x\|_{v_1(B^*)}. \tag{7.13}$$

Remark. For any sequence $x = (x_n)$, let $\hat{x}$ be the shifted sequence defined by $\hat{x}_0 = 0$ and $\hat{x}_n = x_{n-1}$ for all $n \geq 1$. If $x = (x_n)$ and $y = (y_n)$ are both finitely supported, then, taking $B = \mathbb{K}$ for simplicity, Abel summation (or integration by parts) shows $\langle x, y \rangle = -\langle y, \hat{x} \rangle$. A similar identity holds if y (resp. x) is a B-(resp. B^*)-valued sequence (but this requires exchanging the roles of B and B^*).

We will now introduce preduals $u^0_{p'}$ and u^0_p respectively for the spaces v_p or $v_{p'}$ when (we keep this notation throughout)

$$1 < p, p' < \infty \quad \text{and} \quad \frac{1}{p} + \frac{1}{p'} = 1.$$

Recall that for any x in $B^{\mathbb{N}}$, its strong p-variation $V_p(x)$ is defined by

$$V_p(x) = \sup \left\{ \left(\|x_0\|^p + \sum \|x_{n(k)} - x_{n(k-1)}\|^p \right)^{1/p} \right\}$$

where the supremum runs over all $0 = n(0) < n(1) < \cdots$. We set

$$v_p(B) = \{x \in B^{\mathbb{N}} \mid V_p(x) < \infty\} \quad \text{and} \quad \|x\|_{v_p(B)} = V_p(x).$$

When $B = \mathbb{K}$ we denote simply $v_p = v_p(\mathbb{K})$.
Note that, by the Cauchy criterion, $V_p(x) < \infty$ implies that x_n converges to a limit $x_\infty \in B$ when $n \to \infty$.
Let

$$v^0_p(B) = v_p(B) \cap c_0(B) \quad \text{and} \quad v^0_p = v_p \cap c_0.$$

Note that

$$\forall x \in v_p(B) \qquad (x_n - x_\infty)_{n \geq 0} \in v^0_p(B). \tag{7.14}$$

Let $B^{(\mathbb{N})}$ denote the space of finitely supported functions $b = (b(n))_{n \in \mathbb{N}}$ with $b(n) \in B$ for all n. It is easy to see that $B^{(\mathbb{N})}$ is dense in $v^0_p(B)$ for any $1 \leq p < \infty$ (see the proof of Lemma 7.6 below).

For any $b = (b(n)) \in B^{(\mathbb{N})}$ there is a finite partition of $\mathbb{N}$ into disjoint intervals $I_0, I_1, \ldots, I_N$ with $I_0 = [0, n(0)]$, $I_1 =]n(0), n(1)], \ldots, I_N =]n(N-1), n(N)]$ and there are $\xi_0, \ldots, \xi_N \in B$ such that

$$\forall n \in \mathbb{N} \qquad b(n) = \sum\nolimits_0^N \xi_k 1_{I_k}(n). \tag{7.15}$$

We require that $\xi_N \neq 0$ and $\xi_k \neq \xi_{k+1}$ for all $0 \leq k < N$ and we set

$$[b]_{p'} = \left(\sum\nolimits_0^N \|\xi_k\|^{p'} \right)^{1/p'}.$$

Note that the preceding requirement minimizes $\sum \|\xi_k\|^{p'}$.
For any $x \in (B^*)^{\mathbb{N}}$ we have

$$\langle x, b \rangle = \langle x_{n(0)}, \xi_0 \rangle + \sum\nolimits_1^N \langle x_{n(k)} - x_{n(k-1)}, \xi_k \rangle$$

and hence we have

$$\sup\{|\langle x, b\rangle| \mid b \in B^{(\mathbb{N})},\ [b]_{p'} \leq 1\} = \widetilde{V}_p(x) \tag{7.16}$$

where

$$\widetilde{V}_p(x) = \sup_{0 \leq n(0) < n(1) < \cdots} \{(\|x_{n(0)}\|^p + \|x_{n(1)} - x_{n(0)}\|^p + \cdots)^{1/p}\}.$$

Note that

$$V_p(x) \leq \widetilde{V}_p(x) \leq 2^{1/p'} V_p(x). \tag{7.17}$$

Note also that if $I_0 = [0, n(0)]$ we have

$$\langle x, \xi_0 1_{I_0}\rangle = \langle x_{n(0)}, \xi_0\rangle. \tag{7.18}$$

We then set for any $b \in B^{(\mathbb{N})}$

$$\|b\|_{u^0_{p'}(B)} = \inf \left\{ \sum\nolimits_1^m [b_j]_{p'} \right\}$$

where the infimum runs over all decompositions $b = \sum_1^m b_j$ $(b_j \in B^{(\mathbb{N})})$. In other words, $\|\cdot\|_{u^0_{p'}(B)}$ is the gauge of the convex hull of $\{b \mid [b]_{p'} \leq 1\}$. Then we define the Banach space $u^0_{p'}(B)$ as the completion of $B^{(\mathbb{N})}$ equipped with this norm. Note that $\sup \|b(n)\| \leq [b]_{p'}$ and hence $\lim_{n\to\infty} b(n) = 0$ for any b in $u^0_{p'}(B)$ so that we have a bounded inclusion

$$u^0_{p'}(B) \subset c_0(B).$$

In fact, $b \in u^0_{p'}(B)$ iff b can be written as $b = \sum_1^\infty b_j$ with $b_j \in B^{(\mathbb{N})}$ such that $\sum_1^\infty [b_j]_{p'} < \infty$. Moreover (the inf being over all such decompositions):

$$\|b\|_{u^0_{p'}(B)} = \inf \sum\nolimits_1^\infty [b_j]_{p'}.$$

When $B = \mathbb{K}$, we denote simply

$$u^0_{p'} = u^0_{p'}(\mathbb{K}).$$

From (7.16) it is immediate that, with respect to the duality (7.10) we have

$$v_p(B^*) = u^0_{p'}(B)^* \tag{7.19}$$

with equivalent norms. More explicitly, any $x \in v_p(B^*)$ defines a linear form f_x on $u^0_{p'}(B)^*$ by setting $f_x(b) = \langle x, b\rangle$ for any $b \in B^{(\mathbb{N})}$. By (7.16) and (7.17), the latter form admits a unique bounded extension to an element of $u^0_{p'}(B)^*$ satisfying

$$\|x\|_{v_p(B^*)} = V_p(x) \leq \|f_x\|_{u^0_{p'}(B)^*} \leq 2^{1/p'} \|x\|_{v_p(B^*)}.$$

Conversely, to any linear form $f \in u^0_{p'}(B)^*$, we associate the sequence $(x_n) \in B^{*\mathbb{N}}$ defined by $f(\xi 1_{[0,n]}) = x_n(\xi)$ (for all $\xi \in B,\ n \geq 0$). Then, recalling (7.18), we have $f(b) = f_x(b)$ for any $b \in B^{(\mathbb{N})}$ and (7.16) again shows that $x \in v_p(B^*)$ and of course $f = f_x$. Thus we conclude that the correspondence $x \mapsto f_x$ is a surjective isomorphism from $v_p(B^*)$ onto $u^0_{p'}(B)^*$. In this way, we avoid discussing the possible types of convergence of the series (7.10).

Remark 7.3. By an abuse of notation we will continue to denote by $\langle x, y\rangle$ the duality just established for $x \in v_p(B^*)$ and $y \in u^0_{p'}(B)$. (We adopt that notation also for $x \in u^0_{p'}(B^*)$ and $y \in v^0_p(B)$.) Note however, that this is really defined only when y is finitely supported and extended by density and continuity to the whole of $u^0_{p'}(B)$.

In particular, $v_p = (u^0_{p'})^*$. Thus there is a constant $C = C(p)$ such that for all x in $v_p(B^*)$ we have

$$\frac{1}{C}\|x\|_{v_p(B^*)} \leq \sup\{|\langle x, y\rangle| \mid y \in u^0_{p'}(B),\ \|y\|_{u^0_{p'}(B)} \leq 1\} \leq C\|x\|_{v_p(B^*)}. \tag{7.20}$$

Moreover, since $u^0_{p'}(B)^*$ norms $u^0_{p'}(B)$, there is a constant $C' = C'(p)$ such that for any y in $u^0_{p'}(B)$ we have

$$\frac{1}{C'}\|y\|_{u^0_{p'}(B)} \leq \sup\{|\langle x, y\rangle| \mid x \in v^0_p(B^*),\ V_p(x) \leq 1\} \leq C'\|y\|_{u^0_{p'}(B)}. \tag{7.21}$$

Indeed, the last equivalence is clear if we replace $v^0_p(B^*)$ by $v_p(B^*) = u^0_{p'}(B)^*$, but if y is supported say in $[0, \ldots, N]$ then for any x in $v_p(B^*)$

$$\langle x, y\rangle = \langle P_N(x), y\rangle$$

where $P_N(x) = (x_0, x_1, \ldots, x_N, 0 \ldots)$ and we have obviously

$$V_p(P_N(x)) \leq 2V_p(x). \tag{7.22}$$

From this (7.21) follows easily for all y in $u^0_{p'}(B)$.

Let us denote $g_y(x) = \langle x, y\rangle$. Then (7.21) can be rewritten

$$\frac{1}{C'}\|y\|_{u^0_{p'}(B)} \leq \|g_y\|_{v^0_p(B^*)^*} \leq C'\|y\|_{u^0_{p'}(B)}. \tag{7.23}$$

Remark 7.4. Using the $\ell_{p,q}$ norm in place of the ℓ_p norm in the definition of $[\ .\ ]_{p'}$, we can define analogously the space $u^0_{p,q}(B)$ and if $1 \leq p, q < \infty$ the same argument leads to $u^0_{p,q}(B)^* = v_{p',q'}(B^*)$.

By Theorem 3.10 we already know that v^0_p (and a fortiori v_p) is non-reflexive, but it is much less obvious that, if $1 < p < \infty$, it is quasi-reflexive, i.e. of finite codimension in its bidual. This phenomenon was discovered by James. For that reason, the space v^0_2 is usually denoted by J and called the James space. In fact we have

Theorem 7.5. *Let $1 < p < \infty$ with $\frac{1}{p} + \frac{1}{p'} = 1$. By Remark 7.3, the duality (7.8) (or (7.10) in the B valued case) is well defined for $x \in v_p, y \in u^0_{p'}$, and also for $x \in u^0_{p'}, y \in v^0_p$. With respect to that duality, we have*

$$(v^0_p)^* = u^0_{p'} \quad \textit{and} \quad (u^0_{p'})^* = v_p$$

with equivalent norms. More explicitly, the mapping $y \mapsto g_y$ (resp. $x \mapsto f_x$) extends to an isomorphism from $u^0_{p'}$ to $(v^0_p)^*$ (resp. from v_p to $(u^0_{p'})^*$). In particular if X is either $u^0_{p'}$, v^0_p or v_p we have $\dim(X^{**}/X) = 1$. More generally, if $\dim(B) = n$ then if $X = v_p(B)$ or if $X = u^0_p(B)$ we have $\dim(X^{**}/X) = n$.

Lemma 7.6. *The canonical basis (e_n) is a basis of v^0_p satisfying an upper p-estimate in the sense of Remark 3.18. In particular, it is a shrinking basis of v^0_p $(1 < p < \infty)$.*

Proof. By (7.22) we already know that (e_n) is a basic sequence in v^0_p. Let $P_N x = (x_0, x_1, \ldots, x_N, 0, \ldots)$. Assume $V_p(x) < \infty$. We will show that $V_p(x - P_N x) \to 0$ when $N \to \infty$ for any $x \in v^0_p$. Choose $0 = n(0) < n(1) < \cdots < n(K)$ such that

$$V_p(x)^p - \varepsilon < |x_0|^p + |x_{n(1)} - x_0|^p + \cdots + |x_{n(K)} - x_{n(K-1)}|^p.$$

We have then for any $n(K) < n(K+1) < \cdots$

$$\sum\nolimits_{j>K} |x_{n(j)} - x_{n(j-1)}|^p < \varepsilon$$

and hence if $N = n(K)$

$$V_p(x - P_N x)^p \le \sup_{j \ge N} |x_j|^p + \varepsilon,$$

and since we may assume that $N = n(K)$ is as large as we wish we conclude that $V_p(x - P_N x) \to 0$ when $N \to \infty$ for any $x \in v^0_p$.

Note that by (7.17) we have

$$\sup |x_j| \le 2^{1/p'} V_p(x)$$

and hence

$$V_p(P_N x) \le (V_p(x)^p + \sup |x_j|^p)^{1/p} \le 3V_p(x).$$

Thus we conclude that (e_n) is a basis of v^0_p. Note that by (7.17) the norm in v_p is equivalent to

$$\max\left\{\sup |x_j|, \sup_{n(0)<n(1)<\cdots} \left(\sum |x_{n(k)} - x_{n(k-1)}|^p\right)^{1/p}\right\}.$$

From this it is easy to see that there is a constant C so that for any sum of disjoint consecutive blocks $b_1, \ldots, b_N$ on (e_n) we have

$$\|b_1 + \cdots + b_N\| \le C(\|b_1\|^p + \cdots + \|b_N\|^p)^{1/p}.$$

By Remark 3.18 the basis (e_n) must be shrinking. □

Proof of Theorem 7.5. In the duality (7.8) (recall (7.18)), the vectors $\sigma_n = \sum_0^n e_k$ are biorthogonal to e_n, i.e. we have $\langle e_k, \sigma_n\rangle = 0$ for all $k \neq n$ and $= 1$ if $k = n$. By Lemma 7.6, (e_n) is a shrinking basis for v_p^0. Therefore (σ_n) is a basis for $(v_p^0)^*$. Note that $\operatorname{span}(\sigma_0, \dots, \sigma_n) = \operatorname{span}(e_0, \dots, e_n)$. By (7.23) we find $(v_p^0)^* = u_{p'}^0$. We already know by (7.19) that $v_p = (u_{p'}^0)^*$. Thus if $X = v_p^0$ we have $X^{**} = v_p$ and hence by (7.14) $\dim(X^{**}/X) = 1$. If $X = v_p$ (resp. $X = u_{p'}^0$) then $X \simeq v_p^0 \oplus \mathbb{K}$ (resp. $X^* \simeq v_p \simeq v_p^0 \oplus \mathbb{K}$) and hence $X^{**} \simeq (v_p^0)^{**} \oplus \mathbb{K} \simeq X \oplus \mathbb{K}$ (resp. $X^{**} \simeq (v_p^0)^* \oplus \mathbb{K} \simeq X \oplus \mathbb{K}$). The other assertions are proved similarly. $\square$

We will now identify the dual of $\mathcal{W}_{p,q}$. Let $\mathcal{W}_{p,q}^0 = (v_1^0, c_0)_{\theta,q}$ with $1-\theta = 1/p$. Note that

$$\mathcal{W}_{p,q}^0 = \mathcal{W}_{p,q} \cap c_0. \tag{7.24}$$

Indeed, since (see (7.7))

$$v_1 \simeq v_1^0 \oplus \mathbb{K} \quad \text{and} \quad c \simeq c_0 \oplus \mathbb{K}.$$

we obviously have

$$\mathcal{W}_{p,q} \simeq (v_1^0, c_0)_{\theta,q} \oplus \mathbb{K} \tag{7.25}$$

where the second coordinate is $x \mapsto \lim x_n$. Therefore (7.24) follows immediately.

By general interpolation (see Remark 5.1) $v_1^0 \cap c_0$ is dense in $\mathcal{W}_{p,q}^0 = (v_1^0, c_0)_{\theta,q}$ $(0 < \theta < 1, 1 \leq q < \infty)$, from which it is easy to see that finitely supported sequences form a dense subspace of $\mathcal{W}_{p,q}^0$. Thus by (7.25) $\mathcal{W}_{p,q}^0$ can be identified with the closure in $\mathcal{W}_{p,q}$ of the space of finitely supported sequences.

Theorem 7.7. *Let (e_n) denote the canonical basis of $\mathbb{K}^{(\mathbb{N})}$, let (e_n^*) be the biorthogonal functionals and let $\sigma_n = \sum_0^n e_j$. Let $1 < p < \infty$ and $1 \leq q < \infty$. Then (e_n) and (σ_n) each form a basis in $\mathcal{W}_{p,q}^0$. If moreover $q > 1$, (e_n^*) is a basis of $(\mathcal{W}_{p,q}^0)^*$. The linear mapping T defined on span$[\sigma_n]$ by $T\sigma_n = e_n^*$ $(n \geq 0)$ extends to an isomorphism from $\mathcal{W}_{p',q'}^0$ onto $(\mathcal{W}_{p,q}^0)^*$. In particular, $\mathcal{W}_{p'}^0$ is isomorphic to $(\mathcal{W}_p^0)^*$.*

Proof. Each of (e_n) and (σ_n) is a basis for both spaces c_0 and v_1^0. By interpolation applied to the partial sum operators, it follows that each is also a basis in $\mathcal{W}_{p,q}^0$ for any $1 < p < \infty$, $1 \leq q < \infty$. Recall the notion of upper p-estimate from Remark 3.18. Obviously, (e_n) satisfies an upper r-estimate in v_1^0 for $r = 1$, but it also satisfies one in c_0 for any r (or say for $r = \infty$). Therefore, by an interpolation argument based on Theorem 5.7, it follows that (e_n) satisfies an upper r-estimate in $\mathcal{W}_{p,q}^0$ for $1 < r < \min(p, q)$. It follows (see Remark 3.18) that (e_n) is shrinking in $\mathcal{W}_{p,q}^0$. Equivalently this means that (e_n^*) is a basis in $(\mathcal{W}_{p,q}^0)^*$. Define $T\colon\ \operatorname{span}[\sigma_n] \to (\mathcal{W}_{p,q}^0)^*$ by $T(\sigma_n) = e_n^*$. By Corollary 5.21 there is a constant C (independent of n) such that for any n and any x in $\operatorname{span}(\sigma_0, \dots, \sigma_n)$ we have

$$C^{-1}\|x\|_{\mathcal{W}_{p',q'}^0} \leq \|T(x)\|_{(\mathcal{W}_{p,q}^0)^*} \leq C\|x\|_{\mathcal{W}_{p',q'}^0},$$

but since $\mathrm{span}(\sigma_n)$ and $\mathrm{span}(e_n^*)$ are dense respectively in $\mathcal{W}^0_{p',q'}$ and $(\mathcal{W}^0_{p,q})^*, T$ extends to an isomorphism from $\mathcal{W}^0_{p',q'}$ to $(\mathcal{W}^0_{p,q})^*$. $\square$

Remark 7.8. One can check, arguing as for Theorem 7.5, that $\dim(X^{**}/X) = 1$ when X is any of the spaces $\mathcal{W}^0_{p,q}$ or $\mathcal{W}_{p,q}$ with $1 < p < \infty$ and $1 < q < \infty$.

Remark 7.9. Moreover, the B-valued analogue of Theorem 7.7 also holds with the obvious adjustments: the dual space $(\mathcal{W}^0_{p,q}(B))^*$ is isomorphic to $\mathcal{W}^0_{p',q'}(B^*)$.

The following result will be crucial in the sequel ([75, 235]).

Lemma 7.10. *Let $0 < \theta < 1$, $p = (1-\theta)^{-1}$ and let B be an arbitrary Banach space. We have bounded inclusions*

$$(v_1(B), \ell_\infty(B)_{\theta,1} \subset v_p(B) \subset (v_1(B), \ell_\infty(B))_{\theta,\infty} \tag{7.26}$$

$$(v_1^0(B), c_0(B))_{\theta,1} \subset u_p^0(B) \subset v_p^0(B) \subset (v_1^0(B), c_0(B))_{\theta,\infty}. \tag{7.27}$$

Consequently, for any $1 < r < p < s < \infty$ we have bounded inclusions

$$\mathcal{W}_r(B) \subset v_p(B) \subset \mathcal{W}_s(B). \tag{7.28}$$

Proof. The second inclusion in (7.27), namely $u_p^0(B) \subset v_p^0(B)$ is clear from the definition of $u_p^0(B)$, since for any b as in (7.15) we have obviously $V_p(b) \leq 2[b]_p$. Let us show $v_p(B) \subset \mathcal{W}_{\theta,\infty} = (v_1(B), \ell_\infty(B))_{\theta,\infty}$. Let $x \in B^{\mathbb{N}}$ with $V_p(x) \leq 1$. Fix $t > 1$. Then let $n(1) = \inf\{n > 0 \mid \|x_n - x_0\| \geq t^{-(1-\theta)}\}$, and let $n(2) < n(3) < \cdots$ be defined similarly by $n(k) = \inf\{n > n(k-1) \mid \|x_n - x_{n(k-1)}\| > t^{-(1-\theta)}\}$. Whenever the preceding infimum runs over the void set we set $n(k) = \infty$ and we stop the process. Since $V_p(x) \leq 1$, the process has to stop at a certain stage k (so that $n(k) < \infty$ but $n(k+1) = \infty$). We have then on one hand

$$t^{-(1-\theta)} k^{1/p} \leq (\|x_{n(1)} - x_0\|^p + \cdots + \|x_{n(k)} - x_{n(k-1)}\|^p)^{1/p} \leq 1$$

and hence $k \leq t$. But on the other hand we can decompose x as $x = x^0 + x^1$ with $x' = x - x^0$ and x^0 defined by

$$x_n^0 = x_{n(j)} \quad \text{if} \quad n(j) \leq n < n(j+1)$$

where we set by convention $n(0) = 0$ and $n(k+1) = \infty$. By definition of $n(0) < n(1) < \cdots$ we have $\|x^1\|_{\ell_\infty(B)} \leq t^{-(1-\theta)}$ and also (recall $k \leq t$)

$$\begin{aligned}\|x^0\|_{v_1(B)} \leq & \|x_0\| + \|x_{n(1)} - x_0\| + \cdots + \|x_{n(k)} - x_{n(k-1)}\| \\ \leq & (k+1)^{1/p'} V_p(x) \leq 2k^{1/p'} \leq 2t^{1/p'} = 2t^\theta,\end{aligned}$$

so we find $K_t(x; v_1(B), \ell_\infty(B)) \leq \|x^0\|_{v_1(B)} + t\|x^1\|_{\ell_\infty(B)} \leq 3t^\theta$. Thus we conclude

$$\|x\|_{\mathcal{W}_{\theta,\infty}(B)} \leq 3\|x\|_{v_p(B)}.$$

Note that if $x \in c_0(B) \cap v_p(B)$ we find x^0 and x^1 in $c_0(B)$ also, so the same argument gives $\|x\|_{(v_1^0(B), c_0(B))_{\theta,\infty}} \leq 3\|x\|_{v_p^0(B)}$. That yields the third inclusion in (7.27). The inclusion $\mathcal{W}_{\theta,1}(B) \subset v_p(B)$ is an immediate consequence of the following simple Hölder type inequality

$$\|x\|_{v_p(B)} \le \|x\|_{v_1(B)}^{1-\theta}(2\|x\|_{\ell_\infty(B)})^\theta,$$

once one recalls Lemma 5.12. Thus we have established (7.26). It only remains to prove the first inclusion in (7.27). But by duality the latter is equivalent to

$$u_p^0(B)^* \subset (v_1^0(B), c_0(B))_{\theta,1}^*$$

and by the duality for real interpolation spaces (Theorem 5.13) and by (7.13) this boils down to

$$u_p^0(B)^* \subset (\ell_\infty(B^*), v_1(B^*))_{\theta,\infty} = (v_1(B^*), \ell_\infty(B^*))_{1-\theta,\infty}.$$

Equivalently, since $u_p^0(B)^* = v_{p'}(B^*)$ this reduces to

$$v_{p'}(B^*) \subset (v_1(B^*), \ell_\infty(B^*))_{1-\theta,\infty}$$

and this is but the second part of (7.26) with $B^*, p', 1-\theta$ in place of B, p, θ.

The last assertion follows from the general fact (see (5.4)) that for an interpolation pair (A_0, A_1) with $A_0 \subset A_1$, for any $0 < \alpha < \theta < \beta$ we have bounded inclusions $(A_0, A_1)_{\alpha,r} \subset (A_0, A_1)_{\theta,1}$ and $(A_0, A_1)_{\theta,\infty} \subset (A_0, A_1)_{\beta,s}$. □

We will denote by $c(B) \subset \ell_\infty(B)$ the subspace formed of all convergent sequences, equipped with the norm induced by $\ell_\infty(B)$.

Note that for any $0 < \theta < 1$, $1 \le q \le \infty$

$$(v_1(B), \ell_\infty(B))_{\theta,q} = (v_1(B), c(B))_{\theta,q}, \tag{7.29}$$

with identical norms. To check this, we first claim that

$$(v_1(B), \ell_\infty(B))_{\theta,q} \subset c(B).$$

Indeed, a basic fact in interpolation theory asserts that $A_0 \cap A_1$ is dense in $(A_0, A_1)_{\theta,q}$ when $q < \infty$ (see Remark 5.1). Applying this to $(A_0, A_1) = (v_1(B), \ell_\infty(B))$, we find that $A_0 = v_1(B)$ is dense in $(A_0, A_1)_{\theta,q}$, but in our specific case $A_0 \subset A_1$, and hence $(A_0, A_1)_{\theta,q}$ is included in the closure of A_0 in A_1 and since $v_1(B) \subset c(B)$, and $c(B)$ is closed in $\ell_\infty(B)$, the latter closure is included in $c(B)$. This proves our claim for $q < \infty$. But if $q = \infty$, we may choose any θ' with $\theta < \theta' < 1$, then for any finite Q we have $(A_0, A_1)_{\theta,\infty} \subset (A_0, A_1)_{\theta',Q} \subset c(B)$ and we obtain the claim also for $q = \infty$. Obviously, since $v_1(B) \subset c(B)$, for any $x \in c(B)$ that is also in $v_1(B) + \ell_\infty(B)$ we have

$$K_t(x; v_1(B), \ell_\infty(B)) = K_t(x; v_1(B), c(B))$$

and hence the norms of x in $(v_1(B), \ell_\infty(B))_{\theta,q}$ and in $(v_1(B), c(B))_{\theta,q}$ coincide. Now, from our claim that $(v_1(B), \ell_\infty(B))_{\theta,q} \subset c(B)$, (7.29) becomes clear.

Definition. For $b = (b(n)) \in c(B)$, let $b(\infty) = \lim b(n) \in B$. We denote by $u_p(B)$ the subspace of $c(B)$ formed of all $b = (b(n))$ such that

$$(b(n) - b(\infty))_{n\in\mathbb{N}} \in u_p^0.$$

We equip $u_p(B)$ with the norm

$$\|b\|_{u_p(B)} = \|b(\infty)\| + \|(b(n) - b(\infty))\|_{u_p^0(B)}.$$

Remark 7.11. Thus $u_p(B) \simeq B \oplus u_p^0(B)$. In the same decomposition we have $v_1(B) \simeq B \oplus v_1^0(B)$ and $c(B) \simeq B \oplus c_0(B)$. Therefore we must have also

$$(v_1(B), c(B))_{\theta,q} \simeq B \oplus (v_1^0(B), c_0(B))_{\theta,q}$$

for any $0 < \theta < 1$ and $1 \le q \le \infty$. In particular (7.27) and (7.29) imply obviously

$$(7.30) \qquad (v_1(B), \ell_\infty(B))_{\theta,1} \subset u_p(B) \subset v_p(B) \subset (v_1(B), \ell_\infty(B))_{\theta,\infty}.$$

We recall the notation

$$\mathcal{W}_p(B) = (v_1(B), \ell_\infty(B))_{\theta,p}$$

with $p = (1-\theta)^{-1}$.

Lemma 7.12. *Let $1 < r < p < s < \infty$. $0 < \alpha, \beta < 1$ be determined by the equalities $\frac{1}{p} = \frac{1-\alpha}{r} + \frac{\alpha}{\infty}$ and $\frac{1}{p} = \frac{1-\beta}{1} + \frac{\beta}{s}$. Then*

$$(7.31) \qquad \mathcal{W}_p(B) = (v_r(B), \ell_\infty(B))_{\alpha,p}$$

$$(7.32) \qquad \mathcal{W}_p(B) = (v_1(B), u_s(B))_{\beta,p}$$

with equivalent norms. More generally, for any $1 \le q \le \infty$, we have
$\mathcal{W}_{p,q}(B) = (v_r(B), \ell_\infty(B))_{\alpha,q}$ *and* $\mathcal{W}_{p,q}(B) = (v_1(B), u_s(B))_{\beta,q}$.

Proof. The key is to use "reiteration". By Lemma 7.10 the reiteration Theorem 5.9 implies (7.31) and (7.32). □

For simplicity, the following definition was kept implicit until now, but we will need to refer to it.

Let $0 < p < \infty, 1 \le q \le \infty$. We denote by $v_{p,q}(B)$ the space of sequences $x = (x_n)$ in $B^{\mathbb{N}}$ such that

$$V_{p,q}(x) = \sup\{\|(x_0, x_{n(1)} - x_{n(0)}, x_{n(2)} - x_{n(1)}, \ldots)\|_{\ell_{p,q}(B)}\} < \infty$$

where the supremum runs over all sequences $0 = n(0) < n(1) < n(2) < \ldots$ of integers, and we equip it with the quasi-norm $x \mapsto V_{p,q}(x)$. Recall here that by (7.5) (easily extended to $p < 1$) we have

$$V_{p,q}(x) \le \|x\|_{\mathcal{W}_{p,q}}.$$

The space $v_{p,\infty}(B)$ corresponds to the sequences with variation in weak-ℓ_p. It corresponds to an "intersection" between the scales $v_{p,q}$ and $\mathcal{W}_{p,q}$ as formulated in the following lemma:

Lemma 7.13. *For any $1 < p < \infty$ and any B we have*

$$v_{p,\infty}(B) = \mathcal{W}_{p,\infty}(B)$$

with equivalent norms.

Proof. By (7.5) we already observed $\mathcal{W}_{p,\infty}(B) \subset v_{p,\infty}(B)$. Conversely, the proof of (7.26) actually shows $v_{p,\infty}(B) \subset (v_1(B), \ell_\infty(B))_{\theta,\infty}$ where $\frac{1}{p} = 1 - \theta$. $\square$

Let $0 < p < \infty$, $0 < q \le \infty$. Let us denote by $\delta_{p,q}(B)$ the space of sequences $x = (x_n) \in B^{\mathbb{N}}$ such that the sequence $y = (y_n)$, defined by $y_0 = x_0$ and $y_n = x_n - x_{n-1}$ for all $n \ge 1$, is in $\ell_{p,q}(B)$. We equip it with the quasi-norm

$$\|x\|_{\delta_{p,q}(B)} = \|y\|_{\ell_{p,q}(B)}.$$

When $0 < r < 1$, the spaces v_r behave slightly surprisingly with respect to interpolation, as the next statement shows.

Theorem 7.14. *Let B be any Banach space. Fix $0 < r < 1$. Let $0 < \theta < 1$, $1 \le q \le \infty$. Let p be determined by $\frac{1}{p} = \frac{1-\theta}{r}$. Then $(v_r(B), \ell_\infty(B))_{\theta,q}$ can be described as follows:*

(i) *If $r < p < 1$ (i.e. $0 < \theta < \alpha$) we have*

$$(v_r(B), \ell_\infty(B))_{\theta,q} = \delta_{p,q}(B)$$

with equivalent norms.

(ii) *If $1 < p < \infty$ (i.e. $\alpha < \theta < 1$) we have*

$$(v_r(B), \ell_\infty(B))_{\theta,q} = \mathcal{W}_{p,q}(B)$$

with equivalent norms.

Proof. Let $X(\alpha, q) = (v_r(B), \ell_\infty(B))_{\alpha,q}$. Since the operator T taking (x_n) to $(x_0, x_1 - x_0, \cdots, x_n - x_{n-1}, \cdots)$ is bounded simultaneously from $v_r(B)$ to $\ell_r(B)$ and from $\ell_\infty(B)$ to itself, it is also bounded (by Theorem 5.2) from $X(\alpha, 1)$ to $(\ell_r(B), \ell_\infty(B))_{\alpha,1} = \ell_1(B)$. Therefore $X(\alpha, 1) \subset v_1(B)$. Then, by the same argument as for Lemma 7.10 we obtain

$$X(\alpha, 1) \subset v_1(B) \subset X(\alpha, \infty)$$

Therefore, by the reiteration Theorem 5.9 (extended to the quasi-normed case see [5, p. 67]) for any $0 < \gamma < 1$, $0 < \delta < 1$ and $1 \le q \le \infty$, we have

$$(v_r(B), v_1(B))_{\gamma,q} = (v_r(B), \ell_\infty(B))_{\theta,q} \tag{7.33}$$

where $\theta = \gamma\alpha$, and

$$(v_1(B), \ell_\infty(B))_{\delta,q} = (v_r(B), \ell_\infty(B))_{\theta,q} \tag{7.34}$$

where $\theta = (1-\delta)\alpha + \delta \doteq \alpha + \delta(1-\alpha)$. As before, define $y = (y_n)$ by $y_0 = x_0$ and $y_n = x_n - x_{n-1}$ for all $n \geq 1$. Since $\|x\|_{v_r(B)} = \|y\|_{\delta_r(B)}$ for any $0 < r \leq 1$, we may identify $v_r(B)$ with $\delta_r(B)$, or equivalently with $\ell_r(B)$, so that by Theorem 5.7 we have

$$(v_r(B), v_1(B))_{\gamma,q} = \delta_{p,q}(B).$$

Thus (7.33) implies (i). By definition of $\mathcal{W}_{p,q}(B)$, (7.34) is the same as (ii). □

7.2 Type and cotype of $\mathcal{W}_p$

In this section, we will show that the spaces $\mathcal{W}_p$ satisfy an analogue of the Hölder–Minkowski inequality (see Appendix 2 in Chapter 8). The latter refers to the fact that, assuming $1 \leq r \leq p \leq \infty$, for any measure spaces (Ω_1, μ_1), (Ω_2, μ_2) we have a norm 1 inclusion

$$L_r(\mu_1; L_p(\mu_2)) \subset L_p(\mu_2; L_r(\mu_1)). \tag{7.35}$$

Note that the reverse inclusion holds when $p \leq r$ and when $p = r$, Fubini's theorem gives us isometrically

$$L_p(\mu_1; L_p(\mu_2)) \simeq L_p(\mu_2; L_p(\mu_1)) \simeq L_p(\mu_1 \times \mu_2).$$

Although this is very special to (and in some sense characteristic of) L_p-spaces, it turns out that the space $\mathcal{W}_p$ satisfies an analogous property: If $r < p$ we have a bounded inclusion $L_r(\mathcal{W}_p) \subset \mathcal{W}_p(L_r)$, while if $p < r$ we have the reverse $\mathcal{W}_p(L_r) \subset L_r(\mathcal{W}_p)$. There is however (necessarily) a singularity when $r = p$ that reflects the non-reflexivity of $\mathcal{W}_p$.

Theorem 7.15. *Let (Ω, μ) be any measure space and B any Banach space. For simplicity, we write $L_p(B)$ instead of $L_p(\Omega, \mu; B)$. Let $1 < p < \infty$. For any $r < p < s$ we have the following bounded natural inclusions:*

$$L_r(\mathcal{W}_p(B)) \subset \mathcal{W}_p(L_r(B)) \tag{7.36}$$

$$\mathcal{W}_p(L_s(B)) \subset L_s(\mathcal{W}_p(B)). \tag{7.37}$$

Proof. We first observe that this can be easily reduced to the case of an atomic measure space with finitely many atoms, and this allows us to ignore all measurability considerations since we may as well assume $L_r = \ell_r^n$ and $L_s = \ell_s^n$. Let $0 < \alpha, \beta < 1$ be as in Lemma 7.12. Now observe that the following inclusions both hold with norm ≤ 1

$$L_r(v_r(B)) \subset v_r(L_r(B)) \quad \text{and} \quad L_r(\ell_\infty(B)) \subset \ell_\infty(L_r(B)).$$

Therefore by interpolation we have

$$(L_r(v_r(B)), L_r(\ell_\infty(B)))_{\alpha,p} \subset (v_r(L_r(B)), \ell_\infty(L_r(B)))_{\alpha,p}$$

but by (7.31) the last space coincides with $\mathcal{W}_p(L_r(B))$ and by Remark 5.7 since $p > r$ we have

$$L_r(\mathcal{W}_p(B)) \subset (L_r(v_r(B)), L_r(\ell_r(B)))_{\alpha,p}$$

and hence (7.36) follows.

The proof of (7.37) is entirely similar but with the inclusions reversed. By the duality between $v_p(B^*)$ and $u^0_{p'}(B)$ we have

$$u^0_s(L_s(B)) \subset L_s(u^0_s(B)),$$

or equivalently (see Remark 7.11) $u_s(L_s(B)) \subset L_s(u_s(B))$, and obviously also

$$v_1(L_s(B)) \subset L_s(v_1(B)).$$

Therefore by Remark 5.7 again since $s > p$

$$(v_1(L_s(B)), u_s(L_s(B)))_{\beta,p} \subset L_s((v_1(B), u_s(B))_{\beta,p}) = L_s(\mathcal{W}_p(B)),$$

and by (7.32) we obtain (7.37). □

Remark. By Remark 7.9, it is easy to see that (7.36) and (7.37) are actually equivalent by duality.

The next Corollary shows how to apply our study of the spaces $\mathcal{W}_{p,q}$ to the more classical spaces v_p. The main point is the fact that the two scales are intertwined in the form expressed by Lemma 7.10.

Corollary 7.16. *In the situation of Theorem 7.15, let $1 < p_0 < p_1 < \infty$. Then for any r, s such that $1 < r < p_0 < p_1 < s < \infty$ we have the following bounded inclusions:*

$$v_{p_0}(L_s(B)) \subset L_s(v_{p_1}(B))$$
$$L_r(v_{p_0}(B)) \subset v_{p_1}(L_r(B)).$$

Proof. Pick p such that $p_0 < p < p_1$. We have (by (7.6) and Lemma 7.10) bounded inclusions $v_{p_0}(B) \subset \mathcal{W}_p(B) \subset v_1(B)$. Moreover, this holds for any B (and hence also with $L_r(B)$ or $L_s(B)$ in place of B). Therefore the result follows immediately from Theorem 7.15. □

Lemma 7.17. *Let $T\colon\ B_1 \to B_2$ be a bounded operator between Banach spaces. Then, for any $1 < p < \infty$, T extends "naturally" to a bounded operator $\widetilde{T}\colon$ $\mathcal{W}_p(B_1) \to \mathcal{W}_p(B_2)$ taking $x = (x_n)_{n\geq 0} \in \mathcal{W}_p(B_1)$ to $(Tx_n)_{n\geq 0}$, and moreover $\|\widetilde{T}\| = \|T\|$.*

Proof. This is a direct application of the fundamental interpolation principle (cf. Theorem 5.2): indeed we have clearly $\|\widetilde{T}\colon\ v_1(B_1) \to v_1(B_2)\| \leq \|T\|$ and $\|\widetilde{T}\colon\ \ell_\infty(B_1) \to \ell_\infty(B_2)\| \leq \|T\|$, therefore $\|\widetilde{T}\colon\ \mathcal{W}_p(B_1) \to \mathcal{W}_p(B_2)\| \leq \|T\|$. The converse is obvious by considering the action of $\widetilde{T}$ on sequences (x_n) such that $x_0 \in B_1$ and $x_n = 0$ for all $n > 0$. □

Corollary 7.18. *Let $(\Omega_1,\mu_1),(\Omega_2,\mu_2)$ be two measure spaces. With the notation of Theorem 7.15, any bounded linear operator $T\colon\ L_r(\mu_1)\to L_s(\mu_2)$ extends to a bounded operator*

$$\widetilde{T}\colon\ L_r(\mu_1;\mathcal{W}_p)\to L_s(\mu_2;\mathcal{W}_p)$$

such that $T(f\otimes x)=T(f)\otimes x$ $(f\in L_r(\mu_1), x\in\mathcal{W}_p)$.

Proof. By Theorem 7.15, it suffices to show that $\widetilde{T}$ is bounded from $\mathcal{W}_p(L_r)$ to $\mathcal{W}_p(L_s)$, and this follows from the preceding Lemma. □

Corollary 7.19. *Let $1<r<p<s<\infty$. Then $\mathcal{W}_p$ is of type $r\wedge 2$ and of cotype $s\vee 2$ for any $1<p<\infty$.*

Proof. Let $\Omega=\{-1,1\}^{\mathbb{N}}$ equipped with its usual probability μ. Consider the operator

$$T\colon\ \ell^{r\wedge 2}\to L^{s\vee 2}$$

defined by $T((\alpha_n))=\sum\alpha_n\varepsilon_n$. By the Khintchine inequalities (cf. (8.7)), T is bounded, and hence so is $\widetilde{T}$ by the preceding Corollary, and that means $\mathcal{W}_p$ is of type $r\wedge 2$. We may argue similarly with the operator $T\colon\ L_{r\wedge 2}\to\ell_{s\vee 2}$ defined by $T(f)=(\int f\varepsilon_n\, d\mu)_{n\geq 0}$ and this shows that $\mathcal{W}_p$ is of cotype $s\vee 2$. □

Note that, by the classical Kwapień theorem, a Banach space is of type 2 and cotype 2 iff it is isomorphic to a Hilbert space. In particular, type 2 and cotype 2 forces reflexivity. However, we now can state:

Corollary 7.20. *For any $\varepsilon>0$, there are non-reflexive Banach spaces of type 2 and of cotype $2+\varepsilon$.*

Recall that the Banach-Mazur $d(E,F)$ between two (isomorphic) Banach spaces is defined by

$$d(E,F)=\inf\{\|u\|\|u^{-1}\|\}$$

where the infimum runs over all possible isomorphisms $u\colon\ E\to F$.

Remark 7.21. The space $\mathcal{W}_2$ has several remarkable properties reminiscent of Hilbert space: it is isomorphic to its dual and moreover there is a constant C such that any n-dimensional subspace $(n>1)$ $E\subset\mathcal{W}_2$ satisfies $d(E,\ell_2^n)\leq C\mathrm{Log}\ n$. This logarithmic growth is sharp. Indeed $\mathcal{W}_2$ is of course non-reflexive (see Remark 7.1) but any Banach space X for which the function $f(n)=\sup\{d(E,\ell_2^n)\mid E\subset X\}$ is $o(\mathrm{Log}(n))$ must be reflexive! We refer to [235] for more details.

7.3 Strong p-variation in approximation theory

For any $x\in B^{\mathbb{N}}$ we denote

$$V_{p,N}(x)=\sup\{(\|x_0\|^p+\|x_{n(1)}-x_0\|^p+\cdots+\|x_{n(N)}-x_{n(N-1)}\|^p)^{1/p}\}$$

where N is fixed and the supremum runs over all increasing N-tuples of integers $n(1) < n(2) < \cdots < n(N)$. Note that

$$V_{p,N}(x) \leq 2(1+N)^{1/p}\|x\|_\infty. \tag{7.38}$$

Lemma 7.22 ([75])**.** *Let $1 \leq r < \infty$. For any x in $v_r(B) + \ell_\infty(B)$ we have for any $N \geq 1$*

$$2^{-1-1/r}V_{r,N}(x) \leq K_{N^{1/r}}(x, v_r(B), \ell_\infty(B)) \leq 2V_{r,N}(x).$$

Proof. For simplicity we set $K_t(x) = K_t(x, v_r(B), \ell_\infty(B))$. We have obviously $V_{r,N}(x) \leq V_r(x)$ and $V_{r,N}(x) \leq 2(N+1)^{1/r}\|x\|_\infty$. Therefore if $x = x_0 + x_1$ we can write

$$V_{r,N}(x) \leq V_{r,N}(x_0) + V_{r,N}(x_1) \leq \|x_0\|_{v_r(B)} + 2(N+1)^{1/r}\|x_1\|_\infty$$

and hence $V_{r,N}(x) \leq 2^{1+1/r}K_{N^{1/r}}(x)$.

For the converse inequality, we use the same idea as in the above proof of Lemma 7.10. By homogeneity we may assume $V_{r,N}(x) = 1$. We let $n(1) = \inf\{n \mid \|x_n - x_0\| > N^{-1/r}\}$, $n(2) = \inf\{n > n(1) \mid \|x_n - x_{n(1)}\| > N^{-1/r}\}$ and so on. The process will stop at some integer k. Note that $N^{-1/r}k^{1/r} < V_{r,N}(x) = 1$ and hence $k < N$. We then define $x^0_n = x_0$ on $[0, n(1)]$, $x^0_n = x_{n(k-1)}$ if $n \in]n(k-1), n(k)]$ and $x^1 = x - x^0$. Then $\|x^1\|_\infty \leq N^{-1/r}$ and, since $k < N$, $\|x^0\|_{v_r(B)} \leq V_{r,N}(x) = 1$. Thus we obtain

$$K_{N^{1/r}}(x) \leq \|x^0\|_{v_r(B)} + N^{1/r}\|x^1\|_\infty \leq 2. \qquad \square$$

Remark. Actually the preceding Lemma remains valid for $0 < r < 1$ with possibly different constants, with the same proof. In that case, the space v_r is only a quasi-normed space.

Now that we have a more concrete description of the K-functional, we can give a rather nice one for the interpolation spaces $\mathcal{W}_{p,q}(B)$:

Theorem 7.23. *Assume $1 \leq r < p < \infty$, $1 \leq q \leq \infty$. A sequence $x = (x_n)$ in $B^{\mathbb{N}}$ belongs to $\mathcal{W}_{p,q}(B)$ iff the sequence $(N^{-1/r}V_{r,N}(x))_{N\geq 1}$ is in $\ell_{p,q}$ and the corresponding norms (or quasi-norms) are equivalent.*

Proof. We use Lemma 7.12. Let $a_N(x) = N^{-1/r}V_{r,N}(x)$. Simply observe that if $\frac{1}{p} = \frac{1-\alpha}{r} + \frac{\alpha}{\infty}$ and $1 \leq q < \infty$, by "change of variable" (we replace t by $N^{1/r}$)

$$\int_0^\infty (t^{-\alpha}K_t(x, v_r(B), \ell_\infty(B))^q \frac{dt}{t} \simeq \sum\nolimits_{N\geq 1}(N^{-\alpha/r}K_{N^{1/r}}(x, v_r(B), \ell_\infty(B)))^q N^{-1}$$
$$\simeq \sum\nolimits_{N\geq 1}(N^{1/p}a_N(x))^q N^{-1},$$

and the result follows by Remark 5.6. $\square$

Remark 7.24. The preceding result shows that, for any B and any closed subspace $S \subset B$, $\mathcal{W}_{p,q}(S)$ is a closed subspace of $\mathcal{W}_{p,q}(B)$ and its norm is equivalent to the one induced on it by $\mathcal{W}_{p,q}(B)$.

Remark 7.25. One can derive an alternate proof of (7.36) from Theorem 7.23 (or from Lemma 7.22). Then (7.37) follows by a duality argument.

Another useful description of the space $\mathcal{W}_p$ can be given in terms of approximation theory. Actually, it would be more natural (as is done in [75]) to work with functions on $[0,1]$ and to consider approximation by splines, but we prefer to stick to our "discrete" setting.

Let $\mathcal{S}_N \subset \ell_\infty(B)$ be the subset formed of all $b = (b(n))_{n\geq 0}$ such that $\mathbb{N}$ can be partitioned into N intervals on each of which b is constant. Note $\mathcal{S}_N \subset \mathcal{S}_{N+1}$. Then let

$$\forall x \in \ell_\infty(B) \qquad S_N(x) = \inf\{\|x-b\|_\infty \mid b \in \mathcal{S}_N\}.$$

This is simply the distance of x in $\ell_\infty(B)$ to $\mathcal{S}_N$.

For any $x \in \mathcal{S}_N$ we have obviously

$$V_1(x) = \sup_k V_{1,k}(x) = V_{1,N}(x) \leq (1+2N)\|x\|_\infty \tag{7.39}$$

Theorem 7.26. *Let $1 < p < \infty$. The following properties of a sequence $x \in B^{\mathbb{N}}$ are equivalent:*

(i) $x \in \mathcal{W}_p(B)$.

(ii) $\sum_N S_N(x)^p < \infty$.

Moreover the corresponding quasi-norm $x \mapsto (\|x_0\|^p + \sum_{N\geq 1} S_N(x)^p)^{1/p}$ is equivalent to the norm (namely $V_p(x)$) in the space $\mathcal{W}_p(B)$.

Proof. The proof of Lemma 7.22 actually shows that $S_N(x) \leq N^{-1/r} V_{r,N}$ (indeed in that proof $x^0 \in \mathcal{S}_N$). Therefore by Theorem 7.23, (i) implies (ii). Conversely, assume (ii). Note that $\mathcal{S}_n + \mathcal{S}_k \subset \mathcal{S}_{n+k}$ for any $n, k \geq 1$, and also $\sum S_N(x)^p \simeq \sum 2^n S_{2^n}(x)^p < \infty$. Let $x^{(n)} \in \mathcal{S}_n$ be such that $\|x - x^{(n)}\|_\infty < 2S_{2^n}(x)$. Let $\Delta_n = x^{(n)} - x^{(n+1)}$ and $x^{(0)} = 0$. Note $x = \sum_0^\infty \Delta_n$ and $\Delta_n \in \mathcal{S}_{2^n+2^{n+1}} \subset \mathcal{S}_{2^{n+2}}$. Therefore, by (7.39), we have $V_{1,2^k}(\Delta_n) \leq 2^{n+3}\|\Delta_n\|_\infty$ that we will use when $2^k \geq 2^{n+2}$, while, for any k, we already saw in (7.38) that $V_{1,2^k}(\Delta_n) \leq (2^{k+1}+1)\|\Delta_n\|_\infty$. We have

$$\begin{aligned} V_{1,2^k}(x) &\leq V_{1,2^k}\left(\sum_{n\leq k-2} \Delta_n\right) + V_{1,2^k}\left(\sum_{n>k-2} \Delta_n\right), \\ &\leq \sum_{n\leq k-2} V_{1,2^k}(\Delta_n) + \sum_{n>k-2} (2^{k+1}+1)\|\Delta_n\|_\infty \\ &\leq \sum_{n\leq k-2} 2^{n+3}\|\Delta_n\|_\infty + (2^{k+1}+1)\sum_{n>k-2} \|\Delta_n\|_\infty. \end{aligned}$$

But $\|\Delta_n\|_\infty \leq \|x^{(n)} - x\|_\infty + \|x - x^{(n+1)}\|_\infty \leq 4S_{2^n}(x)$ so we find for some constant c an estimate of the form

$$2^{-k} V_{1,2^k}(x) \leq c\left(2^{-k}\sum\nolimits_{n\leq k} 2^n S_{2^n}(x) + \sum\nolimits_{n\geq k} S_{2^n}(x)\right),$$

or equivalently for any $N \geq 1$ (with a different c)

$$N^{-1}V_{1,N}(x) \leq c\left(N^{-1}\sum\nolimits_{n\leq N} S_n(x) + \sum\nolimits_{n\geq N} n^{-1}S_n(x)\right).$$

From this, elementary arguments show that

$$\sum S_N(x)^p < \infty \Rightarrow \sum (N^{-1}V_{1,N}(x))^p < \infty.$$

Indeed, by Hardy's classical inequality, for any $1 < p < \infty$ and for any scalar sequence $(a_1, a_2, \cdots)$, we have

$$\|(N^{-1}\sum\nolimits_{n\leq N} a_n)_{N\geq 1}\|_{\ell_p} \leq p'\|(a_n)\|_{\ell_p}$$

and therefore by duality also for any $1 < p' < \infty$ and any sequence $(b_1, b_2, \cdots)$

$$\|(\sum\nolimits_{N\geq n} b_N/N)_{n\geq 1}\|_{\ell_{p'}} \leq p'\|(b_N)\|_{\ell_{p'}},$$

and hence $\sum S_N(x)^p < \infty \Rightarrow x \in \mathcal{W}_p$ by Theorem 7.23. □

Remark 7.27. The preceding proof shows that the properties in Theorem 7.26 are also equivalent to

(iii) For each integer $n \geq 1$, there are $\Delta_n \in \mathcal{S}_{2^n}$ such that $x = \sum_n \Delta_n$ and $\sum 2^n \sup_{k\geq 2^n} \|\Delta_k\|^p_{\ell_\infty(B)} < \infty$.

Remark 7.28. Let B, B_1 be arbitrary Banach spaces. Let $Q\colon\ B \to B_1$ be a bounded surjection onto B_1 so that $B_1 \simeq B/\ker(Q)$. Then for any $1 < p < \infty$ the associated map $I \otimes Q$ is a surjection from $\mathcal{W}_p(B)$ onto $\mathcal{W}_p(B_1)$. Indeed, this is an easy consequence of the preceding remark. This lifting (or "projective") property can also be proved by duality using Remarks 7.24 and 7.9.

Throughout this chapter we have collected a wealth of information on the real interpolation spaces $\mathcal{W}_{p,q}$. In sharp contrast, the complex analogue remains a long standing open question:

Problem: Describe the complex interpolation spaces between the complex valued versions of v_1 and ℓ_∞.

7.4 Notes and Remarks

This chapter is mainly based on [235]. A key idea comes from Bergh and Peetre's [75]: There they prove Lemma 7.22 and (7.26) in the scalar case but the Banach valued case is identical. As mentioned in the text, the classical James space J is the one that we denote by v_2^0. Theorem 7.5 and Lemma 7.6 are due to James. See [231] for a proof that J^* is of cotype 2.

Our approach can be applied equally well to the couple of function spaces $(V_1(I; B), \ell_\infty(I; B))$ when $I \subset \mathbb{R}$ is an interval (in particular when $I = \mathbb{R}$). Here the definition of $V_p(I;B)$ $(0 < p < \infty)$ is exactly the same as for sequences, or equivalently a function $f\colon\ I \to B$ is in $V_p(I;B)$ iff for any increasing mapping $T\colon\ \mathbf{N} \to \mathbf{I}$, the composition $f \circ T\colon\ n \mapsto f(T(n))$ is in $v_p(B)$ and $\|f\|_{V_p(I;B)}$ is equivalent to $\sup\{\|f \circ T\|_{v_p(B)}\}$ where the sup runs over all possible such increasing mappings T. In case $I = \mathbb{R}$, it is natural to define $V_p^0(I;B)$ as the closure of the subset of compactly supported infinitely differentiable functions, and to replace (as we did for sequence spaces) $V_p(I;B)$ by $V_p^0(I;B)$. See [75] for connections with approximation by splines.

We then define exactly as before $\mathcal{W}_{p,q}(I;B) = (V_1(I;B), \ell_\infty(I;B))_{\theta,q}$. Many results of this chapter remain valid, for instance this is the case for Theorem 7.15, Corollaries 7.16 to 7.19 and those in §**??**. Among the few references we know (besides [231]) that study the Banach spaces of functions with finite strong p-variation, we should mention [11] and [180].

Chapter 8

The UMD property for Banach spaces

8.1 Martingale transforms (scalar case) Burkholder's inequalities

Let $(M_n)_{n\geq 0}$ be a scalar valued martingale on a filtration $(\mathcal{A}_n)_{n\geq 0}$. We will always set $dM_0 = M_0$ (or we make the convention that $M_{-1} \equiv 0$) and

$$\forall n \geq 1 \quad dM_n = M_n - M_{n-1}.$$

When there is no ambiguity, we will often denote by $\mathbb{E}_n$ the conditional expectation relative to $\mathcal{A}_n$. Moreover we will sometimes say n-measurable instead of $\mathcal{A}_n$-measurable.

Let $(\varphi_n)_{n\geq 0}$ be an arbitrary sequence of random variables, that we merely assume to be *adapted* to $(\mathcal{A}_n)_{n\geq 0}$, i.e. we assume that φ_n is $\mathcal{A}_n$-measurable for each $n \geq 0$.

Let then $\widetilde{M}_0 = \varphi_0 M_0$ and

$$\forall n \geq 1 \qquad \widetilde{M}_n = \varphi_0 M_0 + \sum\nolimits_1^n \varphi_{n-1}\, dM_n.$$

Clearly $\widetilde{M}$ is a martingale and the correspondence $M \to \widetilde{M}$ is called a martingale transform.

An adapted sequence (ψ_n) is called *predictable* if ψ_n is $\mathcal{A}_{n-1}$-measurable for each $n \geq 0$. In the above, we crucially used the fact that (ψ_n) defined by $\psi_n = \varphi_{n-1}$ (and say $\psi_0 = 0$) is predictable.

The key property about these transforms is that, if $\sup\limits_n \|\varphi_n\|_\infty < \infty$, then $M \to \widetilde{M}$ is bounded on L_p for all $1 < p < \infty$ and is of weak type (1-1). This is due to Burkholder as well as the corresponding inequalities: for each $1 \leq p < \infty$ there is a constant β_p such that if $1 < p < \infty$

$$\sup \|\widetilde{M}_n\|_p \leq \beta_p \sup \|\varphi_n\|_\infty \sup \|M_n\|_p \tag{8.1}$$

and if $p = 1$

$$\sup_{\lambda>0} \lambda \mathbb{P}(\sup |\widetilde{M}_n| > \lambda) \leq \beta_1 \sup \|\varphi_n\|_\infty \sup \|M_n\|_1. \tag{8.2}$$

For the proof, see Theorem 8.18 and Corollary 8.14 below.

By Doob's maximal inequality, (8.1) implies that, if $\sup \|\varphi_n\|_\infty \leq 1$, we have

$$(\mathbb{E} \sup |\widetilde{M}_n|^p)^{1/p} \leq \beta'_p (\mathbb{E} \sup |M_n|^p)^{1/p}, \tag{8.3}$$

where $\beta'_p = p'\beta_p$. In this form, (8.3) remains valid when $p = 1$. Namely, there is a constant β'_1 such that

$$\mathbb{E} \sup |\widetilde{M}_n| \leq \beta'_1 \mathbb{E} \sup |M_n|. \tag{8.4}$$

This and (8.10) below are known as B. Davis's inequality. See Corollary 8.26 below.

This is already of interest when each of the variables φ_n is constant and in that special case (8.1) expresses the fact that the sequence $(dM_n)_{n\geq 0}$ is an *unconditional* basic sequence in L_p, i.e. the convergence in L_p of the series $\sum dM_n$ is automatically unconditional. Let $\varepsilon = (\varepsilon_n)_n$ be a fixed choice of signs, i.e. $\varepsilon_n = \pm 1$. Then (8.1) implies for any (M_n) converging in L_p

$$\left\| \sum \varepsilon_n \, dM_n \right\|_p \leq \beta_p \left\| \sum dM_n \right\|_p. \tag{8.5}$$

Replacing dM_n by $\varepsilon_n \, dM_n$ in (8.5) we find the reverse inequality

$$\left\| \sum dM_n \right\|_p \leq \beta_p \left\| \sum \varepsilon_n \, dM_n \right\|_p. \tag{8.6}$$

Let us now introduce the uniform probability μ on $\{-1,1\}^{\mathbb{N}}$ and recall the classical Khintchine inequalities: For any $0 < p < \infty$ there are constants $A_p > 0$ and $B_p > 0$ such that for any sequence $x = (x_n)$ in ℓ_2 we have

$$A_p \left(\sum |x_n|^2\right)^{1/2} \leq \left(\int \left|\sum x_n \varepsilon_n\right|^p d\mu(\varepsilon)\right)^{1/p} \leq B_p \left(\sum |x_n|^2\right)^{1/2}. \tag{8.7}$$

Then if we integrate (8.5) and (8.6) (after raising to the p-th power) we find

$$A_p \beta_p^{-1} \|S\|_p \leq \left\| \sum dM_n \right\|_p \leq \beta_p B_p \|S\|_p$$

where S is the so-called "square function" defined by

$$S = \left(|M_0|^2 + \sum\nolimits_1^\infty |dM_n|^2\right)^{1/2}. \tag{8.8}$$

A similar argument can be applied to (8.2) and it yields a constant β'_1 such that

$$\sup_{\lambda>0} \lambda \mathbb{P}(S > \lambda) \leq \beta'_1 \sup \|M_n\|_1. \tag{8.9}$$

Moreover, using the Khintchine inequality (8.7) for $p = 1$, and also (1.23), we find

$$A_1 (\beta'_1)^{-1} \mathbb{E} S \leq \mathbb{E} \sup |M_n| \leq 2\beta'_1 \mathbb{E} S. \tag{8.10}$$

8.2 Kahane's inequalities

In the Banach space, the square function $\sup\limits_n \left(\sum_0^n |d_k|^2\right)^{1/2}$ must be replaced by

$$\sup_n \left(\int\limits_D \left\|\sum\nolimits_0^n \varepsilon_k d_k\right\|_B^2 d\nu\right)^{1/2} \tag{8.11}$$

where $D = \{-1,1\}^{\mathbb{N}}$ equipped with its usual probability ν and where $\varepsilon_n = D \to \{-1,1\}$ denotes the n-th coordinate.

When B is a Hilbert space, for all x_k in B we have

$$\int \left\|\sum\nolimits_0^n \varepsilon_k x_k\right\|^2 = \sum\nolimits_0^n \|x_k\|^2$$

and hence we recover the square function, but in general this is not possible and we must work with (8.11). We will show in the next section that for the Banach spaces with the UMD property, the Burkholder inequality remains valid when the square function is replaced by (8.11).

This motivates a preliminary study of averages such as (8.11) in a general Banach space when d_k are constant.

Theorem 8.1 (Kahane)**.** *For any $0 < p < q < \infty$ there is a constant $K(p,q)$ such that for any Banach space B and any finite subset $x_1, \ldots, x_n$ in B we have*

$$\left\|\sum \varepsilon_k x_k\right\|_{L_q(B)} \le K(p,q) \left\|\sum \varepsilon_k x_k\right\|_{L_p(B)}.$$

In particular $\left\|\sum \varepsilon_k x_k\right\|_{L_2(B)}$ is equivalent to $\left\|\sum \varepsilon_k x_k\right\|_{L_p(B)}$ for any $0 < p < \infty$.

Remark 8.2. Consider the Banach space $\tilde{B}$ formed of all sequences $x = (x_n)$ of elements of B such that $\sup_n \|\sum_0^n x_k\| < \infty$, equipped with the norm

$$\|x\|_{\tilde{B}} = \sup_n \left\|\sum_{k=0}^n x_k\right\|.$$

Then applying Kahane's inequality to the Banach space $\tilde{B}$ we immediately get that, if we denote $S^* = \sup_n \|\sum_{k=0}^n \varepsilon_k x_k\|$, then for any sequence (x_n) in B we have

$$\|S^*\|_q \le K(p,q)\|S^*\|_p.$$

We will base the proof of Kahane's Theorem on the classical hypercontractive inequality on 2-point space made famous by Nelson and Beckner [72] (but first proved in [80]), as follows.

Theorem 8.3. *Let $1 < p < q < \infty$. Let $\xi = ((p-1)/(q-1))^{1/2}$. Let B be an arbitrary Banach space. Then*

$$\forall x, y \in B \qquad \left(\frac{\|x+\xi y\|^q + \|x-\xi y\|^q}{2}\right)^{1/q} \le \left(\frac{\|x+y\|^p + \|x-y\|^p}{2}\right)^{1/p}.$$

Proof. Let $\Omega = \{-1, 1\}, \mathbb{P} = (\delta_1 + \delta_{-1})/2$. Let $\varepsilon_1\colon\ \Omega \to \{-1, 1\}$ be the identity map. The proof actually reduces to the case $B = \mathbb{R}$. Indeed, let $T\colon\ L_p(\Omega, \mathbb{P}) \to L_q(\Omega, \mathbb{P})$ be the operator defined by $T1 = 1$, $T\varepsilon_1 = \xi\varepsilon_1$. Then $T \geq 0$ ($T =$ convolution by $1 + \xi\varepsilon_1$ and $1 + \xi\varepsilon_1 \geq 0$!). Thus the passage from $\mathbb{R}$ to a general B follows from Proposition 1.4. For the proof in the scalar case, we refer to the appendix (still to be written). □

Consider now $D = \{-1, 1\}^{\mathbb{N}}$ equipped with ν as before, let $(\varepsilon_n)_{n\geq 0}$ denote the coordinates and moreover for any finite subset $S \subset \mathbb{N}$, let

$$w_S = \prod_{n\in S} \varepsilon_n.$$

Corollary 8.4. *Let $1 < p < q < \infty$ and let $\xi = ((p-1)/(q-1))^{1/2}$ as before with B arbitrary. Then for any family $\{x_S \mid S \subset \{1, \ldots, n\}\}$ in B we have*

$$\left\|\sum \xi^{|S|} w_S x_S\right\|_{L_q(B)} \leq \left\|\sum w_S x_S\right\|_{L_p(B)}. \tag{8.12}$$

In particular for any $x_1, \ldots, x_n$ in B

$$\left\|\sum \varepsilon_k x_k\right\|_{L_q(B)} \leq ((q-1)/(p-1))^{1/2} \left\|\sum \varepsilon_k x_k\right\|_{L_p(B)}. \tag{8.13}$$

Proof. The proof is based on the following elementary observation: let

$$T_1\colon\ L_p(\Omega_1, \mu_1; B) \to L_q(\Omega_1', \mu_1'; B)$$

and

$$T_2\colon\ L_p(\Omega_2, \mu_2; B) \to L_q(\Omega_2', \mu_2'; B)$$

be s (i.e. with norms ≤ 1). Then, if $q \geq p$, the operator $T_1 \otimes T_2\colon\ L_p(\mu_1 \times \mu_2; B) \to L_q(\mu_1' \times \mu_2'; B)$ also has norm ≤ 1. (To check this one uses the classical Hölder–Minkowski inequality that says that we have a norm 1 inclusion $L_p(\mu; L_q(\mu')) \subset L_q(\mu'; L_p(\mu))$, see (8.79).)

It follows from this observation by iteration that $T \otimes T \otimes \cdots \otimes T$ (n times) is a contraction from $L_p(B)$ to $L_q(B)$, and since this operator multiplies $w_S x_S$ by $\xi^{|S|}$ we obtain (8.12) and hence (8.13). □

Proof of Kahane's Theorem. The preceding corollary already covers the case $1 < p < q < \infty$ with $K(p, q) = ((q-1)/(p-1))^{1/2}$. In particular if we set $f(\cdot) = \left\|\sum \varepsilon_k(\cdot) x_k\right\|$, we have proved for $1 < p < q < \infty$

$$\|f\|_q \leq K(p, q)\|f\|_p.$$

Let $0 < r < 1 < p < q$. Define $0 < \theta < 1$ by the identity $\frac{1}{p} = \frac{1-\theta}{q} + \frac{\theta}{r}$. Then by Hölder we find

$$\|f\|_q \leq K(p, q)\|f\|_q^{1-\theta}\|f\|_r^{\theta}$$

hence after division by $\|f\|_q^{1-\theta}$, we obtain

$$\|f\|_q \leq K(p, q)^{1/\theta}\|f\|_r$$

which yields $K(r, q) \leq K(p, q)^{1/\theta}$. □

We might as well record here an elementary "contraction principle": Assume $1 \le p \le \infty$. Let B be an arbitrary Banach space. Then $\forall x_1, \ldots, x_n \in B$ $\forall \alpha_1, \ldots, \alpha_n \in \mathbb{R}$

$$\left\| \sum \alpha_k \varepsilon_k x_k \right\|_{L_p(B)} \le \sup |\alpha_k| \left\| \sum \varepsilon_k x_k \right\|_{L_p(B)}; \tag{8.14}$$

moreover $\forall \beta_1, \ldots, \beta_n \in \mathbb{C}$

$$\left\| \sum \beta_k \varepsilon_k x_k \right\|_{L_p(B)} \le 2 \sup |\beta_k| \left\| \sum \varepsilon_k x_k \right\|_{L_p(B)}. \tag{8.15}$$

To verify this, note that by convexity the supremum of the left side of (8.14) over all (α_k) in $\mathbb{R}^n$ with $\sup |\alpha_k| \le 1$ is attained on an extreme point, i.e. an element of $\{-1, 1\}^n$, for which (8.14) becomes an equality. This proves (8.14). To verify (8.15), simply write $\beta_k = \alpha'_k + i\alpha'_k$ and use the triangle inequality.

Lemma 8.5. *Let $0 < p < q < \infty$. Let $\mathcal{F}$ be a subset of $L_q(\Omega, \mathcal{A}, \mathbb{P})$. Assume that there is $C > 0$ such that*

$$\forall f \in \mathcal{F} \qquad \|f\|_q \le C\|f\|_p.$$

Then there are $\delta > 0$ and $R > 0$ such that

$$\forall f \in \mathcal{F} \qquad \mathbb{P}(|f| > R\|f\|_q) \ge \delta.$$

Proof. Let r be such that $p^{-1} = q^{-1} + r^{-1}$. Replacing f by $f\|f\|_q^{-1}$ we may assume that $\|f\|_q = 1$ for all f in $\mathcal{F}$. By Hölder's inequality for any $R > 0$ we have

$$\|f 1_{\{|f|>R\}}\|_p \le (\mathbb{P}(|f| > R))^{1/r}.$$

Hence we can write

$$\begin{aligned} 1 = \|f\|_q \le C\|f\|_p &\le C\|f 1_{\{|f| \le R\}}\|_p + C\|f 1_{\{|f|>R\}}\|_p \\ &\le CR + C(\mathbb{P}(|f| > R))^{1/r}. \end{aligned}$$

Thus if we choose $R = (2C)^{-1}$ and $\delta = (2C)^{-r}$ we obtain the announced result. □

Let f_n and f be B-valued random variables. Recall that, by definition, f_n converges to f in probability if

$$\forall \varepsilon > 0 \qquad \mathbb{P}(\|f_n - f\| > \varepsilon) \to 0 \quad \text{when} \quad n \to \infty.$$

This convergence is in general strictly weaker than a.s. convergence. However, by Corollary 1.24, it is equivalent for sums of *independent* random variables, in particular for the sums considered in Theorem 8.6 below.

The corresponding topology is the natural one on the topological vector space $L_0(B)$ of B-valued Bochner measurable functions. The preceding Lemma is an extrapolation principle: If the L_q-topology on a linear space $\mathcal{F}$ coincides with the L_p-topology for *some* $p < q$, then it also coincides with the topology of convergence in probability (i.e. the L_0-topology).

In particular, we obtain

Theorem 8.6. *(i) Let (α_n) be a scalar sequence. Then $\sum_0^\infty \varepsilon_n\alpha_n$ converges in probability iff $\sum |\alpha_n|^2 < \infty$, and then it converges a.s. and in L_p for all $p < \infty$.*

(ii) Let (x_n) be a sequence in a Banach space B. Then the series $\sum_0^\infty \varepsilon_n x_n$ converges in probability iff it converges in $L_p(B)$ for some $0 < p < \infty$. Then it converges a.s. and in $L_p(B)$ for all $0 < p < \infty$.

Proof. (i) By the Khintchine inequalities and the preceding Lemma, the L_p- and L_0-topologies coincide on the span of $\{\varepsilon_n\}$ for any $0 < p < \infty$. Then the a.s. convergence follows either from Theorem 1.22 or from the martingale convergence theorem since $f_n = \sum_0^n \varepsilon_k\alpha_k$ is a martingale and we may choose $p > 1$.

(ii) Same argument as for (i) but using the Kahane inequalities instead of the Khintchine ones.

□

Applications

(i) Let $0 < r < \infty$. Let (x_n) be a sequence in the Banach (or quasi-Banach) space $B = L_r(T, \mu)$ over a measure space. Then the series $\sum \varepsilon_n x_n$ converges a.s. in B iff

$$\int (\sum |x_n|^2)^{r/2} d\mu < \infty$$

or equivalently iff $(\sum |x_n|^2)^{1/2} \in B$. Indeed, choosing $p = r$ in the last Theorem this is an easy consequence.

(ii) In particular, if $B = \ell_r$, with canonical basis (e_k) and if for each n, we set $x_n = \sum x_n(k)e_k$, then $\sum \varepsilon_n x_n$ converges a.s. in B iff

$$\sum_k (\sum_n |x_n(k)|^2)^{r/2} < \infty.$$

(iii) Let (a_n) be a scalar sequence indexed by $\mathbb{Z}$. Consider the *formal* Fourier series $\sum_{n\in\mathbb{Z}} a_n e^{int}$. Let B be a Banach space of functions over the circle group $\mathbb{T}$, such as for instance the space of continuous functions $C(\mathbb{T})$ or the space $L_r(\mathbb{T}, \mu)$ with respect to the normalized Haar measure μ. Note that in both cases the Fourier transform of an element of B determines the element. By convention, we will write $\sum_{n\in\mathbb{Z}} a_n e^{int} \in B$ if there is an element $f \in B$ with Fourier transform equal to (a_n), i.e. such that $\forall n \in \mathbb{Z}\ \hat{f}(n) = a_n$. Then, $\sum_{n\in\mathbb{Z}} \varepsilon_n a_n e^{int} \in L_r(\mathbb{T}, \mu)$ for almost all choice of signs (ε_n) iff $\sum |a_n|^2 < \infty$. Note that the condition we find does not depend on r, which is surprising at first glance.

(iv) With the same notation, $f = \sum_{n\in\mathbb{Z}} \varepsilon_n a_n e^{int} \in C(\mathbb{T})$ for almost all choice of signs (ε_n) iff a.s. the partial sums of the random Fourier series

$$\sum_{|n|\le N} \varepsilon_n a_n e^{int}$$

converge uniformly over $\mathbb{T}$ when $N \to \infty$; moreover, this holds iff both (unilateral) series $\sum_0^N \varepsilon_n a_n e^{int}$ and $\sum_{-1}^{-N} \varepsilon_n a_n e^{int}$ converge uniformly over $\mathbb{T}$. By the preceding theorem, we then have for any $p < \infty$

$$\mathbb{E} \sup_{t \in \mathbb{T}} \left| \sum_{|n|>N} \varepsilon_n a_n e^{int} \right|^p \to 0.$$

Let $f_+ = \sum_{n \geq 0} \varepsilon_n a_n e^{int}$ and $f_- = \sum_{n<0} \varepsilon_n a_n e^{int}$. Observe that $f = f_+ + f_-$ has the same distribution as $f = f_+ - f_-$, and hence: $f \in B$ a.s. iff both $f_+ \in B$ a.s. and $f_- \in B$ a.s. . Then, the last point (iv) follows from (the Ito-Nisio) Theorem **??**, taking for D the countable collection of all measures μ such that $|\mu|(\mathbb{T}) < 1$ with finitely supported Fourier transform taking values in any fixed dense countable subset of $\mathbb{C}$, say in $\mathbb{Q} + i\mathbb{Q}$.

8.3 Extrapolation. Gundy's decomposition. UMD

The central idea to prove the Burkholder inequalities is usually described as "extrapolation". Schematically, the main point in the scalar case is:

- the L_2-case is obvious by the orthogonality of martingale differences,

-the L_p-case can be deduced from the L_2-one by extrapolation. The basic principles of extrapolation go back to [108]. there are several ways to implement the extrapolation technique. We use the Gundy decomposition because it will be adaptable easily to all the other situations of interest to us in these notes. We will use similar ideas in the vector-valued case. However, although we use classical ideas we will need to be careful about certain details because we are interested in the precise relations between certain constants, such as the UMD constants of a Banach space.

Definition 8.7. Let $1 < p < \infty$. A Banach space B is called UMD$_p$ if there is a constant C such that for any martingale (f_n) converging in $L_p(B)$ we have for any choice of signs $\varepsilon_n = \pm 1$

$$\sup_n \left\| \sum\nolimits_0^n \varepsilon_k df_k \right\|_{L_p(B)} \leq C \sup_n \|f_n\|_{L_p(B)}. \tag{8.16}$$

We will denote the best C in (8.16) by $C_p(B)$. We will say that B is UMD if this holds for some $1 < p < \infty$ (we will see below that it then holds for all $1 < p < \infty$).

Clearly, any Hilbert space is UMD$_2$.

Remark. Equivalently, we may restrict (8.16) to "finite" martingales, i.e. one for which there is an integer n such that $df_k = 0$ for all $k > n$.

By an elementary duality argument one easily checks:

Proposition 8.8. *B is UMD$_p$ iff B^* is UMD$_{p'}$ with $\frac{1}{p} + \frac{1}{p'} = 1$. Moreover, we have*

$$C_{p'}(B^*) = C_p(B). \tag{8.17}$$

Let $x = \{x_n\}$ be a sequence in a Banach space B. We define

$$R(x) = \sup_n \left\| \sum\nolimits_0^n \varepsilon_k x_k \right\|_{L_2(D,\nu;B)}.$$

Let (f_n) be any martingale on $(\Omega, \mathcal{A}, \mathbb{P})$ that is bounded in $L_p(B)$. For any $\omega \in \Omega$, we define

$$R_{df}(\omega) = R(\{df_n(\omega) \mid n \geq 0\})$$

or equivalently

$$R_{df}(\omega) = \sup_n \left\| \sum\nolimits_0^n \varepsilon_k df_k(\omega) \right\|_{L_2(D,\nu;B)}$$

where the L_2-norm is with respect to the variables (ε_n) defined on (D, ν).

The B-valued version of Burkholder's inequalities reads as follows:

Proposition 8.9. *B is UMD$_p$ iff there are positive constants C_1, C_2 such that for any martingale (f_n) converging in $L_p(B)$ we have*

$$C_1^{-1} \|R_{df}\|_p \leq \sup_n \|f_n\|_{L_p(B)} \leq C_2 \|R_{df}\|_p. \tag{8.18}$$

Proof. Fix a choice of signs $\varepsilon_n = \pm 1$. Let $g_n = \sum_0^n \varepsilon_k df_k$. Note the pointwise equality (recall the proof of (8.14))

$$R_{dg} = R_{df}.$$

The latter immediately implies the "if-part". Conversely, assume B UMD$_p$. Then we have (8.16). But actually, applying (8.16) to g in place of f, we also find the converse inequality

$$\sup_n \|f_n\|_{L_p(B)} \leq C \sup_n \left\| \sum\nolimits_0^n \varepsilon_k df_k \right\|_{L_p(B)}.$$

Assume for simplicity that (f_n) is a finite martingale. Then if we elevate to the p-th power and average both (8.16) and its converse over all choices of signs, we obtain (8.18); note that we use Kahane's inequalities (Theorem 8.1) to replace the L_p-norm over the signs by the L_2-norm, i.e. by R_{df}. □

As before, let $(\varphi_n)_{n\geq 0}$ be a sequence of scalar valued r.v.'s, adapted to a filtration $(\mathcal{A}_n)_{n\geq 0}$. Let $(f_n)_{n\geq 0}$ be a B-valued martingale relative to $(\mathcal{A}_n)_{n\geq 0}$. Just as in the scalar case, the sequence defined by

$$\tilde{f}_n = \varphi_0 f_0 + \sum\nolimits_1^n \varphi_{k-1}(f_k - f_{k-1})$$

forms a martingale, called a "martingale transform" of $(f_n)_{n\geq 0}$.

Proposition 8.10. *If B is UMD_p and (φ_n) real valued, then with the preceding notation we have*

$$\sup_{n\geq 0} \|\tilde{f}_n\|_{L_p(B)} \leq C_1 C_2 \sup_{n\geq 0} \|\varphi_n\|_\infty \sup_{n\geq 0} \|f_n\|_{L_p(B)}.$$

If the (φ_n)'s are complex valued, this holds with $2C_1C_2$ instead of C_1C_2.

Proof. By (8.14) we have for a.a. ω

$$R_{d\tilde{f}}(\omega) \le \sup_{n\ge 0} \|\varphi_n\|_\infty R_{df}(\omega),$$

so the announced inequality follows from (8.18). In the complex case, we use (8.15) instead. $\square$

Remark 8.11. Consider the dyadic case, i.e. we take $\Omega = \{-1,1\}^{\mathbb{N}_*}$, with $\varepsilon_k\colon\ \Omega \to \{-1,1\}$ the k-th coordinate for $k = 1,2,\dots$ and we set $\mathcal{A}_0 = \{\phi,\Omega\}$ and $\mathcal{A}_n = \sigma(\varepsilon_1,\dots,\varepsilon_n)$ for all $n \ge 1$. We claim that, in that case, Proposition 8.10 holds (for real multipliers) with C_1C_2 replaced by the UMD$_p$ constant of B. For notational convenience, we will use the Rademacher functions defined in §1.3 i.e. we take $\Omega = [0,1]$ and $\varepsilon_n = r_n$. In that case we have $\mathcal{A}_n = \mathcal{B}_{2^n}$ where $(\mathcal{B}_k)$ is the Haar filtration (see §1.3). Consider then a martingale transform $\tilde{f}_n = \varphi_0 f_0 + \sum_1^n \varphi_{k-1}(f_k - f_{k-1})$ associated to the predictable family (φ_{k-1}) with respect to the filtration $(\mathcal{A}_n)$. Let then $F_k = \mathbb{E}(f_n|\mathcal{B}_k)$ for $k = 1,\cdots,2^n$, in particular so that $F_{2^n} = f_n$. We have then

$$\tilde{f}_n - \tilde{f}_{n-1} = \sum\nolimits_{2^{n-1}<k\le 2^n} \varphi_{n-1}(F_k - F_{k-1}).$$

But now φ_{n-1} is constant on the support of $F_k - F_{k-1}$ for each $2^{n-1} < k \le 2^n$. Indeed, by (1.8), $F_k - F_{k-1}$ and h_k have the same support and the functions $r_1,\cdots,r_{n-1}$ are all constant on that support if $2^{n-1} < k \le 2^n$. Therefore, with respect to the filtration $(\mathcal{B}_k)$, $\tilde{f}_n$ appears as a martingale transform relative to constant multipliers (and not only predictable ones). This shows that, in the dyadic case, Proposition 8.10 holds with C_1C_2 replaced by the UMD$_p$ constant of B.

Then we will prove:

Theorem 8.12. *Consider a Banach space B. Then for any $1 < p,q < \infty, B$ is UMD$_p$ iff it is UMD$_q$ and we have positive constants $\alpha(p,q)$ and $\beta(p,q)$ depending only on p and q such that*

$$\alpha(p,q)C_p(B) \le C_q(B) \le \beta(p,q)C_q(B).$$

To prove Theorem 8.12 we will use Gundy's decomposition of martingales. This is a martingale analogue of the classical Calderón–Zygmund decomposition (see the next section and also [57]).

Theorem 8.13. *Let B be a Banach space. Let $(f_n)_{n\ge 0}$ be a martingale adapted to $(\mathcal{A}_n)_{n\ge 0}$, and converging in $L_1(B)$ to a limit f with $\|f\|_{L_1(B)} \le 1$. Then for any $\lambda > 0$ there is a decomposition*

$$f = a + b + c$$

with $a,b,c \in L_1(B)$ such that:

(i) $\|a\|_{L_1(B)} \leq 2$ *and*

$$\mathbb{P}(\sup_n \|da_n\| \neq 0) \leq 3\lambda^{-1}$$

(ii) $\big\| \sum \|db_n\| \big\|_1 \leq 4$

(iii) $\|c\|_{L_\infty(B)} \leq 2\lambda$ *and* $\|c\|_{L_1(B)} \leq 5$.

Note that (iii) implies for any $1 < p < \infty$

$$\|c\|_{L_p(B)} \leq 5^{1/p}(2\lambda)^{1-1/p}. \tag{8.19}$$

Proof. We follow Gundy's original proof closely. Recall that by convention we denote $\mathbb{E}_n$ instead of $\mathbb{E}^{\mathcal{A}_n}$, so that $f_n = \mathbb{E}_n f$ and $df_n = f_n - f_{n-1}$ $\forall n \geq 1$ and $df_0 = f_0$. Let $r = \inf\{n \mid \|f_n\| > \lambda\}$. Let $v_n = \|df_n\| \cdot 1_{\{r=n\}}$. Then let

$$s = \inf\left\{n \mid \sum_{k=0}^{n} \mathbb{E}_k(v_{k+1}) > \lambda\right\}.$$

Finally let

$$T = r \wedge s.$$

Clearly r, s and T are stopping times. Let $a = f - f_T$ so that $a_n = f_n - f_{T\wedge n}$. We have clearly (since $f_T = \mathbb{E}^{\mathcal{A}_T} f$) $\|a\|_{L_1(B)} \leq 2$.

Moreover, obviously $T = \infty$ implies $da_n = 0$ so

$$\{\sup \|da_n\| \neq 0\} = \bigcup_n \{da_n \neq 0\} \subset \{T < \infty\} = \{s < \infty\} \cup \{r < \infty\}$$

and hence

$$\mathbb{P}\{\sup \|da_n\| \neq 0\} \leq \mathbb{P}(r < \infty) + \mathbb{P}(s < \infty). \tag{8.20}$$

Now by Doob's inequality (see Theorem 1.9)

$$\mathbb{P}(r < \infty) = \mathbb{P}(\sup_n \|f_n\| > \lambda) \leq \lambda^{-1} \tag{8.21}$$

and also

$$\mathbb{P}(s < \infty) = \mathbb{P}\left(\sum_{k=0}^{\infty} \mathbb{E}_k(v_{k+1}) > \lambda\right) \leq \lambda^{-1} \sum\nolimits_0^\infty \mathbb{E}(\mathbb{E}_k(v_{k+1})) = \lambda^{-1} \sum\nolimits_1^\infty \mathbb{E} v_k.$$

But now

$$\mathbb{E} v_k = \mathbb{E}(\|df_k\| 1_{\{r=k\}})$$

and $r = k$ implies $\|f_{k-1}\| \leq \lambda < \|f_k\|$ hence $\|df_k\| \leq \|f_k\| + \|f_{k-1}\| \leq 2\|f_k\|$. This implies

$$\mathbb{E} v_k \leq \mathbb{E}(2\|f_k\| 1_{\{r=k\}}) = 2\mathbb{E}\|\mathbb{E}_k(f 1_{\{r=k\}})\|$$

hence since by Jensen $\|\mathbb{E}_k(f1_{\{r=k\}})\| \le \mathbb{E}_k(\|f\|1_{\{r=k\}})$, we have

$$\mathbb{E}v_k \le 2\mathbb{E}(\|f\|1_{\{r=k\}}) \tag{8.22}$$

and hence

$$\mathbb{P}(s < \infty) \le \lambda^{-1}\sum\nolimits_0^\infty \mathbb{E}v_k \le 2\lambda^{-1}\|f\|_{L_1(B)} \le 2\lambda^{-1}. \tag{8.23}$$

Thus combining (8.20), (8.21) and (8.23) we obtain (i).

Note that $f - a = f_T$ so we must have a priori

$$b_n + c_n = f_{T\wedge n},$$

which will guarantee that

$$\|b + c\|_{L_1(B)} \le 1. \tag{8.24}$$

Also

$$f_{T\wedge n} - f_{T\wedge(n-1)} = df_n 1_{\{n\le T\}} = df_n \cdot 1_{\{n\le r\}} \cdot 1_{\{n\le s\}} = \gamma_n + \delta_n$$

where

$$\begin{aligned} \gamma_n &= df_n \cdot 1_{\{n<r\}}1_{\{n\le s\}} \\ \delta_n &= df_n \cdot 1_{\{n=r\}}1_{\{n\le s\}}. \end{aligned}$$

Obviously since $(f_{T\wedge n})$ is a martingale we have

$$\mathbb{E}_{n-1}(\gamma_n + \delta_n) = 0 \qquad \forall n \ge 1$$

so we can define $db_0 = \delta_0$, $dc_0 = \gamma_0$ and for all $n \ge 1$

$$\begin{aligned} db_n &= \delta_n - \mathbb{E}_{n-1}(\delta_n) \\ dc_n &= \gamma_n + \mathbb{E}_{n-1}(\delta_n). \end{aligned}$$

Since $\mathbb{E}_{n-1}(\delta_n) = -\mathbb{E}_{n-1}(\gamma_n)$ these are indeed martingale differences.

Note that by Jensen

$$\mathbb{E}\sum \|db_n\| \le 2\mathbb{E}\sum \|\delta_n\| \le 2\mathbb{E}\sum |v_n|$$

hence by (8.23) we have (ii).

We now turn to (iii). First note that by (8.24), (ii) and the triangle inequality we have

$$\|c\|_{L_1(B)} \le 5.$$

Finally, $\sum \gamma_n = \sum\limits_{n\le s} df_n 1_{\{n<r\}} = f_{(r-1)\wedge s}$ if $r \ge 1$ and $\sum \gamma_n = 0$ if $r = 0$, so that by definition of r

$$\left\|\sum \gamma_n\right\|_{L_\infty(B)} \le \lambda. \tag{8.25}$$

Moreover, since $\{n \le s\}$ is $(n-1)$-measurable

$$\Big\|\sum_{n\ge 1}\mathbb{E}_{n-1}(\delta_n)\Big\| = \Big\|\sum_{n\le s}\mathbb{E}_{n-1}(df_n 1_{\{n=r\}})\Big\|$$

hence by Jensen

$$\begin{aligned}&\le \sum_{n\le s}\mathbb{E}_{n-1}(\|df_n\|\cdot 1_{\{n=r\}})\\ &= \sum_{k<s}\mathbb{E}_k(\|df_{k+1}\|1_{\{r=k+1\}})\end{aligned}$$

which by definition of s is $\le \lambda$.

Thus we conclude $\Big\|\sum\limits_{n\ge 1}\mathbb{E}_{n-1}(\delta_n)\Big\|_{L_\infty(B)} \le \lambda$ and (iii) follows from (8.25) by the triangle inequality. □

Corollary 8.14. *Assume that B is UMD$_p$ for some $1 < p < \infty$. Then all martingale transforms are of weak-type (1-1). More precisely there is a constant C such that for all martingales $(f_n)_{n\ge 0}$ bounded in $L_1(B)$ and for all choices of signs $\varepsilon_n = \pm 1$ the transformed martingale $\tilde f_n = \sum_0^n \varepsilon_k df_k$ satisfies*

$$\sup_{\lambda>0}\lambda\mathbb{P}(\sup_{n\ge 0}\|\tilde f_n\| > \lambda) \le C\sup_{n\ge 0}\|f_n\|_{L_1(B)}. \tag{8.26}$$

More generally, the same holds when

$$\tilde f_n = \sum\nolimits_0^n \varphi_{k-1}df_k$$

and $(\varphi_n)_{n\ge 0}$ is an adapted sequence of scalar valued variables such that

$$\forall n \ge 0 \qquad \|\varphi_n\|_\infty \le 1,$$

with the usual convention $\varphi_{-1} \equiv 0$.

Proof. By homogeneity, we may assume $\|f\|_{L_1(B)} \le 1$. We have $\tilde f_n = \tilde a_n + \tilde b_n + \tilde c_n$ and

$$\mathbb{P}(\sup\|\tilde f_n\| > 3\lambda) \le \mathbb{P}(\sup\|\tilde a_n\| > \lambda) + \mathbb{P}(\sup\|\tilde b_n\| > \lambda) + \mathbb{P}(\sup\|\tilde c_n\| > \lambda). \tag{8.27}$$

We estimate each term on the right side separately: since $\sup\|\tilde a_n\| > \lambda$ implies $\sup\limits_n\|da_n\| \ne 0$ we have by (i)

$$\mathbb{P}(\sup\|\tilde a_n\| > \lambda) \le 3\lambda^{-1}.$$

By Chebyshev's inequality, since $\sup\|\tilde b_n\| \le \sum\|db_n\|$

$$\mathbb{P}(\sup\|\tilde b_n\| > \lambda) \le \lambda^{-1}\sum\mathbb{E}\|db_n\| \le 4\lambda^{-1}.$$

Finally, by (8.19), UMD_p and Doob's inequality, we have

$$\left\| \sup \|\tilde{c}_n\| \right\|_p \leq p' C_p(B) 5^{1/p} \lambda^{1-1/p}$$

hence by Chebyshev again

$$\mathbb{P}(\sup \|\tilde{c}_n\| > \lambda) \leq (p' C_p(B) 5^{1/p})^p \lambda^{-1}$$

so by (8.27) we obtain (8.26) with the constant $C \leq 3(7 + 5(p' C_p(B))^p)$.

For the more general case of predictable multipliers (φ_n), the same argument works using Proposition 8.10. $\square$

Remark 8.15. Note that the preceding argument also shows

$$\sup_{n \geq 0} \sup_{\lambda > 0} \lambda \mathbb{P}(\|\tilde{f}_n\| > \lambda) \leq C' \sup_{n \geq 0} \|f_n\|_{L_1(B)}$$

with $C' \leq 3(7 + 5C_p(B)^p)$.

Corollary 8.16. *In the scalar case (i.e. $B = \mathbb{R}$ or $\mathbb{C}$) we find* (8.26) *with $C \leq 81$. Moreover, for any martingale $(f_n)_{n \geq 0}$ bounded in L_1 we have*

$$\sup_{\lambda > 0} \lambda \mathbb{P}\left(\left(\sum |df_n|^2 \right)^{1/2} > \lambda \right) \leq 81 \sup_{n \geq 0} \|f_n\|_1.$$

More generally, if B is a Hilbert space we find

$$\sup_{\lambda > 0} \lambda \mathbb{P}\left(\left(\sum \|df_n\|^2 \right)^{1/2} > \lambda \right) \leq 81 \sup \|f_n\|_{L^1(B)}. \tag{8.28}$$

Proof. The first assertion is clear since $C_2(\mathbb{R}) = C_2(\mathbb{C}) = 1$. To prove the second one, one simply observes that

$$\left\| \left(\sum |df_n|^2 \right)^{1/2} \right\|_2 = \|f\|_2$$

so that the same argument applies when we substitute $S(f) = \left(\sum |df_n|^2 \right)^{1/2}$ to $\sup_{n \geq 1} |\tilde{f}_n|$. $\square$

Remark 8.17. Note that (8.28) only requires that B is UMD and of cotype 2.

Theorem 8.18 (Burkholder's inequalities)**.** *For any $1 < p < \infty$, there is a positive constant β_p such that, for any scalar martingale (M_n) in L_p and for any predictable uniformly bounded scalar sequence (φ_n), we have*

$$\sup_n \| \sum_0^n \varphi_k dM_k \|_p \leq \beta_p \sup_n \|M_n\|_p \sup_n \|\varphi_n\|_\infty.$$

Let S be the square function defined by (8.8)*. There are positive constants a_p and b_p such that any scalar martingale (M_n) in L_p satisfies*

$$a_p^{-1} \|S\|_p \leq \sup_n \|M_n\|_p \leq b_p \|S\|_p. \tag{8.29}$$

Proof. By homogeneity we may assume $\sup_n \|\varphi_n\|_\infty \le 1$. Let T_φ be the transformation taking $f \in L_2$ to $\sum \varphi_n df_n$. Clearly, by Parseval, $\|T_\varphi \colon\ L_2 \to L_2\| \le 1$. A fortiori, T_φ is of weak type (2-2). By Corollary 8.14 applied to $B = \mathbb{C}$, T_φ is of weak type (1-1), hence by Marcinkiewicz Theorem 8.51, for any $1 < p < 2$, there is β_p so that $\|T_\varphi \colon\ L_p \to L_p\| \le \beta_p$. By duality, since T_φ is (essentially) self-adjoint we have $\|T_\varphi \colon\ L_{p'} \to L_{p'}\| \le \beta_p$ for any $p' > 2$. This establishes the first assertion. To prove the second one, we first fix a choice of signs $\varepsilon = (\varepsilon_n)$ and we use $\varphi_n = \varepsilon_n$. Let $T_\varepsilon M_n = \sum_0^n \varepsilon_k dM_k$. The first assertion gives us $\|T_\varepsilon M_n\|_p \le \beta_p \|M_n\|_p$ but since $T_\varepsilon(T_\varepsilon M_n) = M_n$ we have by iteration $\|M_n\|_p \le \beta_p \|T_\varepsilon M_n\|_p$. Therefore,

$$(\beta_p)^{-1}\|T_\varepsilon M_n\|_p \le \|M_n\|_p \le \beta_p \|T_\varepsilon M_n\|_p.$$

But if we now integrate with respect to ε and use the Khintchine inequalities (8.7), letting $S_n = (\sum_o^n |dM_k|^2)^{1/2}$ we find

$$A_p(\beta_p)^{-1}\|S_n\|_p \le \|M_n\|_p \le \beta_p B_p \|S_n\|_p,$$

and we conclude by taking the supremum over n.
Note that, for $1 < p < 2$, we can also deduce the square function inequality $\|S\|_p \le b_p \sup_n \|M_n\|_p$ from Corollary 8.16 by the sublinear version of Marcinkiewicz Theorem (see Remark 8.52). □

Proof of Theorem 8.12. Assume B UMD$_p$. Then by Corollary 8.14 and by the Marcinkiewicz interpolation Theorem 8.51, B is UMD$_q$ for all $1 < q < p$. But now, by (8.17) B^* is UMD$_{p'}$, and hence we may repeat the preceding argument for B^*, and obtain that B^* is UMD$_{q'}$ for any $1 < q' < p'$. By (8.17) again, this means that B is UMD$_q$ for all $q > p$, and hence finally for all $1 < q < \infty$. □

We now give the basic examples of UMD spaces.

Corollary 8.19. *Let (S, Σ, m) be an arbitrary measure space and let $1 < p < \infty$. Then the Banach space $B = L_p(S, \Sigma, m)$ is UMD. More generally, if B is any UMD space, then the space $L_p(S, \Sigma, m; B)$ is UMD.*

Proof. By Fubini's theorem, it is easy to see that, if B is UMD$_p$, then the space $L_p(S, \Sigma, m; B)$ is UMD$_p$. The case $B = \mathbb{C}$ corresponds to the first assertion. Since UMD$_p$ does not depend on p, this proves these two assertions. □

We will now give a different approach to Theorem 8.12 based on "extrapolation". This is particularly efficient in the dyadic case: Indeed, the dyadic filtration has the advantage that $\|df_{n+1}(\omega)\|_B$ is actually $\mathcal{A}_n$-measurable for each $n \ge 0$. In other words the lengths of the increments are "predictable".

We start by a rather general version of the extrapolation principle.

Lemma 8.20. *Let $(v_n)_{n\ge 0}$ and $(w_n)_{n\ge 0}$ be adapted sequences of non-negative random variables, converging a.s. to limits denoted by v_∞ and w_∞. Fix $p > 0$.*

Let $C > 0$ be a constant. Assume that for any stopping time $T\colon\ \Omega \to \mathbb{N} \cup \{\infty\}$ we have

$$\|1_{\{T>0\}} v_T\|_p \leq \|1_{\{T>0\}} w_T\|_p. \tag{8.30}$$

Moreover, assume that there is an adapted non-negative sequence $(\psi_n)_{n\geq 0}$ such that

$$\forall n \geq 0 \qquad w_{n+1} - w_n \leq \psi_n.$$

Let $w^* = \sup_n w_n$ and $\psi^* = \sup_n \psi_n$. Then for any $t > 0$

$$\mathbb{P}\{v_\infty > t\} \leq t^{-p}\mathbb{E}(t^p \wedge w^{*p}) + \mathbb{P}\{w^* + \psi^* > t\}, \tag{8.31}$$

and hence for any $0 < q < p$

$$\mathbb{E}v_\infty^q \leq (p/(p-q))\mathbb{E}w^{*q} + \mathbb{E}(w^* + \psi^*)^q. \tag{8.32}$$

Proof. Let $T = \inf\{n \mid w_n + \psi_n > t\}$. Note that on $\{T > 0\}$ we have

$$w_T \leq w_{T-1} + \psi_{T-1} \leq t,$$

and hence

$$1_{\{T>0\}} w_T \leq t \wedge w^*. \tag{8.33}$$

We have obviously

$$\begin{aligned}
\mathbb{P}(v_\infty > t) &\leq \mathbb{P}(v_T > t, T = \infty) + \mathbb{P}(v_\infty > t, T < \infty) \\
&\leq \mathbb{P}(v_T > t, T > 0) + \mathbb{P}(T < \infty) \\
&\leq t^{-p}\mathbb{E}(1_{\{T>0\}} v_T^p) + \mathbb{P}(\sup(w_n + \psi_n) > t)
\end{aligned}$$

and hence by (8.30) and (8.33)

$$\begin{aligned}
&\leq t^{-p}\mathbb{E}(1_{\{T>0\}} w_T^p) + \mathbb{P}(w^* + \psi^* > t) \\
&\leq t^{-p}\mathbb{E}(t^p \wedge w^{*p}) + \mathbb{P}(w^* + \psi^* > t),
\end{aligned}$$

which proves (8.31). Then, using $\mathbb{E}v_\infty^q = \int_0^\infty qt^{q-1}\mathbb{P}\{v_\infty > t\}dt$ and $1 \wedge (w^*/t)^p = 1_{\{w^*>t\}} + (w^*/t)^p 1_{\{w^*\leq t\}}$ we obtain (8.32) by an elementary computation.

$\square$

Remark. Note that actually we only use $\|1_{\{T=\infty\}} v_T\|_p \leq \|1_{\{T>0\}} w_T\|_p$ which is a priori weaker than (8.30)

Second Proof of Theorem 8.12. Let $1 < q < p < \infty$. Assume B UMD$_p$. Consider a finite dyadic martingale (f_n) and a fixed choice of signs (ε_n). We will apply Lemma 8.20. Let $\tilde f_\infty = \sum \varepsilon_n df_n$, $\tilde f_n = \mathbb{E}_n(\tilde f_\infty)$ and let T be a stopping time. We set $v_n = \|\tilde f_n\|_B$ and $w_n = C_p(B)\|f_n\|_B$. By (8.16) applied to the martingale $(1_{\{T>0\}} f_{n\wedge T})$, we have

$$\|1_{\{T>0\}} v_T\|_p = \|1_{\{T>0\}}\tilde f_T\|_{L_p(B)} \le C_p(B)\|1_{\{T>0\}} f_T\|_{L_p(B)} = \|1_{\{T>0\}} w_T\|_p$$

and hence (8.30) holds.

For dyadic martingales, $\|df_{n+1}\|_B$ is n-measurable, so we can take simply $\psi_n = C_p(B)\|df_{n+1}\|_B$. Note that $w^*+\psi^* \le 3f^*$. By Doob's maximal inequality, if $1 < q < p$, (8.32) implies

$$\|\tilde f_\infty\|_{L_q(B)} = \|v_\infty\|_q \le C_p(B)\left(3 + (p/(p-q))^{1/q}\right) q'\|f\|_{L_q(B)}.$$

This shows B is UMD$_q$. By duality (see (8.17)), the preceding argument applied to B^* shows B^* is UMD$_{q'}$ for $1 < q' < p' < \infty$, and hence that B also is UMD$_q$ for any $1 < p < q < \infty$. All this is restricted to the dyadic filtration, but it is known ([202], see section 8.6 below), that it suffices. □

Remark 8.21. Assume B UMD$_p$ again. The preceding argument shows that, for any q with $0 < q < p$, there is a constant $D(q,p)$ such that for any dyadic B-valued martingale we have

$$\|\sup_n \|\tilde f_n\|_B\|_q \le D(q,p)C_p(B)\|\sup_n \|f_n\|_B\|_q.$$

Indeed, this follows easily from Lemma 8.20 setting $v_n = \sup_{k\le n}\|\tilde f_k\|_B$ and $w_n = \sup_{k\le n}\|f_k\|_B$.

8.4 The UMD property for $p = 1$ Burgess Davis decomposition

The following classical decomposition due to Burgess Davis is very useful to control the "jumps" of a martingale when a priori their length is not predictable.

Lemma 8.22. *A general B-valued martingale $(f_n)_{n\ge0}$ with $\mathbb{E}\sup_n\|f_n\| < \infty$ can be decomposed as a sum*

$$f_n = h_n + g_n$$

with $h_0 = 0$, $\sum_1^\infty \mathbb{E}\|dh_n\| \le 6\,\mathbb{E}\sup_n\|f_n\|$ and where almost surely for all $n \ge 1$

$$\|dg_n\| \le 6\ f^*_{n-1},$$

with the notation $f^*_n = \sup_{k\le n}\|f_k\|$. More generally, for any $p \ge 1$ we have

$$\left\|\sum\nolimits_1^\infty \|dh_n\|\right\|_p \le 6p\,\|\sup\|f_n\|\|_p \tag{8.34}$$

and hence

$$\|\sup\|g_n\|\|_p \leq (1+6p)\|\sup\|f_n\|\|_p. \tag{8.35}$$

Proof. We define h and g via their increments by setting $h_0 = 0, g_0 = f_0$ and

$$dh_n = df_n \cdot 1_{\{f_n^* > 2f_{n-1}^*\}} - \mathbb{E}_{n-1}(df_n \cdot 1_{\{f_n^* > 2f_{n-1}^*\}})$$

and

$$dg_n = df_n \cdot 1_{\{f_n^* \leq 2f_{n-1}^*\}} - \mathbb{E}_{n-1}(df_n \cdot 1_{\{f_n^* \leq 2f_{n-1}^*\}}).$$

Note that when $f_n^* > 2f_{n-1}^*$ then $f_n^* \leq 2(f_n^* - f_{n-1}^*)$ hence we must have

$$\|df_n\|1_{\{f_n^* > 2f_{n-1}^*\}} \leq (f_n^* + f_{n-1}^*)1_{\{f_n^* > 2f_{n-1}^*\}} \leq (3/2)f_n^* 1_{\{f_n^* > 2f_{n-1}^*\}} \leq 3(f_n^* - f_{n-1}^*).$$

Hence

$$\sum \|dh_n\|_{L^1(B)} \leq \sum 6\mathbb{E}(f_n^* - f_{n-1}^*) \leq 6\mathbb{E}\sup_n \|f_n\|.$$

On the other hand, we have

$$\|df_n \cdot 1_{\{f_n^* \leq 2f_{n-1}^*\}}\| \leq (f_n^* + f_{n-1}^*) \cdot 1_{\{f_n^* \leq 2f_{n-1}^*\}} \leq 3f_{n-1}^*$$

hence $\|dg_n\| \leq 6f_{n-1}^*$. Finally, (8.34) follows from the dual to Doob's inequality (namely Theorem 1.10), since we have

$$\sum \|dh_n\| \leq \sum 3(f_n^* - f_{n-1}^*) + \sum \mathbb{E}_{n-1}[3(f_n^* - f_{n-1}^*)],$$

and (8.35) follows from the triangle inequality. □

Theorem 8.23. *Let B be a UMD Banach space. Let $C = 54\ C_2(B)$. Then for all filtrations $(\mathcal{A}_n)_{n\geq 0}$ and all choices of signs $\varepsilon_n = \pm 1$ we have for all martingales $(f_n)_{n\geq 0}$*

$$\mathbb{E}\sup_n \left\|\sum_0^n \varepsilon_k df_k\right\| \leq C\mathbb{E}\sup_n \|f_n\|. \tag{8.36}$$

Proof. We will use Lemma 8.22. Let $\tilde{f}_n = \sum_{k\leq n} \varepsilon_k df_k$ and let $\tilde{f}_n^* = \sup_{k\leq n}\|\tilde{f}_k\|$ and $\tilde{f}^* = \sup_n \tilde{f}_n^*$, and similarly for (g_n) and (h_n). By the UMD property and Doob's maximal inequality, we have for any stopping time T (note that $(1_{\{T>0\}}\tilde{g}_{n\wedge T})$ is a martingale)

$$\|\tilde{g}_T^* 1_{\{T>0\}}\|_2 \leq 2\|\tilde{g}_T 1_{\{T>0\}}\|_{L_2(B)} \leq 2\ C_2(B)\|g_T 1_{\{T>0\}}\|_{L_2(B)}.$$

By the triangle inequality we have on one hand

$$\|\tilde{f}^*\|_1 \leq \sum \|dh_n\|_{L^1(B)} + \|\tilde{g}^*\|_1 \leq 6\|f^*\|_1 + \|\tilde{g}^*\|_1.$$

On the other hand, by (8.32), applied with $v_n = \tilde{g}_n^*$, $w_n = 2\ C_2(B)g_n^*$ and $\psi_n = 2\ C_2(B)6f_n^*$, we have

$$\|\tilde{g}^*\|_1 \leq 12\ C_2(B)\|f^*\|_1 + 6\ C_2(B)\|g^*\|_1$$

hence, using $\|g^*\|_1 \leq \|f^*\|_1 + \|h^*\|_1 \leq 7\|f^*\|_1$, we obtain the announced result after some arithmetic. □

Remark 8.24. Conversely, any space B satisfying (8.36) for some C must be UMD. More generally, for $\varepsilon = (\varepsilon_n) \in D$, let us denote $T_\varepsilon(f) = \sum_0^\infty \varepsilon_n df_n$. If we have for any ε

$$\|T_\varepsilon(f)\|_1 \le C\|f\|_{L_\infty(B)},$$

then B is UMD. Actually even if an inequality of the form $\|T_\varepsilon(f)\|_r \le C\|f\|_{L_\infty(B)}$, holds for some $0 < r < 1$ then B is UMD. This follows from the type/cotype theory (see [207, 206]). Let us briefly sketch the argument. Indeed, this inequality implies that B is of type $p > 1$ and of cotype $q < \infty$. The cotype $< \infty$ implies that for some $q < \infty$ and some C_1 we have $\|T_\varepsilon(f)\|_r \le C_1\|f\|_{L_q(B)}$, then by Gundy's decomposition (Theorem 8.13) we find for some $0 < t < 1$ an inequality of the form $\|T_\varepsilon(f)\|_t \le C_2\|f\|_{L_1(B)}$. A fortiori we have $\|T_\varepsilon(f)\|_t \le C_2\|f\|_{L_p(B)}$ for $p > 1$. Therefore using the type > 1, we conclude that, for a suitable $p > 1$, we have $\|T_\varepsilon(f)\|_p \le C_3\|f\|_{L_p(B)}$, and hence B is UMD.

Third Proof of Theorem 8.12. This is merely a variant of the second proof that avoids the restriction to dyadic martingales by making use of the B. Davis decomposition. Let $1 < q < \infty$. Assume B UMD$_p$. Consider a finite martingale (f_n) in $L_q(B)$ and let (g_n) be as in Lemma 8.22. Note that (g_n) is also a finite martingale. We may assume for simplicity $f_0 = g_0 = 0$. Fix a choice of signs $\varepsilon = (\varepsilon_n)$. Let $g_\infty = \sum_1^\infty \varepsilon_k dg_k$, $\tilde g_n = \sum_1^n \varepsilon_k dg_k$ and let $\tilde g^* = \sup_n \|\tilde g_n\|$, $\tilde g_n^* = \sup_{k\le n} \|\tilde g_k\|$ and $g_n^* = \sup_{k\le n} \|g_k\|$. Since B is UMD$_p$, by (8.16) and Doob's inequality, we have

$$\|\tilde g^*\|_p \le p' C_p(B)\|g_\infty\|_{L_p(B)}.$$

Since this also holds for all the stopped martingales $(1_{\{T>0\}} g_{n\wedge T})$ for any stopping time T, we have

$$\|1_{\{T>0\}}\tilde g_T^*\|_p \le p' C_p(B)\|1_{\{T>0\}} g_T\|_{L_p(B)},$$

and a fortiori if we set $v_n = \tilde g_n^*$ we have

$$\|1_{\{T>0\}} v_T\|_p \le p' C_p(B)\|1_{\{T>0\}} g_T\|_{L_p(B)}.$$

We will use Lemma 8.20 with $w_n = p' C_p(B) g_n^*$ and $\psi_n = p' C_p(B) 6 f_n^*$. Therefore by (8.32), we have for any $1 \le q < p$

$$\|\tilde g^*\|_q \le p' C_p(B)[6\|f^*\|_q + (1 + (p/(p-q))^{1/q})\|g^*\|_q]$$

and hence by (8.35)

$$\|\tilde g^*\|_q \le p' C_p(B)(6 + (1 + (p/(p-q))^{1/q})(1+6q))\|f^*\|_q = C_p(B)C(p,q)\|f^*\|_q$$

Finally, since we have trivially

$$\tilde f^* \le \tilde g^* + \sum \|dh_n\|,$$

recalling (8.34) and assuming $1 \le q < p$ we obtain

$$\begin{aligned}\|\tilde f^*\|_q &\le \|\tilde g^*\|_q + \left\|\sum \|dh_n\|\right\|_q \\ &\le (C_p(B)C(p,q) + 6q)\|f^*\|_q.\end{aligned}$$

Now, when $1 < q < p$, Doob's maximal inequality yields

$$\|\tilde f\|_{L_q(B)} \le \|\tilde f^*\|_q \le q'(C_p(B)C(p,q) + 6q)\|f\|_{L_q(B)},$$

and hence

$$C_q(B) \le q' C_p(B)C(p,q) + 6q.$$

This shows that $\mathrm{UMD}_p \Rightarrow \mathrm{UMD}_q$. The converse is proved by duality as in the second proof. $\square$

Definition. A Banach space B is called UMD_1 if there is a constant C such that for any martingale in $L_1(B)$ we have for any choice of signs $\varepsilon_n = \pm 1$

$$\mathbb{E}\sup_n \left\|\sum\nolimits_0^n \varepsilon_k df_k\right\|_B \le C\mathbb{E}\sup_n \|f_n\|_B. \tag{8.37}$$

We will denote the best C in (8.37) by $C_1(B)$.

The preceding (or Theorem 8.23) shows that for any $p \neq 1$ $\mathrm{UMD}_p \Rightarrow \mathrm{UMD}_1$. (Just take $q = 1$ in the preceding "third" proof, and stop the proof before the last step.) The converse, namely $\mathrm{UMD}_1 \Rightarrow \mathrm{UMD}$ is also true by the preceding Remark 8.24.

Here is the analogue of Proposition 8.9 for the case $p = 1$:

Proposition 8.25. *A Banach space B is* UMD_1 *(or equivalently UMD) iff there are constants C_1' and C_2' such that for any martingale (f_n) in $L_1(B)$ we have*

$$(C_1')^{-1}\mathbb{E}R_{df} \le \mathbb{E}\sup_n \|f_n\|_B \le C_2'\mathbb{E}R_{df}.$$

Here we recall that

$$R_{df}(\omega) = \sup_n \left\|\sum\nolimits_0^n \varepsilon_k df_k(\omega)\right\|_{L_2(B)}.$$

Proof. Since $\tilde{\tilde f} = f$, we may apply (8.37) with $\tilde f$ in place of f and we obtain

$$C^{-1}\mathbb{E}\sup\|f_n\| \le \mathbb{E}\sup\left\|\sum\nolimits_0^n \varepsilon_k df_k\right\| \le C\mathbb{E}\sup\|f_n\|.$$

After averaging over the choices of signs $\varepsilon = (\varepsilon_n)$, this becomes

$$C^{-1}\mathbb{E}\sup\|f_n\| \le \mathbb{E}\Phi \le C\mathbb{E}\sup\|f_n\| \tag{8.38}$$

where $\Phi(\omega) = \int \sup_n \|\sum_0^n \varepsilon_k df_k(\omega)\|_B d\mu$. By Kahane's inequality (see Remark 8.2) and by Doob's inequality (or Corollary 1.23 with an extra $\sqrt{2}$ factor), we have

$$R_{df} \le \Phi \le 2K(2,1)R_{df}$$

and hence the Proposition follows from (8.38) $\square$

Corollary 8.26. *In the scalar (or Hilbert space) valued case there is a positive constant β_1' such that for any scalar martingale (M_n) and any predictable sequence (φ_n) with $\|\varphi_n\|_\infty \le 1$ we have*

$$\mathbb{E}\sup_n \left|\sum\nolimits_0^n \varphi_k dM_k\right| \le \beta_1' \mathbb{E}\sup|M_n|.$$

Let S be the square function defined by (8.8). *There are constants a_1', b_1' such that for any martingale (M_n) in L_1 we have*

$$(a_1')^{-1}\|S\|_1 \le \mathbb{E}\sup|M_n| \le b_1'\|S\|_1. \tag{8.39}$$

8.5 Examples

Consider again the martingale $M_n = \prod_1^n(1+\varepsilon_k)$ on $D = \{-1,1\}^{\mathbb{N}}$ with respect to the filtration $\mathcal{A}_n = \sigma(\varepsilon_1,\dots,\varepsilon_n)$. We set $M_0 = 1$ and let $\mathcal{A}_0$ be the trivial σ-algebra. Note that $M_k = 2^k 1_{\{\varepsilon_1=\dots=\varepsilon_k=1\}}$ and $dM_k = \varepsilon_k M_{k-1}$ for all $k \ge 1$. Let $\Omega_0 = \{\varepsilon_1 = -1\}$, $\Omega_k = \{\varepsilon_1 = \dots = \varepsilon_k = 1, \varepsilon_{k+1} = -1\}$ for all $0 < k < n$ and finally

$$\Omega_n = \{\varepsilon_1 = \varepsilon_1 = \dots = \varepsilon_n = 1\}.$$

Note that $\Omega_0, \Omega_1, \dots, \Omega_n$ form a partition of our probability space D. We have for any $n \ge 1$

$$\sup_{k\le n}|M_k| = 1_{\Omega_0} + \sum_{0<k<n} 1_{\Omega_k}2^k + 1_{\Omega_n}2^n$$

and hence

$$\mathbb{E}\sup_{k\le n}|M_k| = \sum_{0<k<n} 2^{-k-1}2^k + 1 = (n+1)/2.$$

Let

$$\begin{aligned} S_n &= (|M_0|^2 + |dM_1|^2 + \cdots + |dM_n|^2)^{1/2} \\ &= (1 + |M_0|^2 + \cdots + |M_{n-1}|^2)^{1/2}. \end{aligned}$$

We have

$$S_n = 1_{\Omega_0}2^{1/2} + \sum_{0<k<n} 1_{\Omega_k}(1+1+2^2+\cdots+2^{2k})^{1/2} + 1_{\Omega_n}(1+1+2^2+\cdots+2^{2n-2})^{1/2}$$

and hence

$$\mathbb{E}S_n = 2^{-1/2} + \sum_{0<k<n} 2^{-k-1}\left(1 + \frac{2^{2k+2}-1}{3}\right)^{1/2} + 2^{-n}\left(1+\frac{2^{2n}-1}{3}\right)^{1/2},$$

which shows that there is $\alpha > 0$ independent of n such that

$$n/\alpha \le \mathbb{E}S_n \le \alpha n.$$

As a consequence, we may infer that

$$nA_1/\alpha \leq \sup_{\xi_k=\pm1} \left\|\sum\nolimits_0^n \xi_k dM_k\right\|_1. \tag{8.40}$$

Indeed, by Khintchine's inequality we have for any ω

$$A_1 S_n(\omega) \leq \int \left|\sum\nolimits_0^n \xi_k dM_k(\omega)\right| d\mu(\xi)$$

and hence after integration in ω

$$A_1 \mathbb{E} S_n \leq \int \left\|\sum\nolimits_0^n \xi_k dM_k\right\|_1 d\mu(\xi)$$

which obviously implies (8.40).

In particular this shows that the inequality $\|S\|_p \leq a_p \sup_n \|M_n\|_p$ (see Theorem 8.18), as well as Doob's maximal inequality do not remain valid for $p = 1$. We will now show that the spaces ℓ_1, c_0, L_1, $C_{[0,1]}$ and L_∞ all fail UMD.

The proof is based on the following

Proposition 8.27. *For each $1 < p < \infty$, there is $\delta > 0$ such that the* UMD_p-*constant of ℓ_1^N satisfies*

$$\forall N \geq 1 \qquad C_p(\ell_1^N) \geq \delta \mathrm{Log}(N).$$

Proof. It suffices to show this for $N = 2^n$. Consider $B = L_1(D, \mu)$ and let (f_n) be the B-valued martingale defined by $f_n(\omega) = \prod_1^n (1 + \varepsilon_k(\omega)\varepsilon_k)$. Note that by the translation invariance of μ (indeed μ is the Haar measure on $\{-1,1\}^{\mathbb{N}}$) we have for any ω

$$\|f_n(\omega)\|_B = \|M_n\|_1 \tag{8.41}$$

where (M_n) is the scalar valued martingale in the previous paragraph. Similarly, for any choice of signs $\xi = (\xi_n)_n$ we have for any ω

$$\left\|\sum\nolimits_0^n \xi_k df_k(\omega)\right\|_B = \left\|\sum\nolimits_0^n \xi_k dM_k\right\|_1. \tag{8.42}$$

Now observing that $(f_1, \ldots, f_n)$ actually takes values in a subspace of B that is isometric to $\ell_1^{2^n}$, we have

$$\left\|\sum\nolimits_0^n \xi_k df_k\right\|_{L_p(B)} \leq C_p(\ell_1^{2^n}) \left\|\sum\nolimits_0^n df_k\right\|_{L_p(B)}$$

and hence by (8.41) and (8.42)

$$\left\|\sum\nolimits_0^n \xi_k dM_k\right\|_1 \leq C_p(\ell_1^{2^n})\|M_n\|_1 \leq C_p(\ell_1^{2^n}).$$

which implies by (8.40)

$$nA_{1/\alpha} \leq C_p(\ell_1^{2^n}).$$

□

Remark. Note that by (8.42) we have for any fixed $\omega \in \Omega$

$$R_{df}(\omega) = \left(\int \left\| \sum_0^n \xi_k dM_k \right\|_1^2 d\mu(\xi) \right)^{1/2},$$

and hence by Jensen's inequality and by (8.7)

$$A_1 \mathbb{E} S_n \le R_{df}(\omega) \le \mathbb{E} S_n.$$

Thus, we obtain for any ω

$$(A_1/\alpha)n \le R_{df}(\omega) \le \alpha n.$$

We say that a Banach space B contains ℓ_1^n's uniformly (equivalently in the terminology of §3.1), ℓ_1 is finitely representable in B) if for any $\varepsilon > 0$ there is a subspace $E \subset B$ that is $(1+\varepsilon)$-isomorphic to ℓ_1^n. We then have

Corollary 8.28. *The spaces ℓ_1, L_1 (and also $c_0, \ell_\infty, C[0,1]$ and $L_\infty[0,1]$) all fail the UMD property. More generally, any space B that contains ℓ_1^n's uniformly must fail UMD.*

Proof. The last assertion is an obvious consequence of the Proposition. Then the other assertions all follow since each of the spaces listed contains ℓ_1^n's uniformly. Indeed, note that $\ell_1^n \subset \ell_\infty^{2^n}$ isometrically. Alternatively, one can use duality (we have $C_p(\ell_1^N) = C_{p'}(\ell_\infty^N)$) to deduce from the Proposition that $c_0, \ell_\infty, C[0,1]$ and $L_\infty([0,1])$ fail UMD$_{p'}$. □

8.6 Dyadic UMD implies UMD

We wish to show that we may restrict ourselves to the dyadic filtration in the definition of UMD spaces. For that purpose, the following Lemmas will be convenient.

Lemma 8.29. *Let $(f_n)_{n\ge 0}$ be a martingale in $L_p(\Omega, \mathcal{A}; B)$ $(1 \le p < \infty)$. Let $\varepsilon > 0$. Then there is a martingale $(f'_n)_{n\ge 0}$ formed of step functions such that*

$$\forall n \ge 0 \qquad \|f_n - f'_n\|_{L_p(B)} < \varepsilon.$$

Proof. Fix $\delta_n > 0$ with $\sum \delta_n < \delta$. Let $\mathcal{A}_n = \sigma(f_0, \ldots, f_n)$. For each $n \ge 0$, let F_n be an $\mathcal{A}_n$-measurable step function such that

$$\|df_n - F_n\|_{L_p(B)} < \delta_n.$$

Let $\mathcal{B}_n = \sigma(F_0, \ldots, F_n)$. Note $\mathcal{B}_n \subset \mathcal{A}_n$ and $\mathcal{B}_n$ is finite since the F_k's are step functions. Let $f'_n = F_0 + \sum_1^n F_k - \mathbb{E}^{\mathcal{B}_{k-1}} F_k$. Note that for all $n \ge 1$

$$\|df_n - df'_n\|_{L_p(B)} \le \|df_n - F_n\|_{L_p(B)} + \|E^{\mathcal{B}_{n-1}}(F_n - df_n)\|_{L_p(B)} < 2\delta_n$$

and hence
$$\|f_n - f'_n\|_{L_p(B)} \le 2\sum\nolimits_0^n \delta_k < 2\delta.$$
Since $\mathcal{B}_n$ is finite, f'_n is a step function and finally $(f'_n)_{n\ge 0}$ is clearly a martingale, so the result follows with $\delta = \varepsilon/2$. □

Recall that the dyadic filtration on $([0,1[, dt)$ is defined by $\mathcal{A}_0 = (\emptyset, [0,1])$ and $\mathcal{A}_m$ is generated by the partition into the 2^m intervals
$$[(j-1)2^{-m}, j2^{-m}[\qquad j = 1, 2, \ldots, 2^m.$$
Given $(\Omega, \mathcal{A}, \mathbb{P})$ and $C \in \mathcal{A}$ with $\mathbb{P}(C) > 0$, We will denote by $\mathbb{P}_C$ the conditional probability on C i.e.
$$\forall A \in \mathcal{A} \quad \mathbb{P}_C(A) = \mathbb{P}(C)^{-1}\mathbb{P}(A \cap C).$$
Note that if I_j is a finite (measurable) partition of C,
$$\mathbb{P}_C(A) = \sum_j \mathbb{P}_C(I_j)\mathbb{P}_{I_j}(A).$$
For simplicity, we use the notation $\mathbb{P}(F = x)$ instead of $\mathbb{P}(\{F = x\})$.

Lemma 8.30. *Let f be a B-valued variable on $(\Omega, \mathcal{A}, \mathbb{P})$ taking values in a finite set $V \subset B$. We assume $\mathbb{P}(f = x) > 0$ for all x in V. Then for any $\varepsilon > 0$ we can find positive integers m and $\{k_x \mid x \in V\}$ with $\sum_x 2^{-m}k_x = 1$ such that*
$$\forall x \in V \qquad |\mathbb{P}(f = x) - 2^{-m}k_x| < \varepsilon.$$
Moreover, if this holds, then, denoting by $\mathbb{P}'$ the Lebesgue (probability) measure on $[0,1]$

(i) *There is an $\mathcal{A}_m$-measurable variable φ on $([0,1], \mathbb{P}')$, taking the same values as f, such that, for all x in V, $\mathbb{P}'(\varphi = x) = 2^{-m}k_x$.*

(ii) *Let $m' \ge 1$ be any integer and let C be any set in $\mathcal{A}_{m'}$. Then there is a $\mathcal{A}_{m'+m}$-measurable variable φ on C such that for any non-negligible $I \subset C$ with $I \in \mathcal{A}_{m'}$*
$$\mathbb{P}'_I(\varphi = x) = 2^{-m}k_x$$
and hence
$$|\mathbb{P}(f = x) - \mathbb{P}'_I(\varphi = x)| < \varepsilon.$$

Proof. The first assertion as well as (i) are essentially obvious. For (ii), assume that $|C| = 2^{-m'}K$, so that C is a disjoint union of (atomic) intervals of $\mathcal{A}_{m'}$ say $\{I_j \mid j \le K\}$ with $|I_j| = 2^{-m'}$. Then on each I_j, by (i) transported on I_j, we can find an $\mathcal{A}_{m'+m}$-measurable variable φ_j such that
$$\mathbb{P}'_{I_j}(\varphi_j = x) = 2^{-m}k_x$$

so it suffices to set

$$\varphi = \sum\nolimits_j 1_{I_j} \varphi_j$$

to obtain

$$\mathbb{P}'_C(\varphi = x) = \sum\nolimits_j \mathbb{P}'_C(I_j)\mathbb{P}'_{I_j}(\varphi_j = x) = 2^{-m}k_x,$$

and the same with C replaced by any $I \subset C$ ($I \in \mathcal{A}_{m'}$, $|I| > 0$). $\square$

Lemma 8.31. *Let $(f_n)_{n\geq 0}$ be a B-valued martingale formed of step functions, so that for each $N \geq 0$, $(f_0, \ldots, f_N)$ takes values in a finite subset $V_N \subset B^{N+1}$ and f_0 is constant. Let $\delta_n \geq 0$ and $\delta > 0$ be such that $\sum \delta_n < \delta$. Then there are integers*

$$0 = m(0) \leq m(1) \leq \cdots \leq m(N) \leq \cdots$$

and a sequence $(f'_n)_{n\geq 0}$ on $[0,1]$ adapted to $(\mathcal{A}_{m(n)})_{n\geq 0}$ such that for each N, $F'_N = (f'_0, \ldots, f'_N)$ takes the same values V_N as $F_N = (f_0, \ldots, f_N)$ and such that

$$\forall x \in V_N \qquad |\mathbb{P}((f_0, \ldots, f_N) = x) - \mathbb{P}'((f'_0, \ldots, f'_N) = x)| < \delta_1 + \cdots + \delta_N$$

and

$$\forall n \geq 1 \qquad \|f'_{n-1} - \mathbb{E}^{\mathcal{A}_{m(n-1)}} f'_n\|_{L_\infty(B)} < \delta_n.$$

Proof. We prove this by induction on N. The case $N = 0$ is obvious (since f_0 is constant). Assume this proved up to N for a given $\delta > 0$ and let us produce $m(N+1)$ and f'_{N+1}. Fix a value x in V_N. Consider $A(x) = \{(f_0, \ldots, f_N) = x\}$. Then f_{N+1} is a step function on $(A(x), \mathbb{P}_{A(x)})$. Applying Lemma 8.30 to it, we find an integer m as in Lemma 8.30. Since we have only finitely many x's to consider, we choose m large enough so it is suitable for all x's simultaneously. Consider $A'(x) = \{(f'_0, \ldots, f'_N) = x\} \subset [0,1]$. By Lemma 8.29 (ii) applied to $f_{N+1|A(x)}$ with $(\Omega, P) = (A(x), \mathbb{P}_{A(x)})$ and $m' = m(N)$, we can find an $\mathcal{A}_{m(N)+m}$-measurable variable φ_x on $A'(x)$ such that

$$\forall y \in B \qquad |\mathbb{P}'_{A'(x)}(\varphi_x = y) - \mathbb{P}_{A(x)}(f_{N+1} = y)| < \varepsilon.$$

We then define f'_{N+1} on $A'(x)$ by

$$f'_{n+1|A'(x)} = \varphi_x.$$

Then

$$\mathbb{P}((F_N, f_{N+1}) = (x,y)) = \mathbb{P}(A(x))\mathbb{P}_{A(x)}(f_{N+1} = y)$$

and similarly for (F'_N, f'_{N+1}). Recall

$$|\mathbb{P}(A(x)) - \mathbb{P}'(A'(x))| \leq \delta_1 + \cdots + \delta_N$$
$$|\mathbb{P}_{A(x)}(f_{N+1} = y) - \mathbb{P}'_{A'(x)}(f'_{N+1} = y)| < \varepsilon$$

hence

$$|\mathbb{P}((F_N, f_{N+1}) = (x,y)) - \mathbb{P}'((F', f'_{N+1}) = (x,y))| < \delta_1 + \cdots + \delta_N + \varepsilon.$$

Finally, the martingale condition for (F_N, f_{N+1}) means that if $x = (x_0, \cdots, x_N)$

$$x_N = \sum\nolimits_y \mathbb{P}_{A(x)}(f_{N+1} = y)y,$$

so that

$$\Big\| x_N - \sum\nolimits_y \mathbb{P}'_{A'(x)}(f'_{N+1} = y)y \Big\| < K\varepsilon$$

where $K = \sup \sum \|y\|$ with the sum running over all the (finitely many) values of f_{N+1} on $A(x)$ (i.e. those y's such that $\mathbb{P}_{A(x)}(f_{N+1} = y) > 0$) and the sup is over all x's. Since (see Lemma 8.30) we actually can obtain the same with $A(x)$ replaced by any $\mathcal{A}_{m(N)}$-atom $I \subset A(x)$, this implies

$$\|f'_N - \mathbb{E}^{\mathcal{A}_{m(N)}} f'_{N+1}\|_\infty < K\varepsilon.$$

Thus it suffices to choose $\varepsilon > 0$ small enough to obtain the $(N+1)$-th step of the induction. □

Theorem 8.32. *Let $1 < p < \infty$. To compute the UMD_p constant of a Banach space B, we may restrict ourselves to martingale differences relative to the dyadic filtration on [0,1] i.e. their unconditionality constant dominates that of any martingale difference sequence.*

Proof. Assume, we know that

$$\Big\| \sum \varepsilon_n dg_n \Big\|_{L^p(B)} \le C \Big\| \sum dg_n \Big\|_{L^p(B)} \tag{8.43}$$

for any $\varepsilon_n = \pm 1$ and any finite dyadic martingale $(g_0, \ldots, g_N)$ relative to $\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_N$. Then, by an obvious blocking argument, the same still holds for martingales relative to a subsequence $\mathcal{A}_0, \mathcal{A}_{m(1)}, \mathcal{A}_{m(2)}, \ldots$ of the dyadic filtration.

Now fix $\varepsilon > 0$. Consider an arbitrary martingale $(f_0, \ldots, f_N)$ in $L_p(B)$. We claim that it satisfies (8.43). By Lemma 8.29, we may assume that $F = (f_0, \ldots, f_N)$ are step functions. Let then $F' = (f'_0, \ldots, f'_N)$ be as in Lemma 8.31. Note that we have

$$\mathbb{E}\|f_N\|^p - \mathbb{E}\|f'_N\|^p = \sum \|x\|^p (\mathbb{P}(f_N = x) - \mathbb{P}(f'_N = x))$$

where the sum runs over the range R_N of f_N, hence

$$|\mathbb{E}\|f_N\|^p - \mathbb{E}\|f'_N\|^p| \le C_1 \varepsilon \tag{8.44}$$

with $C_1 = \sum_{x \in R_N} \|x\|^p$.

For any $x \in B^{N+1}$, let $T(x) = (y_0, \ldots, y_N)$ where $y_0 = x_0$ and $y_n - y_{n-1} = \varepsilon_n(x_n - x_{n-1})$. Note that T is its own inverse. Then

$$\mathbb{P}(T(F) = x) = \mathbb{P}(F = T^{-1}(x))$$

hence

$$|\mathbb{P}(T(F) = x) - \mathbb{P}(T(F') = x)| < \varepsilon$$

which implies by the same reasoning as above for (8.44) that

$$\left|\mathbb{E}\Big\|\sum \varepsilon_n df_n\Big\|^p - \mathbb{E}\Big\|\sum \varepsilon_n df'_n\Big\|^p\right| < C_2\varepsilon. \tag{8.45}$$

Thus we are essentially reduced to proving (8.43) for (f'_n). But since (f'_n) satisfies the quasi-martingale estimate:

$$\sum_1^N \|f'_{n-1} - \mathbb{E}^{\mathcal{A}_{m(n-1)}} f'_n\|_\infty \le N\varepsilon,$$

we know by Remark 2.18 that there is a bona fide dyadic martingale g_n such that for $k = 1, \ldots, N$

$$\|f'_k - g_{m(k)}\|_{L_p(B)} < C_3\varepsilon.$$

Therefore, we deduce from (8.43)

$$\|\sum_0^N \varepsilon_n df'_n\|_{L_p(B)} \le C\|f'_N\|_{L_p(B)} + C_4\varepsilon,$$

and from (8.44) and (8.45) we get finally

$$\|\sum_0^N \varepsilon_n df_n\|_{L_p(B)} \le C\|f_N\|_{L_p(B)} + C_5\varepsilon.$$

Letting $\varepsilon \to 0$ proves our claim. □

8.7 The Burkholder–Rosenthal inequality

We now turn to what we call the Burkholder–Rosenthal inequality, because Burkholder apparently was inspired by Rosenthal's discovery of this inequality for sums of independent random variables.

Let $(f_n)_{n\ge 0}$ be a scalar (or Hilbert space valued) martingale in L_2. We will denote by $\sigma(f)$ the "conditioned square function", namely

$$\sigma(f) = \left(\|f_0\|^2 + \sum\nolimits_1^\infty \mathbb{E}_{n-1}\|df_n\|^2\right)^{1/2}. \tag{8.46}$$

We will also denote

$$d^*(f) = \sup_{n\ge 0} |df_n|.$$

We have then

Theorem 8.33 (Burkholder–Rosenthal inequality)**.** *For any $2 \le p < \infty$, there are positive constants α'_p, β'_p such that any scalar or Hilbert space valued martingale $(f_n)_{n\ge 0}$ in L_p satisfies*

$$\alpha'_p(\|\sigma(f)\|_p + \|d^*(f)\|_p) \le \sup_{n\ge 0}\|f_n\|_p \le \beta'_p[\|\sigma(f)\|_p + \|d^*(f)\|_p]$$

Proof. We will prove this in the scalar case only. The Hilbert space case is identical. For short, we will write σ and d^* instead of $\sigma(f)$ and $d^*(f)$. By convention, we set $\mathbb{E}_{-1}|d_0|^2 = |d_0|^2$. Recall that we have

$$\alpha_p\|S\|_p \le \sup\|f_n\|_p \le \beta_p\|S\|_p. \tag{8.47}$$

Since $p/2 \ge 1$, by Theorem 1.10 (the dual to Doob's inequality) we have on one hand

$$\left\|\sum \mathbb{E}_{n-1}|d_n|^2\right\|_{p/2} \le (p/2)\left\|\sum |d_n|^2\right\|_{p/2}$$

therefore

$$\|\sigma\|_p \le (p/2)^{1/2}\|S\|_p.$$

On the other hand, by Doob we have

$$\|f^*\|_p \le p' \sup\|f_n\|_p;$$

since $d^* \le 2f^*$, this last inequality implies

$$\|d^*\| \le 2p' \sup\|f_n\|_p.$$

Therefore we obtain

$$\|\sigma\|_p + \|d^*\|_p \le ((p/2)^{1/2}\alpha_p^{-1} + 2p')\sup\|f_n\|_p.$$

For the other side, we will estimate

$$S^2 - \sigma^2 = \sum\nolimits_1^\infty |d_n|^2 - \mathbb{E}_{n-1}|d_n|^2.$$

Since $d_n' = |d_n|^2 - \mathbb{E}_{n-1}|d_n|^2$ are martingale differences, we have

$$\|S^2 - \sigma^2\|_{p/2} \le \beta_{p/2}\left\|\left(\sum |d_n'|^2\right)^{1/2}\right\|_{p/2} \le \beta_{p/2}(\mathrm{I} + \mathrm{II}),$$

where

$$\mathrm{I} = \left\|\left(\sum |d_n|^4\right)^{1/2}\right\|_{p/2}$$

and

$$\mathrm{II} = \left\|\left(\sum (\mathbb{E}_{n-1}|d_n|^2)^2\right)^{1/2}\right\|_{p/2} \le \left\|\sum \mathbb{E}_{n-1}|d_n|^2\right\|_{p/2} = \|\sigma\|_p^2.$$

But now

$$\left(\sum |d_n|^4\right)^{1/4} \le (Sd^*)^{1/2}$$

hence

$$\left\|\left(\sum |d_n|^4\right)^{1/4}\right\|_p \leq (\|S\|_p\|d^*\|_p)^{1/2}$$

so that

$$\mathrm{I} = \left\|\left(\sum |d_n|^4\right)^{1/4}\right\|_p^2 \leq \|S\|_p\|d^*\|_p,$$

hence by the arithmetic/geometric mean inequality for any $t > 0$

$$\sqrt{\mathrm{I}} \leq 2^{-1}(t\|S\|_p + t^{-1}\|d^*\|_p).$$

Recapitulating, this gives us since $\sqrt{\mathrm{I}+\mathrm{II}} \leq \sqrt{\mathrm{I}} + \sqrt{\mathrm{II}}$

$$\||S^2-\sigma^2|^{1/2}\|_p \leq (\beta_{p/2})^{1/2}(\sqrt{\mathrm{I}}+\sqrt{\mathrm{II}}) \leq (\beta_{p/2})^{1/2}(2^{-1}t\|S\|_p+t^{-1}\|d^*\|_p+\|\sigma\|_p).$$

But now

$$S = \sqrt{S^2} \leq |S^2 - \sigma^2|^{1/2} + \sigma$$

therefore

$$\|S\|_p \leq (\beta_{p/2})^{1/2}(2^{-1}t\|S\|_p + 2^{-1}t^{-1}\|d^*\|_p + \|\sigma\|_p) + \|\sigma\|_p.$$

Finally, if we choose t so that $(\beta_{p/2})^{1/2}t = 1$ we find

$$\begin{aligned} & \|S\|_p \leq 2^{-1}\|S\|_p + 2^{-1}\beta_{p/2}\|d^*\|_p + (\beta_{p/2})^{1/2}\|\sigma\|_p + \|\sigma\|_p \\ \Rightarrow \quad & \|S\|_p \leq \beta_{p/2}\|d^*\|_p + 2(\beta_{p/2})^{1/2}\|\sigma\|_p + 2\|\sigma\|_p, \end{aligned}$$

so that we obtain the desired inequality with $\beta'_p = \max\{\beta_{p/2}, 2(\beta_{p/2})^{1/2}+2\}$. □

Lemma 8.34. *Let $2 \leq p \leq \infty$. Any scalar (or Hilbert space) valued martingale $(f_n)_{n\geq 0}$ satisfies*

$$\left(\sum_0^\infty \|df_n\|_p^p\right)^{1/p} \leq 2^{1/p'}\|f\|_p$$

Proof. Consider f in L_∞. Let $f_n = \mathbb{E}_n f$. We have trivially both

$$\left(\sum\nolimits_0^\infty \|df_n\|_2^2\right)^{1/2} \leq \|f\|_2$$

and

$$\sup_n \|df_n\|_\infty \leq 2\|f\|_\infty.$$

Therefore the inequality follows by complex interpolation. □

Let us denote for (f_n) as in (8.46)

$$\sigma_p(f) = \left(\|f_0\|^p + \sum\nolimits_1^\infty \|df_n\|^p\right)^{1/p}.$$

Note that

$$\|\sigma_p(f)\|_p = \left(\sum\nolimits_0^\infty \|df_n\|_p^p\right)^{1/p}.$$

Then the following variant of the Burkholder–Rosenthal inequality is particularly useful:

Theorem 8.35. *Let $2 \le p < \infty$. Let $\alpha''_p = 2^{-1}\min(\alpha'_p, 2^{-1/p'})$. Any scalar (or Hilbert space) valued martingale $(f_n)_{n\ge 0}$ satisfies*

$$\alpha''_p(\|\sigma(f)\|_p + \|\sigma_p(f)\|_p) \le \sup \|f_n\|_p \le \beta_p(\|\sigma(f)\|_p + \|\sigma_p(f)\|_p). \tag{8.48}$$

Proof. The first inequality follows from Lemma 8.34. Moreover, we have trivially $d^*(f) \le \sigma_p(f)$ and hence

$$\|d^*(f)\|_p \le \|\sigma_p(f)\|_p.$$

Thus we obtain the second inequality. □

Corollary 8.36. *Let (Y_n) be independent random variables in L_p, $2 < p < \infty$, with mean zero, i.e. $\mathbb{E}Y_n = 0$ for all n. Then the series $\sum Y_n$ converges in L_p iff both $\sum \|Y_n\|_2^2 < \infty$ and $\sum \|Y_n\|_p^p < \infty$. Moreover, we have*

$$\alpha''_p[(\sum \|Y_n\|_2^2)^{1/2} + (\sum \|Y_n\|_p^p)^{1/p}] \le \|\sum Y_n\|_p \le \beta_p[(\sum \|Y_n\|_2^2)^{1/2} + (\sum \|Y_n\|_p^p)^{1/p}].$$

Proof. Let $\mathcal{A}_n = \sigma(Y_0, Y_1, \cdots, Y_n)$. Clearly, since the (Y_n)'s are independent, we have $\mathbb{E}^{\mathcal{A}_{n-1}}|Y_n|^2 = \mathbb{E}|Y_n|^2$ hence if $f = \sum Y_n$ (i.e. $df_n = Y_n$), we have $\sigma(f) = (\sum \|Y_n\|_2^2)^{1/2}$ and $\|\sigma_p(f)\|_p = (\sum \|Y_n\|_p^p)^{1/p}$. Thus the result follows from (8.48). □

Corollary 8.37. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let $(\mathcal{C}_n)$ be a sequence of independent σ-subalgebras of $\mathcal{A}$. Let $\Sigma_p \subset L_p(\Omega, \mathcal{A}, \mathbb{P})$ be the closure of the linear space of all the finite sums $\sum Y_n$ with $Y_n \in L_p$, $\mathcal{C}_n$-measurable and with $\mathbb{E}Y_n = 0$ for all n. Then the orthogonal projection $Q\colon\ L_2 \to \Sigma_2$, defined by*

$$\forall f \in L_2 \qquad Q(f) = \sum (\mathbb{E}^{\mathcal{C}_n}(f) - \mathbb{E}(f))$$

is bounded on L_p for all $1 < p < \infty$.

Proof. By duality, it suffices to show this for $2 < p < \infty$. Let $f \in L_p$ and let $\mathcal{A}_n = \sigma(\mathcal{C}_0, \mathcal{C}_1, \cdots, \mathcal{C}_n)$. As usual we set $df_n = \mathbb{E}_n f - \mathbb{E}_{n-1} f$. We may as well assume $\mathcal{A} = \mathcal{A}_\infty$. Then Σ_p clearly coincides with the set of all f in L_p such that df_n is $\mathcal{C}_n$-measurable for all n, and we have $Qf = \sum \mathbb{E}^{\mathcal{C}_n} df_n$ for all f in L_2. Assuming $p > 2$, we have

$$(\sum \|\mathbb{E}^{\mathcal{C}_n} df_n\|_2^2)^{1/2} \le (\sum \|df_n\|_2^2)^{1/2} = \|f\|_2 \le \|f\|_p,$$

and by interpolation between the cases $p = \infty$ and $p = 2$ (as in Lemma 8.34 above) we have

$$(\sum \|\mathbb{E}^{\mathcal{C}_n} df_n\|_p^p)^{1/p} \le 2^{1/p'}\|f\|_p. \tag{8.49}$$

Therefore, by Corollary 8.36 we find

$$\|Qf\|_p \le \beta_p[(\sum \|\mathbb{E}^{\mathcal{C}_n} df_n\|_2^2)^{1/2} + (\sum \|\mathbb{E}^{\mathcal{C}_n} df_n\|_p^p)^{1/p}] \le \beta_p(1+2^{1/p'})\|f\|_p,$$

which means $\|Q\colon\ L_p \to L_p\| \le \beta_p(1+2^{1/p'})$. □

Corollary 8.38. *Let $p \ge 2$. Let (Y_n) be a sequence of independent mean zero random variables in $L_p(\Omega, \mathcal{A}, \mathbb{P})$ with $\|Y_n\|_p = 1$. Let $w_n = \|Y_n\|_2$ and $w = (w_n)$. Let $x = (x_n)$ be a scalar sequence. Then the series $\sum x_n Y_n$ converges in L_p iff both $\sum w_n^2|x_n|^2 < \infty$ and $\sum |x_n|^p < \infty$. Let $X_{p,w}$ be the space of all such sequences with norm $\|x\|_{p,w} = (\sum w_n^2|x_n|^2)^{1/2} + (\sum |x_n|^p)^{1/p}$. We have then*

$$\alpha_p''\|x\|_{p,w} \le \|\sum x_n Y_n\|_p \le \beta_p\|x\|_{p,w}.$$

Therefore, as a Banach space, the span in L_p of (Y_n) depends only on $w = (w_n)$.

Proof. This is immediate from (8.48) (see Corollary 8.36). □

Corollary 8.39. *Let $p \ge 2$. Let $w_n > 0$. Let (Y_n) be a sequence of independent symmetric random variables with $\|Y_n\|_2 = w_n$, $\|Y_n\|_p = 1$ and such that, for each n, $|Y_n|$ has only one non-zero value. Then the orthogonal projection P onto the closed span of (Y_n) in L_2 is bounded on L_p. Consequently, the space $X_{p,w}$ is isomorphic to a complemented subspace of L_p.*

Proof. An elementary calculation shows that, since $|Y_n|$ is a multiple of an indicator function we have

$$\|Y_n\|_p\|Y_n\|_{p'} = \|Y_n\|_2^2,$$

and hence, since $\|Y_n\|_p = 1$, $\|Y_n\|_{p'} = \|Y_n\|_2^2$. Let $\mathcal{C}_n$ be the σ-algebra generated by Y_n. Let Q be as in Corollary 8.37. Note that $\langle f, Y_n\rangle = \langle \mathbb{E}^{\mathcal{C}_n}(f) - \mathbb{E}(f), Y_n\rangle$ for all n. We have

$$\forall f \in L_2 \qquad Pf = \sum \|Y_n\|_2^{-2}\langle f, Y_n\rangle Y_n = \sum \|Y_n\|_2^{-2}\langle Qf, Y_n\rangle Y_n.$$

We have on one hand clearly $\|Pf\|_2 \le \|f\|_2 \le \|f\|_p$, and on the other one

$$\sum |\|Y_n\|_2^{-2}\langle f, Y_n\rangle|^p = \sum \|Y_n\|_2^{-2p}|\langle \mathbb{E}^{\mathcal{C}_n}(f) - \mathbb{E}(f), Y_n\rangle|^p$$

$$\le \sum (\|Y_n\|_2^{-2}\|Y_n\|_{p'})^p\|\mathbb{E}^{\mathcal{C}_n}(f) - \mathbb{E}(f)\|_p^p = \sum \|\mathbb{E}^{\mathcal{C}_n}(f) - \mathbb{E}(f)\|_p^p \le (2^{1/p'}\|f\|_p)^p,$$

where at the last step we used (8.49); therefore Corollary 8.38 yields

$$\|Pf\|_p \le \beta_p(1+2^{1/p'})\|f\|_p.$$

□

Remark 8.40. In the preceding statement, the dual of the space $X_{p,w}$ can be identified with the closed span in $L_{p'}$ of the variables (Y_n). Since $X_{p,w}$ is the intersection of ℓ_p with a weighted ℓ_2-space, its dual is the sum of the respective duals. It follows that, the series $\sum x_n Y_n \|Y_n\|_{p'}^{-1}$ converges in $L_{p'}$ iff (x_n) admits a decomposition of the form $x_n = a_n + b_n$ with both $\sum |a_n|^{p'} < \infty$ and $\sum |w_n^{-1} b_n|^2 < \infty$, and the corresponding norms are equivalent.

Remark 8.41. Fix $p > 2$. Let $q = 2p/(p-2)$ so that $1/2 = 1/p + 1/q$. By Hölder, we have

$$(\sum w_n^2 |x_n|^2)^{1/2} \leq (\sum |x_n|^p)^{1/p} (\sum |w_n|^q)^{1/q},$$

so that, on one hand, if $\sum |w_n|^q < \infty$, then $X(p,w) = \ell_p$ and on the other hand, if $\inf w_n > 0$, obviously $X(p,w) = \ell_2$. Now if $w = (w_n)$ splits as the disjoint union of a sequence such that $\sum |w_n|^q < \infty$ and one such that $\inf w_n > 0$, then $X(p,w)$ is isomorphic to $\ell_p \oplus \ell_2$.

If none of these three cases happens, w must satisfy both $\liminf w_n = 0$ and $\sum_{n : w_n < \varepsilon} |w_n|^q = \infty$ for any $\varepsilon > 0$. Rosenthal proved that the resulting space $X(p,w)$ is actually independent of w up to isomorphism. More precisely, if w and w' are two sequences both satisfying this, then $X(p,w)$ and $X(p,w')$ are isomorphic to the same Banach space, which therefore can be denoted simply by X_p. Historically, this space was the first example of a genuinely new $\mathcal{L}_p$-space, one that was not obtained by direct sums from the classical examples ℓ_2, ℓ_p or L_p. Shortly after that breakthrough, uncountably many examples of $\mathcal{L}_p$-spaces were produced in [91].

8.8 Stein Inequalities in UMD spaces

Bourgain [83] observed that the UMD property of a Banach space B implies a certain B-valued version of Stein's inequality. In its most classical form, Stein's inequality is as follows. Consider a filtration $(\mathcal{A}_n)_{n \geq 0}$ on a probability space (Ω, P) and let $1 < p < \infty$. Then for *any* sequence $(F_n)_{n \geq 0}$ in L_p we have

$$\left\| \left(\sum |\mathbb{E}_n F_n|^2 \right)^{1/2} \right\|_p \leq C(p) \left\| \left(\sum |F_n|^2 \right)^{1/2} \right\|_p \tag{8.50}$$

where $C(p)$ is a constant depending only on p.
When $p = 1$ this is no longer valid.

As usual in the B-valued case, the "square function" $\left(\sum |F_n|^2\right)^{1/2}$ must be replaced by an average of $\|\sum \varepsilon_n F_n\|_B$ over all signs $\varepsilon = (\varepsilon_n)$. In particular, Bourgain proved that, if B is UMD, if $F = (F_n)_{n \geq 0}$ is an *arbitrary* sequence in $L_p(\Omega, P; B)$ and if we define as before

$$R(\{F_n\}) = \left(\int \left\| \sum \varepsilon_n F_n \right\|_B^2 d\nu(\varepsilon) \right)^{1/2}, \tag{8.51}$$

then we have

$$\|R(\{\mathbb{E}_n F_n\})\|_p \leq C(p,B)\|R(\{F_n\})\|_p \tag{8.52}$$

where $C(p,B)$ depends only on p and B (see Theorem 8.42 below).

When $B = \mathbb{R}$ or $\mathbb{C}$ or when B is Hilbertian, we recover Stein's inequality.

Theorem 8.42. *Let $F = (F_n)_{n\geq 0}$ be an arbitrary sequence in $L_p(\Omega, P; B)$. We have for any $1 < p < \infty$*

$$\left\|\sum \varepsilon_n \mathbb{E}_n(F_n)\right\|_{L_p(d\nu\times dP;B)} \leq C_p(B)\left(\int \left\|\sum \varepsilon_n F_n\right\|^p d\nu\ dP\right)^{1/p}. \tag{8.53}$$

Proof. Consider as usual $D = \{-1,+1\}^{\mathbb{N}}$ equipped with the filtration

$$\mathcal{B}_n = \tau(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n).$$

Then we define a filtration $(\mathcal{C}_n)_{n\geq 0}$ on $\Omega \times D$ by setting

$$\mathcal{C}_{2j} = \mathcal{A}_j \otimes \mathcal{B}_j$$
$$\mathcal{C}_{2j+1} = \mathcal{A}_{j+1} \otimes \mathcal{B}_j.$$

Note that this is an increasing filtration. Now consider $f \in L_p(\Omega\times D; B)$ defined by

$$f = \sum_{n\geq 0} F_n \varepsilon_n.$$

We will apply the preceding results to the martingale

$$f_n = \mathbb{E}^{\mathcal{C}_n}(f).$$

Note that we have

$$f_{2j} = \sum_{n\leq j} \mathbb{E}_j(F_n)\varepsilon_n$$

and

$$f_{2j+1} = \sum_{n\leq j} \mathbb{E}_{j+1}(F_n)\varepsilon_n.$$

This implies that the increments are of two kinds: on one hand

$$df_{2j+1} = \sum_{n\leq j} d(F_n)_{j+1}\varepsilon_n$$

and on the other

$$df_{2j} = \mathbb{E}_j(F_j)\varepsilon_j.$$

Thus, by the definition of UMD$_p$, we find

$$\left\|\sum \mathbb{E}_j(F_j)\varepsilon_j\right\|_{L_p(B)} = \left\|\sum df_{2j}\right\|_{L_p(B)} \leq C_p(B)\|f\|_{L_p(B)}. \tag{8.54}$$

□

When B is isomorphic to a Hilbert space (and in some sense only then, see [184]), then $R(\{x_n\})$ is equivalent to

$$\left(\sum \|x_n\|^2\right)^{1/2},$$

but in a general Banach space these two ways to measure the "quadratic variation" of a sequence are quite different.

8.9 H^1 spaces. Atoms. BMO

The Hardy space H^1 has many analogues in martingale theory. The main one is probably as follows: we define the space "martingale-H^1" relative to a filtration $(\mathcal{A}_n)_{n\geq 0}$ to be the space of scalar valued martingales $(f_n)_{n\geq 0}$ such that the maximal function $f^* = \sup_{n\geq 0} |f_n|$ is in L^1. By convention, we always set $\mathcal{A}_{-1} = \{\Omega, \phi\}$. The space "martingale-BMO" is then defined as the space of all martingales $(f_n)_{n\geq 0}$ converging in L_1 such that

$$\sup_{n\geq 1} \|\mathbb{E}_n|f_\infty - f_{n-1}|\|_\infty < \infty.$$

Equivalently, let us assume $\mathcal{A} = \mathcal{A}_\infty$. For any f in L_1, we set

$$\|f\|_{BMO} = \sup_{n\geq 1} \|\mathbb{E}_n|f - f_{n-1}|\|_\infty < \infty.$$

We will identify BMO with the space of all $\mathcal{A}_\infty$-measurable f's for which this is finite. Note that, strictly speaking this is not a norm, only one "modulo the constants", i.e. it becomes one if we pass to the quotient modulo the (one dimensional) subspace spanned by the constant functions.

We will need to work with "regular" filtrations

Definition 8.43. A filtration $(\mathcal{A}_n)_{n\geq 0}$ is called regular if there is a constant $C \geq 1$ such that, for all $n \geq 1$ and for all $f \geq 0$ in $L_1(\Omega, \mathcal{A}, \mathbb{P})$, we have

$$\mathbb{E}_n(f) \leq C\mathbb{E}_{n-1}(f). \tag{8.55}$$

We will also assume for convenience that the initial σ-field $\mathcal{A}_0$ is trivial.

For example it is easy to see that the dyadic filtration is regular. More generally, if $\mathcal{A}_n$ is finite for all n and there is $\delta > 0$ such that, for all $n \geq 1$, for all atoms α of $\mathcal{A}_{n-1}$ and all atoms $\alpha' \subset \alpha$ of $\mathcal{A}_n$, we have $\mathbb{P}(\alpha')\mathbb{P}(\alpha)^{-1} \geq \delta$, then the filtration is regular (the dyadic case corresponds to $\delta = 1/2$).

If the filtration is regular and (8.55) holds, then

$$C^{-1}\|f\|_{BMO} \leq \sup_{n\geq 0} \|\mathbb{E}_n|f - f_n|\|_\infty \leq 2\|f\|_{BMO}. \tag{8.56}$$

Indeed, we have

$$\mathbb{E}_n|f - f_{n-1}| \leq C\mathbb{E}_{n-1}|f - f_{n-1}| \leq C \sup_{n\geq 0} \|\mathbb{E}_n|f_\infty - f_n|\|_\infty,$$

whence the first inequality. Also $|f_n - f_{n-1}| \leq \mathbb{E}_n|f - f_{n-1}| \leq \|f\|_{BMO}$ and since $|f - f_n| \leq |f - f_{n-1}| + |f_n - f_{n-1}|$, we obtain the other side.

The martingale version of the Fefferman duality theorem then says that, in the regular case, the space BMO can be identified with the dual of H^1, the duality being:

$$\langle g, f\rangle = \lim_{n\to\infty} \mathbb{E}(g_n f_n).$$

It is well known that all this can be extended rather easily to the Banach space valued case, as follows (cf. [78], see also [79]).

Let $(\mathcal{A}_n)_{n\geq 0}$ be a fixed filtration on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let B be a Banach space. We will denote by $H^1_{max}(B)$ the space of all B-valued martingales $(f_n)_{n\geq 0}$ converging in $L_1(B)$ and such that

$$\mathbb{E} \sup_{n\geq 0} \|f_n\|_B < \infty,$$

equipped with the norm $\|f\|_{H^1_{max}(B)} = \mathbb{E} \sup_{n\geq 0} \|f_n\|_B$. We will again denote by f^* the maximal function, i.e. we set

$$f^*(\cdot) = \sup_{n\geq 0} \|f_n(\cdot)\|_B.$$

Remark. Note that, if $p > 1$, we could also define more generally the Banach space $H^p_{max}(B)$ as the space of all B-valued martingales $(f_n)_{n\geq 0}$ such that $\mathbb{E}\sup_{n\geq 0} \|f_n\|_B^p < \infty$, equipped with the norm $\|f\|_{H^p_{max}(B)} = (\mathbb{E}\sup_{n\geq 0} \|f_n\|_B^p)^{1/p}$. But then, by Doob's maximal inequality (1.20), this simply would coincide with the space of L_p-bounded martingales denoted earlier by $h_p(\Omega, (\mathcal{A}_n)_{n\geq 0}, \mathbb{P}; B)$; moreover, if B has the RNP, this can be identified (see Remark 2.12) with the space $L_p(\Omega, \mathcal{A}_\infty, \mathbb{P}; B)$.

The duality between H^1 and BMO can be reformulated nicely using "atoms", as follows. A function $a\colon\ \Omega \to B$ in $L^1(\Omega, \mathbb{P}; B)$ is called an atom (relative to our fixed filtration $(\mathcal{A}_n)_{n\geq 0}$) if there is an integer $n \geq 0$ and a set $A \in \mathcal{A}_n$ such that

$$\{a \neq 0\} \subset A$$

$\mathbb{E}_n(a) = 0$ and $\|a\|_{L_\infty(B)} \leq 1/\mathbb{P}(A)$. The space $H^1_{at}(B)$ (relative to our fixed filtration) is then defined as the space of all functions f in $L_1(B)$ which can be written as an absolutely convergent series of the form

$$f = \mathbb{E}_0(f) + \sum_{n=1}^{\infty} \lambda_n a_n$$

where a_n are atoms and $\sum |\lambda_n| < \infty$. We define

$$\|f\|_{H^1_{at}(B)} = \|\mathbb{E}_0(f)\| + \inf\left\{\sum |\lambda_n|\right\}$$

where the infimum runs over all possible such representations of f.

Theorem 8.44. *In the regular case (for example in the dyadic case), the spaces $H^1_{max}(B)$ and $H^1_{at}(B)$ are identical and their norms are equivalent, with equivalence constants independent of B.*

Proof. If a is an atom, we have clearly $\|a\|_{L_1(B)} \le \|a\|_{L_2(B)} \le 1$ hence Doob's inequality yields $\|a^*\|_2 \le 2$ and hence $\|a\|_{H^1_{\max}(B)} \le 2$, so that we find

$$\|f\|_{H^1_{max}(B)} \le 2\|f\|_{H^1_{at}(B)}.$$

To prove the converse we will use the following consequence of (8.55): for any $\mathcal{A}_n$-measurable $f \ge 0$ we have

$$\mathbb{E}_{n-1}(f) \le C^3(\mathbb{E}_{n-1}(f^{1/2}))^2. \tag{8.57}$$

Indeed, (8.55) obviously implies $f^2 \le C^2\mathbb{E}_{n-1}(f)^2$ and hence $(\mathbb{E}_{n-1}f^2)^{1/2} \le C\mathbb{E}_{n-1}(f)$. But now by Hölder's inequality we have if $1 = \frac{1-\theta}{p} + \frac{\theta}{q}$

$$\mathbb{E}_{n-1}(f) \le (\mathbb{E}_{n-1}f^p)^{\frac{1-\theta}{p}}(\mathbb{E}_{n-1}f^q)^{\frac{\theta}{q}}$$

hence choosing $p = 1/2$, $q = 2$, $\theta = 2/3$ we find

$$(\mathbb{E}_{n-1}f^2)^{1/2} \le C\mathbb{E}_{n-1}(f) \le C(\mathbb{E}_{n-1}\sqrt{f})^{2/3}(\mathbb{E}_{n-1}f^2)^{1/3},$$

which implies (8.57) (once we divide by $(\mathbb{E}_{n-1}f^2)^{1/3}$, raise to the power 3 and note that $\mathbb{E}_{n-1}(f) \le (\mathbb{E}_{n-1}f^2)^{1/2}$).

Consider now f with

$$\|f\|_{H^1_{\max}(B)} = \mathbb{E}f^* \le 1.$$

We will prove that

$$\|f\|_{H^1_{at}(B)} \le \|\mathbb{E}(f)\| + 9C + 64C^4 \le 1 + 9C + 64C^4. \tag{8.58}$$

Clearly (replacing f by $f - \mathbb{E}_0 f$) we may assume $f_0 = \mathbb{E}f = 0$. As usual we let $d_n = f_n - f_{n-1}$. Then for any $m \ge 0$ we introduce the stopping time

$$T_m = \inf\{n \ge 0 \mid \|f_n\| + \mathbb{E}_n\|d_{n+1}\| > 2^m\}.$$

Note that $T_m > 0$ since $f_0 = 0$ and $\mathbb{E}_0\|d_1\| = \mathbb{E}\|f_1\| \le 1$. We claim that

$$\|f_{T_m}\| \le C2^m. \tag{8.59}$$

Indeed, if $T_m = n > 0$ we have $\|f_{n-1}\| + \mathbb{E}_{n-1}\|d_n\| \le 2^m$ and hence by (8.55) $\|f_n\| \le \|f_{n-1}\| + \|d_n\| \le C2^m$. We can now conclude: since $T_m \uparrow \infty$ we can write

$$f = f_{T_0} + \sum_{m \ge 1} f_{T_m} - f_{T_{m-1}}.$$

Let $a_0 = C^{-1}f_{T_0}$. Then $\|a_0\|_\infty \le 1$, $\mathbb{E}_0(a_0) = \mathbb{E}f = 0$, therefore a_0 is an atom relative to $\mathcal{A}_0$ (with support included in Ω). For any $m \ge 1$ and $n > 0$ we set

$$a_{m,n} = (f_{T_m} - f_{T_{m-1}}) \cdot 1_{\{T_{m-1}=n\}}(C2^{m+1}\mathbb{P}\{T_{m-1} = n\})^{-1}.$$

Then $a_{m,n}$ is an atom: indeed it is supported on $\{T_{m-1} = n\}$ and (8.59) implies $\|a_{m,n}\|_\infty \le \mathbb{P}\{T_{m-1} = n\}^{-1}$. Here we assume $\mathbb{P}\{T_{m-1} = n\} > 0$ otherwise we set for notational convenience $a_{m,n} = 0$. Finally $\mathbb{E}_n(a_{m,n}) = 0$. Indeed, we have $T_{m-1} \le T_m$, so that $T_m \wedge n = T_{m-1} \wedge n$ when $T_{m-1} = n$, and since $\{T_{m-1} = n\} \in \mathcal{A}_n$, by (1.10) we may write

$$\mathbb{E}_n(1_{\{T_{m-1}=n\}}(f_{T_m} - f_{T_{m-1}})) = 1_{\{T_{m-1}=n\}}(f_{T_m\wedge n} - f_{T_{m-1}\wedge n}) = 0.$$

We can now complete the proof of (8.58). We have

$$f = Ca_0 + \sum_{m,n>0} C2^{m+1}\mathbb{P}\{T_{m-1} = n\}a_{m,n}$$

therefore

$$\begin{aligned}\|f\|_{H^1_{\text{at}}(B)} &\le C + C\sum_{m,n>0} 2^{m+1}\mathbb{P}\{T_{m-1} = n\}\\ &= C + C\sum_{m>0} 2^{m+1}\mathbb{P}\{T_{m-1} < \infty\}.\end{aligned}$$

Note that

$$\mathbb{P}(T_{m-1} < \infty) \le \mathbb{P}\{f^* + \sup_n \mathbb{E}_n\|d_{n+1}\| > 2^{m-1}\}$$

and hence if we set $Z = f^* + \sup_n \mathbb{E}_n\|d_{n+1}\|$

$$\sum 2^{m+1}\mathbb{P}(T_{m-1} < \infty) \le \mathbb{E}\left(\sum_n 2^{m+1}1_{\{2^{m-1}<Z\}}\right) \le 8\mathbb{E}Z.$$

Finally, since $\|d_{n+1}\| \le 2f^*$, by (8.57) we have

$$\mathbb{E}_n\|d_{n+1}\| \le C^3(\mathbb{E}_n(\|d_{n+1}\|^{1/2}))^2 \le 2C^3(\mathbb{E}_n\sqrt{f^*})^2$$

therefore by Doob's inequality (i.e. (1.12) with $p = 2$)

$$\mathbb{E}\sup_n \mathbb{E}_n\|d_{n+1}\| \le 2C^3\mathbb{E}(\sup_n \mathbb{E}_n\sqrt{f^*})^2 \le 8C^3\mathbb{E}f^* \le 8C^3.$$

Thus we finally obtain as announced

$$\|f\|_{H^1_{at}(B)} \le C + 8C\mathbb{E}Z \le C + 8C(1 + 8C^3) = 9C + 64C^4.$$

□

Let us now assume that each $\mathcal{A}_n$ is finite and the σ-algebra $\mathcal{A}$ is generated by the $\mathcal{A}_n$'s (so that $\mathcal{A} = \mathcal{A}_\infty$).

Assume $1 \le p < \infty$, and $p^{-1} + (p')^{-1} = 1$. Recall that the dual of $L_p(B)$ can be identified with the space $h_{p'}(\Omega, (\mathcal{A}_n)_{n\ge0}, P; B^*)$ formed of all the B^*-valued martingales which are bounded in $L_{p'}(B^*)$, equipped with the norm

$\|\varphi\| = \sup_{n\geq 0} \|\varphi_n\|_{L_{p'}(B^*)}$. For all f in $L_p(B)$ and all such φ, the duality is defined by setting

$$\langle \varphi, f\rangle = \lim_{n\to\infty} \mathbb{E}\langle \varphi_n, f_n\rangle = \lim_{n\to\infty} \mathbb{E}\langle \varphi_n, f\rangle.$$

It is easy to see (by the density of $\bigcup_{n\geq 0} L_p(\Omega, \mathcal{A}_n, \mathbb{P}; B)$ in $L_p(B)$) that the preceding limits exist. Then, by the above Proposition 2.14, we have isometrically

$$L_p(B)^* = h_{p'}(\Omega, (\mathcal{A}_n)_{n\geq 0}, P; B^*).$$

The B^*-valued analogue of BMO can then be defined. We will denote by $BMO_\sigma(B^*)$ the space of all martingales $\varphi = (\varphi_n)_{n\geq 0}$ in $L_1(B^*)$ such that $\sup_{n\geq 0}\sup_{m\geq n} \|\mathbb{E}_n\|\varphi_m - \varphi_n\|_{B^*}\|_\infty < \infty$ and we equip it with the "norm" (modulo constants)

$$\|\varphi\| = \sup_{n\geq 0}\sup_{m\geq n} \|\mathbb{E}_n\|\varphi_m - \varphi_n\|_{B^*}\|_\infty.$$

We have then the following easy result:

Theorem 8.45. *In the same duality as above, in the regular case, we have $H^1_{at}(B)^* = BMO_\sigma(B^*)$. More precisely if we let*

$$\|\varphi\|_* = \sup\{|\langle\varphi, f\rangle| \mid f \in H^1_{at}(B) \quad \|f\|_{H^1_{at}(B)} \leq 1\},$$

then, assuming $\varphi_0 = 0$, we have

$$\|\varphi\|_* \leq \|\varphi\|_{BMO_\sigma(B^*)} \leq 2\|\varphi\|_*. \tag{8.60}$$

Sketch. The preceding supremum can be restricted to the set of atoms. Moreover we can restrict the atoms (by martingale convergence) to be in

$$\bigcup_{m\geq 0} L_1(\Omega, \mathcal{A}_m, \mathbb{P}; B).$$

Let

$$\|\varphi\|_* = \sup\{|\langle\varphi, f\rangle| \mid f \in H^1_{at}(B)\ \|f\|_{H^1_{at}(B)} \leq 1\}.$$

If an atom a is $\mathcal{A}_m$-measurable with $\{a \neq 0\} \subset A$, $A \in \mathcal{A}_n$, $\mathbb{E}_n(a) = 0$ and $\|a\|_{L_\infty(B)} \leq \mathbb{P}(A)^{-1}$, we have

$$\begin{aligned}\langle\varphi, a\rangle &= \lim_{k\to\infty} \mathbb{E}\langle\varphi_k, a_k\rangle = \mathbb{E}\langle\varphi_m, a\rangle\\ &= \mathbb{E}\langle\varphi_m - \mathbb{E}_n(\varphi_m), a\rangle\end{aligned}$$

hence

$$|\langle\varphi, a\rangle| \leq \frac{1}{\mathbb{P}(A)}\int_A \|\varphi_m - \mathbb{E}_n(\varphi_m)\|_{B^*} d\mathbb{P} \leq \text{ess sup } \mathbb{E}_n\|\varphi_m - \mathbb{E}_n(\varphi_m)\|_{B^*}$$

(where the last inequality uses $A \in \mathcal{A}_n$).

Thus we obtain for all atoms a

$$|\langle \varphi, a\rangle| \le \|\varphi\|_{BMO_\sigma(B^*)}. \tag{8.61}$$

Conversely, it is easy to check that

$$\|\varphi\|_{BMO_\sigma(B^*)} \le 2\sup\{|\langle \varphi, a\rangle| \mid \text{ a atom}\}. \tag{8.62}$$

Indeed, if $|\langle \varphi, a\rangle| \le 1$ for all a, then we deduce that, for all A in $\mathcal{A}_n$ and all functions b in the unit ball of $L_\infty(\mathcal{A}, \mathbb{P}; B)$, since $2^{-1}\mathbb{P}(A)^{-1}1_A(b - \mathbb{E}_n(b))$ is an atom, we have

$$\mathbb{P}(A)^{-1}|\langle \varphi_m, 1_A(b - \mathbb{E}_n(b))\rangle| \le 2$$

or equivalently

$$\mathbb{P}(A)^{-1}\left|\int_A \langle \varphi_m - \mathbb{E}_n(\varphi_m), b\rangle\right| \le 2$$

which implies (taking the sup over all b's) that

$$\mathbb{P}(A)^{-1}\int_A \|\varphi_m - \mathbb{E}_n(\varphi_m)\|_{B^*} \le 2,$$

completing the proof of (8.62). □

By a famous theorem of John and Nirenberg, any φ in $BMO_\sigma(B^*)$ automatically is in $L_q^\sigma(B^*)$ for all $q < \infty$. More precisely if we define

$$\|\varphi\|_{[q]} = \sup_m \sup_{n\ge 0} \|(\mathbb{E}_n\|\varphi_m - \mathbb{E}_{n-1}\varphi_m\|^q)^{1/q}\|_\infty$$

then we have $\|\varphi\|_{[q]} < \infty$ for any $q < \infty$ and there is a numerical constant K such that for all $1 \le q < \infty$

$$\|\varphi\|_{BMO_\sigma(B^*)} \le \|\varphi\|_{[q]} \le Kq\|\varphi\|_{BMO_\sigma(B^*)}.$$

8.10 Burkholder's geometric characterization of UMD space

In [103], Burkholder found a somewhat geometric condition, that he called ζ-convexity that is equivalent to the UMD property.

Definition 8.46. A Banach space B is called ζ-convex if there is a function $\zeta\colon\ B \times B \to \mathbb{R}$ that is symmetric (i.e. $\zeta(x, y) = \zeta(y, x)$), separately concave in each of the two variables, satisfying $\zeta(0, 0) > 0$ and such that

$$\forall x, y \in B \qquad \zeta(x, y) \le \|x + y\| \quad \text{whenever} \quad \|x\| \le 1 \le \|y\|.$$

Theorem 8.47. *A Banach space B is UMD iff it is ζ-convex.*

The following result from [178] is closely connected to Burkholder's characterization but the meaning of the condition is somewhat easier to grasp, since it involves only the rather standard class of (differences of) convex continuous functions.

Theorem 8.48. *Let B be a real Banach space. We set $X = B \oplus B^*$. Then B is UMD iff the function $\varphi\colon\ X \to \mathbb{R}$ defined by*

$$\varphi(x, x^*) = x^*(x)$$

is the difference of two convex continuous functions on X.

Remark 8.49. Let X be a Banach space. Then $\varphi\colon\ X \to \mathbb{R}$ is the difference of two convex continuous functions on X iff there is a convex continuous $\psi\colon\ X \to \mathbb{R}$ such that $\psi \pm \varphi$ are both convex and continuous. Indeed, if $\psi \pm \varphi$ are convex and continuous then $\psi = \varphi_1 - \varphi_2$ with $\varphi_1 = (\psi + \varphi)/2$, $\varphi_2 = (\psi - \varphi)/2$. Conversely, if $\varphi = \varphi_1 - \varphi_2$ with φ_1, φ_2 convex continuous then if $\psi = \varphi_1 + \varphi_2, \psi$ is convex continuous and both $\psi + \varphi$ and $\psi - \varphi$ are also convex and continuous. □

In the rest of this section we set $\Omega = \{-1, 1\}^{\mathbb{N}_*}$, denote by $\varepsilon_k\colon\ \Omega \to \{-1, 1\}$ the k-th coordinate for $k = 1, 2, \ldots$. We set $\mathcal{A}_0 = \{\phi, \Omega\}$ and $\mathcal{A}_n = \sigma(\varepsilon_1, \ldots, \varepsilon_n)$ for all $n \geq 1$.

We will use the following.

Lemma 8.50. *Let $V\colon\ X \to X^*$ be a bounded linear operator. Assume that there is a constant C such that for all finite X-valued dyadic martingales with $f_0 = 0$ we have*

$$\sum\nolimits_1^\infty \mathbb{E}|V(df_n)(df_n)| \leq C\|f\|^2_{L_\infty(X)}. \tag{8.63}$$

Then there is a constant C' such that for all such (f_n) we actually have

$$\sum\nolimits_1^\infty \mathbb{E}|V(df_n)(df_n)| \leq C'\mathbb{E}\|f\|^2_X. \tag{8.64}$$

Proof. Fix $k \geq 0$. For simplicity we set $d_n = df_n$. We first claim that

$$\mathbb{E}_k \sum_{n>k} |V(d_n)(d_n)| \leq 4C\|f\|^2_{L_\infty(B)}. \tag{8.65}$$

Indeed, consider $f - f_k$ fix $(\varepsilon_1, \ldots, \varepsilon_k)$, and let

$$\forall \omega \in \{-1, 1\}^{\mathbb{N}} \qquad F(\omega) = f(\varepsilon_1, \ldots, \varepsilon_k, \omega) - f_k(\varepsilon_1, \ldots, \varepsilon_k).$$

Applying (8.63) to F with $(\varepsilon_1, \ldots, \varepsilon_k)$ fixed we find

$$\sum\nolimits_{n>k} \mathbb{E}_\omega |V(d_n)(d_n)|(\varepsilon_1, \ldots, \varepsilon_k, \omega) \leq C \sup\nolimits_\omega \|f(\varepsilon_1, \ldots, \varepsilon_k, \omega) - f_k(\varepsilon_1, \ldots, \varepsilon_k)\|^2,$$

and hence for any stopping time T_0

$$\int_{\{T_0=k\}} \sum\nolimits_{n>T_0} |V(d_n)(d_n)| \leq C\mathbb{P}\{T_0 = k\}\|f - f_{T_0}\|^2_{L_\infty(X)},$$

so that summing over k we find

$$\mathbb{E}\left(1_{\{T_0<\infty\}}\sum\nolimits_{n>T_0}|V(d_n)(d_n)|\right)\le C\mathbb{P}\{T_0<\infty\}\|f-f_{T_0}\|^2_{L_\infty(X)}. \tag{8.66}$$

Now let $T_1\ge T_0$ be another stopping time. Replacing f by f_{T_1} in (8.66) we find

$$\mathbb{E}\left(1_{\{T_0<\infty\}}\sum\nolimits_{T_0<n\le T_1}|V(d_n)(d_n)|\right)\le C\mathbb{P}\{T_0<\infty\}\|f_{T_1}-f_{T_0}\|^2_{L_\infty(X)}. \tag{8.67}$$

We will now prove (8.64). We may assume by homogeneity that $\mathbb{E}\|f\|_X^2=1$. We define $T_0=\inf\{n\ge 0\mid \|f_n\|+\|d_{n+1}\|>1\}$ and for all $m\ge 1$

$$T_m=\inf\{n>T_{m-1}\mid \|f_n-f_{T_{m-1}}\|+\|d_{n+1}\|>2^m\}.$$

We have then

$$\mathbb{E}\sum\nolimits_1^\infty|V(d_n)(d_n)|=\mathrm{I}+\mathrm{II}$$

where $\mathrm{I}=\mathbb{E}\sum_{n\le T_0}|V(d_n)(d_n)|$ and $\mathrm{II}=\sum_{m\ge1}\mathbb{E}\sum_{T_{m-1}<n\le T_m}|V(d_n)(d_n)|$. But by (8.67) we have

$$\mathrm{II}\le\sum\nolimits_{m\ge1}C\mathbb{P}\{T_{m-1}<\infty\}\|f_{T_m}-f_{T_{m-1}}\|^2_{L_\infty(X)}. \tag{8.68}$$

Since $\|f_{T_m}-f_{T_{m-1}}\|\le 2^m$ if $T_m>T_{m-1}$ (because $\|f_{T_m-1}-f_{T_{m-1}}\|+\|d_{T_m}\|\le 2^m$) and also $f_{T_m}-f_{T_{m-1}}=0$ if $T_m=T_{m-1}$, we have $\|f_{T_m}-f_{T_{m-1}}\|_{L_\infty(X)}\le 2^m$. Moreover

$$\mathbb{P}\{T_{m-1}<\infty\}\le\mathbb{P}\{2\sup\nolimits_{n,k}\|f_n-f_k\|+\|d_{n+1}\|>2^m\}\le\mathbb{P}\{4\sup\|f_n\|>2^m\}$$

so (8.68) implies (setting $f^*=\sup\|f_n\|$)

$$\begin{aligned}\mathrm{II}&\le C\sum\nolimits_{m\ge1}\mathbb{P}\{4f^*>2^m\}2^{2m}\\&\le C'\mathbb{E}f^{*2}\le 4C'\mathbb{E}\|f\|_X^2\le 4C',\end{aligned}$$

where at the last step we used Doob's inequality. Moreover by (8.63) applied with f_{T_0} in place of f we have $\mathrm{I}\le C\|f_{T_0}\|^2_{L_\infty(X)}$ and if $T_0\ge1$ we again have $\|f_{T_0}\|\le\|f_{T_0-1}\|+\|d_{T_0}\|\le 1$ while if $T_0=0$ we have $f_{T_0}=0$. Therefore we find $\mathrm{I}\le C$ and we conclude $\mathbb{E}\sum_1^\infty|V(d_n)(d_n)|\le 1+4C'$. By homogeneity this proves the announced result. □

Proof of Theorem 8.48. Let $V\colon X\to X^*$ be the unique self-adjoint linear map such that $V(x)(x)=\varphi(x)$ for all x in X. Equivalently, this means

$$\forall x,y\in X\quad V(x)(y)=V(y)(x)=(\varphi(x+y)-\varphi(x-y))/4. \tag{8.69}$$

If $x=(b,b^*)$ and $y=(c,c^*)$, we have $V(x)(y)=(b^*(c)+c^*(b))/2$. Note that $\|V\|\le 1$ (actually $\le 1/2$). Assume that B and hence X is UMD. We claim that if C is the UMD_2 constant of X then any finite X-valued martingale (f_n) in $L_2(X)$ satisfies

$$\sum\nolimits_1^\infty\mathbb{E}|V(d_n)(d_n)|\le C\mathbb{E}\|f\|_X^2.$$

Indeed, since $V(d_n)(d_n)$ is predictable (recall that in the dyadic case $d_n = \varepsilon_n\varphi_{n-1}$ with φ_{n-1} $\mathcal{A}_{n-1}$-measurable), the random variable $\xi_n = \operatorname{sign}(V(d_n)(d_n))$ is $\mathcal{A}_{n-1}$-measurable ("predictable") so that we can write

$$\sum\nolimits_1^n \mathbb{E}|V(d_n)(d_n)| = \mathbb{E}\sum\nolimits_1^\infty \xi_n V(d_n)(d_n) = \mathbb{E}\left(V\left(\sum\nolimits_1^\infty \xi_n d_n\right)\left(\sum\nolimits_0^\infty d_n\right)\right)$$

and hence recalling Remark 8.11

$$\begin{aligned}&\le \|V\|\mathbb{E}\left\|\sum\nolimits_1^\infty \xi_n d_n\right\|\left\|\sum\nolimits_0^\infty d_n\right\|\\ &\le \|V\|\left\|\sum\nolimits_1^\infty \xi_n d_n\right\|_{L_2(X)}\left\|\sum\nolimits_0^\infty d_n\right\|_{L_2(X)}\\ &\le C\|f\|^2_{L_2(X)}.\end{aligned}$$

This proves our claim.

Using this claim we define for any x in X

$$\psi(x) = \inf\left\{C\mathbb{E}\|f\|_X^2 - \sum\nolimits_1^\infty |\mathbb{E}V(df_n)(df_n)|\right\}$$

where the infimum runs over all finite dyadic martingales (f_n) with $f_0 = x$.

We will now show that for any y in X

$$|V(y)(y)| \le 2^{-1}(\psi(x+y) + \psi(x-y)) - \psi(x). \tag{8.70}$$

Let $\varepsilon > 0$ and let (f_n), (g_n) be such that $f_0 = x + y$, $g_0 = x - y$ and

$$\begin{aligned}C\mathbb{E}\|f\|^2 - \sum\nolimits_1^\infty \mathbb{E}|V(df_n)(df_n)| &< \psi(x+y) + \varepsilon\\ C\mathbb{E}\|g\|^2 - \sum\nolimits_1^\infty \mathbb{E}|V(dg_n)(dg_n)| &< \psi(x-y) + \varepsilon.\end{aligned}$$

We then define a dyadic martingale F_n by setting $F_0 = x$, $F_1 = x + \varepsilon_1 y$ and then for $n > 1$ $F_n(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = f(\varepsilon_2, \ldots, \varepsilon_n)$ if $\varepsilon_1 = 1$ and $= g(\varepsilon_2, \ldots, \varepsilon_n)$ if $\varepsilon_1 = -1$. We then find since $|V(dF_1)(dF_1)| = |V(y)(y)|$

$$\begin{aligned}\psi(x) &\le C\mathbb{E}\|F\|^2 - \sum\nolimits_1^\infty \mathbb{E}|V(dF_n)(dF_n)|\\ &< 2^{-1}(\psi(x+y) + \psi(x-y)) + \varepsilon - |V(y)(y)|\end{aligned}$$

and hence we obtain (8.70). But then (8.70) can be rewritten as

$$|V(y)(y)| = |2^{-1}(\varphi(x+y)+\varphi(x-y))-\varphi(x)| \le 2^{-1}(\psi(x+y)+\psi(x-y))-\psi(x) \tag{8.71}$$

so by Remark 8.49 this implies that $\psi \pm \varphi$ is convex, or more precisely mid-convex, but since we obviously have (consider $f_n = x$ $\forall n \ge 1$) $\psi(x) \le C\|x\|^2$, it follows that $\psi \pm \varphi$ are bounded on bounded sets and hence by classical results (cf. [53, p. 215]) they are actually both convex and continuous. This completes the proof of the "only if" part.

Conversely, assume that φ is the difference of two convex continuous functions. By Remark 8.49 there is ψ convex continuous such that $\psi \pm \varphi$ is convex and hence (8.71) holds. Let (f_n) be a finite X-valued dyadic martingale, and let $d_n = df_n$. Applying (8.71) with $x = f_{n-1}$, $y = d_n$ we find for all $n \geq 1$

$$|V(d_n)(d_n)| \leq \mathbb{E}_{n-1}(\psi(f_n) - \psi(f_{n-1}))$$

and hence after integration

$$\sum\nolimits_1^\infty \mathbb{E}|V(d_n)(d_n)| \leq \mathbb{E}(\psi(f) - \psi(0)).$$

Assume $f_0 = 0$. Since ψ is continuous, there is $r > 0$ such that $\|x\| \leq r$ implies $|\psi(x) - \psi(0)| \leq 1$. Therefore if $\|f\|_{L_\infty(X)} \leq r$ we find $\sum_1^\infty \mathbb{E}|V(d_n)(d_n)| \leq 1$. By homogeneity, this implies that

$$\sum\nolimits_1^\infty \mathbb{E}|V(d_n)(d_n)| \leq (1/r)^2 \|f\|^2_{L_\infty(X)}, \tag{8.72}$$

and hence, by Lemma 8.50, we obtain (8.64). Let (g_n) be another finite dyadic X-valued martingale. Let $d'_n = dg_n$. Clearly by polarization

$$V(d_n)(d'_n) = 4^{-1}(V(d_n + d'_n)(d_n + d'_n) - V(d_n - d'_n)(d_n - d'_n))$$

and hence (8.64) implies

$$\sum\nolimits_1^\infty \mathbb{E}|V(d_n)(d'_n)| \leq (C'/4)(\|f+g\|^2_{L_2(X)} + \|f-g\|^2_{L_2(X)}). \tag{8.73}$$

Now let $\xi_n = \pm 1$ be arbitrary signs. Let $\tilde{f} = \sum \xi_n d_n$. Since $V\colon\ X \to X^*$ is isometric, we have

$$\begin{aligned}\|\tilde{f}\|_{L_2(X)} &= \|V(\tilde{f})\|_{L_2(X^*)},\\ &= \sup\{\mathbb{E}V(\tilde{f})(g) \mid g \in B_{L_2(X)}\}.\end{aligned}$$

But for $g \in B_{L_2(X)}$ we have

$$\mathbb{E}V(\tilde{f})(g) = \mathbb{E}\sum\nolimits_1^\infty V(d\tilde{f}_n)(dg_n) = \mathbb{E}\sum\nolimits_1^\infty \xi_n V(d_n)(dg_n)$$

and hence by (8.73) if $g, f \in B_{L_2(X)}$

$$|\mathbb{E}V(\tilde{f})(g)| \leq \mathbb{E}\sum\nolimits_1^\infty |V(d_n)(dg_n)| \leq 2C'.$$

Thus we obtain $\|\tilde{f}\|_{L_2(X)} \leq 2C'$. By homogeneity, $\|\tilde{f}\|_{L_2(X)} \leq 2C'\|f\|_{L_2(X)}$. In other words, $\mathrm{UMD}_2(X) \leq 2C'$. □

8.11 Notes and Remarks

The inequalities (8.1) and (8.2) were obtained in a 1966 paper by Burkholder. We refer the reader to the classical papers [101] and [108] for more on this. See also the book [25].

The best constants in the Khintchine inequalities are known: see [254, 156]. Szarek [254] proved that $A_1 = 2^{-1/2}$. More generally, let γ_p be the L_p-norm of a standard Gaussian distribution (with mean zero and variance 1). It is well known that

$$\gamma_p = 2^{1/2} \left(\Gamma((p+1)/2)/\sqrt{\pi}\right)^{1/p} \quad 0 < p < \infty.$$

Let $p_0 = 1.87...$ be the unique solution in the interval $]1, 2[$ of the equation $2^{1/2-1/p} = \gamma_p$ (or explicitly $\Gamma((p+1)/2) = \sqrt{\pi}/2$), then Haagerup (see [156]) proved :

$$A_p = 2^{1/2-1/p} \quad 0 < p \le p_0, \tag{8.74}$$

$$A_p = \gamma_p \quad p_0 \le p \le 2, \tag{8.75}$$

$$B_p = \gamma_p \quad 2 \le p < \infty. \tag{8.76}$$

The lower bounds $A_p \ge \max\{\gamma_p, 2^{1/2-1/p}\}$ for $p \le 2$ and $B_p \ge \gamma_p$ for $p \ge 2$ are easy exercises (by the Central Limit Theorem).

For Kahane's inequalities, some of the optimal constants are also known, in particular (see [185]), if $0 < p \le 1 \le q \le 2$, we have $K(p, q) = 2^{\frac{1}{p}-\frac{1}{q}}$.

Kahane's inequalities follow from the results in the first edition of [31]. The idea to derive them from the 2-point hypercontractive inequality is due to C. Borell.

The property UMD was introduced by B. Maurey and the author (see [202]), together with the observations that Burkholder's ideas could be extended to show that $\text{UMD}_p \Leftrightarrow \text{UMD}_q$ for any $1 < p, q < \infty$. It was also noted that UMD_p implies reflexivity (and even super-reflexivity), see Chapter 3 below for more on this. The Gundy decomposition appearing in Theorem 8.13 comes from [154].

The extrapolation principle appearing in Lemma 8.20 (sometimes called "good λ-inequality") is based on the early ideas of Burkholder and Gundy ([108]), but our presentation was influenced by the refinements from [186].

§8.4 is a simple adaptation to the B-valued case of Burgess Davis classical results from [123]. §8.5 is "folkloric". §8.6 is due to B. Maurey [202]. The Burkholder–Rosenthal inequality in Theorem 8.33 appears in [101]. It was preceded by Rosenthal's paper [242] from which Corollaries 8.36 to 8.38 are extracted. §8.8 is due to Bourgain [83], but the original Stein inequality comes from [56]. The extension to the Banach valued case of the atomic decomposition of functions in $H^1(\mathbb{R})$ or $H^1(\mathbb{T})$ (related to §8.9) is due independently to Garcia–Cuerva and Bourgain ([83]). We refer to O. Blasco's [78] for a detailed account of the H^1-BMO duality in the B-valued case but in the classical setting of functions on $\mathbb{T}$ or $\mathbb{R}$. Our §8.9 is just the martingale analogue of his main result (see also [79]).

In §**??**, the main inequality (**??**) is due to Azuma [68]. §8.10 is motivated by Burkholder's characterization of UMD spaces in terms of ζ-convexity, for which we refer to [103, 104, 105, 107]. For Theorem 8.48, we refer to [178].

Appendix 1: Marcinkiewicz theorem

In the next statement, it will be convenient to use the following terminology. Let X, Y be Banach spaces, let (Ω, μ), $((\Omega', \mu')$ be measure spaces and let $T\colon\ L_p(\mu; X) \to L_0(\mu'; Y)$ be a linear operator. We say that T is of weak type (p, p) with constant C if we have for any f in $L_p(\mu; X)$

$$(\sup_{\lambda>0} \lambda^p \mu'(\|Tf\| > \lambda))^{1/p} \le C\|f\|_{L_p(X)}.$$

We say that T is of strong type (p, p) if it bounded from $L_p(X)$ to $L_p(Y)$. We invoke repeatedly the following famous classical result due to Marcinkiewicz.

Theorem 8.51 (Marcinkiewicz)**.** *Let $0 < p_0 < p_1 \le \infty$ in the above situation, assume that T is both of weak type (p_0, p_0) with constant C_0 and of weak type (p_1, p_1) with constant C_1. Then for any $0 < \theta < 1$, T is of strong type (p_θ, p_θ) with $p_\theta^{-1} = (1-\theta)p_0^{-1} + \theta p_1^{-1}$, and moreover we have*

$$\|T\colon\ L_{p_\theta}(X) \to L_{p_\theta}(Y)\| \le K(p_0, p_1, p) C_0^{1-\theta} C_1^{\theta}$$

where $K(p_0, p_1, p)$ is a constant depending only on p_0, p_1, p.

Proof. Let $f \in L_{p_0}(X) \cap L_{p_1}(X)$. Consider a decomposition

$$\begin{aligned} f &= f_0 + f_1 \quad \text{with} \\ f_0 &= f \cdot 1_{\{\|f\| > \gamma\lambda\}} \quad \text{and} \\ f_1 &= f \cdot 1_{\{\|f\| \le \gamma\lambda\}}, \end{aligned}$$

where $\gamma > 0$ and $\lambda > 0$ are fixed. We have by our assumptions

$$\mu'(\|T(f_0)\| > \lambda) \le (C_0\lambda^{-1})^{p_0} \int\limits_{\{\|f\| > \gamma\lambda\}} \|f\|^{p_0}\ d\mu$$

$$\mu'(\|T(f_1)\| > \lambda) \le (C_1\lambda^{-1})^{p_1} \int\limits_{\{\|f\| \le \gamma\lambda\}} \|f\|^{p_1}\ d\mu$$

hence since $\|T(f)\| \le \|T(f_0)\| + \|T(f_1)\|$

$$\mu'(\|T(f)\| > 2\lambda) \le C_0^{p_0}\lambda^{-p_0} \int\limits_{\|f\| > \gamma\lambda} \|f\|^{p_0} d\mu + C_1^{p_1}\lambda^{-p_1} \int\limits_{\|f\| \le \gamma\lambda} \|f\|^{p_1} d\mu. \tag{8.77}$$

Let $p = p_\theta$. If we now multiply (8.77) by $2^p p\lambda^{p-1}$ and integrate with respect to λ, using

$$\int_{\{\|f\|>\gamma\lambda\}} \lambda^{p-p_0-1} d\lambda = (p-p_0)^{-1}(\|f\|/\gamma)^{p-p_0}$$

and

$$\int_{\{\|f\|\le\gamma\lambda\}} \lambda^{p-p_1-1} d\lambda = (p_1-p)^{-1}(\|f\|/\gamma)^{p-p_1},$$

we find

$$\begin{aligned}\int \|T(f)\|^p \, d\mu' \le\ & 2^p p C_0^{p_0}(p-p_0)^{-1}\gamma^{p_0-p}\int \|f\|^p \, d\mu \\ & + 2^p p C_1^{p_1}(p_1-p)^{-1}\gamma^{p_1-p}\int \|f\|^p \, d\mu.\end{aligned}$$

Hence, we obtain the estimate

$$\begin{aligned}\|T\colon\ L_p(X) \to L_p(Y)\| \le\ & 2C_0^{p_0/p}(p-p_0)^{-1/p}\gamma^{(p_0-p)/p} \\ & + 2C_1^{p_1/p}(p_1-p)^{-1}\gamma^{(p_1-p)/p},\end{aligned}$$

so that choosing γ so that

$$C_0^{p_0}\gamma^{p_0-p} = C_1^{p_1}\gamma^{p_1-p}$$

we finally find the announced result with

$$K = 2(p-p_0)^{-1/p} + 2(p_1-p)^{-1/p}. \qquad \square$$

Remark 8.52. It is fairly obvious and well known that the preceding proof remains valid for "sublinear" operators. Indeed, all that we need for the operator T is the pointwise inequalities

$$\|T(f_0+f_1)\|_B \le \|T(f_0)\|_B + \|T(f_1)\|_B$$

for any pair f_0, f_1 in $L_{p_0}(X)\cap L_{p_1}(X)$, and also the positive homogeneity, i.e.

$$\forall\lambda\ge 0,\ \forall f\in L_{p_0}(X)\cap L_{p_1}(X)\quad \|T(\lambda f)\|_B = \lambda\|T(f)\|_B.$$

Appendix 2: Hölder-Minkowski inequality

For further reference, we wish to review here a classical set of inequalities usually referred to as "the Hölder–Minkowski inequality". Let $0 < q \le p \le \infty$ and let $(\Omega, \mathcal{A}, \mu)$ be any measure space. Consider a sequence (x_n) in $L_p(\Omega, \mathcal{A}, \mu)$. Then

$$\left\|\left(\sum |x_n|^q\right)^{1/q}\right\|_p \le \left(\sum \|x_n\|_p^q\right)^{1/q}. \tag{8.78}$$

Indeed, this is an easy consequence of the fact (since $p/q > 1$) that $L_{p/q}$ is a normed space. In particular, when $q = 1$ we find

$$\left\|\sum |x_n|\right\|_p \le \sum \|x_n\|_p$$

that is but the triangle inequality in L_p. If $0 < p \le q \le \infty$, the inequality is reversed: we have

$$\left\|\left(\sum |x_n|^q\right)^{1/q}\right\|_p \ge \left(\sum \|x_n\|_p^q\right)^{1/q}. \tag{8.79}$$

In particular, when $q = \infty$, we find simply the obvious inequality

$$\left\|\sup_n |x_n|\right\|_p \ge \sup_n \|x_n\|_p.$$

One way to check (8.79) is to set $r = q/p$, $r' = r(r-1)^{-1}$ and $y_n = |x_n|^p$. Then (8.79) is the same as

$$\left\|\left(\sum |y_n|^r\right)^{1/r}\right\|_1 \ge \left(\sum \|y_n\|_1^r\right)^{1/r}$$

that is easy to derive from

$$\left(\sum |y_n|^r\right)^{1/r} = \sup\left\{\sum \alpha_n |y_n| \;\middle|\; \alpha_n \ge 0 \sum |\alpha_n|^{r'} \le 1\right\}.$$

Indeed, we find

$$\int \left(\sum |y_n|^r\right)^{1/r} \ge \sup_{\alpha_n \ge \sum |\alpha_n|^{r'} \le 1} \int \sum |\alpha_n|\,|y_n| = \left(\sum \|y_n\|_1^r\right)^{1/r}.$$

In its simplest form (8.78) and (8.79) reduce to: $\forall x, y \in L_p$

$$\|(|x|^q + |y|^q)^{1/q}\|_p \le (\|x\|_p^q + \|y\|_p^q)^{1/q} \quad \text{if} \quad p \ge q$$
$$\|(|x|^q + |y|^q)^{1/q}\|_p \ge (\|x\|_p^q + \|y\|_p^q)^{1/q} \quad \text{if} \quad p \le q.$$

It is easy to see that actually the preceding inequalities imply conversely (8.78) and (8.79).

In the opposite direction, one can easily deduce from (8.78) and (8.79) the following refinements of (8.78) and (8.79). Let $(\Omega', \mathcal{A}', \mu')$ be another measure space. Consider a measurable function $F\colon\ \Omega \times \Omega' \to \mathbb{R}$. Then (8.78) and (8.79) become

$$\|F\|_{L_p(\mu; L_q(\mu'))} \ge \|F\|_{L_q(\mu'; L_p(\mu))} \quad \text{if} \quad p \ge q \tag{8.80}$$

$$\|F\|_{L_p(\mu; L_q(\mu'))} \le \|F\|_{L_q(\mu'; L_p(\mu))} \quad \text{if} \quad p \le q. \tag{8.81}$$

Essentially the same proof as for (8.78) and (8.79) establishes (8.80) and (8.81). Note that (8.78) and (8.79) correspond to $\Omega' = \mathbb{N}$ equipped with the counting measure $\mu' = \sum \delta_n$.

Appendix 3: Reverse Hölder principle

The classical Hölder inequality implies that for any measurable function $Z \geq 0$ on a probability space and any $0 < q < p < \infty$ we have $\|Z\|_q \leq \|Z\|_p$. By the "reverse Hölder principle" we mean the following two statements (closely related to [100]) in which the behaviour of Z in L_q controls conversely its belonging to L_p. We will use the notation

$$\|Z\|_{p,\infty} = \left(\sup_{\lambda>0} t^p \mathbb{P}\{Z > t\}\right)^{1/p}.$$

Our first principle corresponds roughly to the case $q = 0$.

Proposition 8.53. *Let $0 < p < \infty$. For any $0 < \delta < 1$ and any $R > 0$ there is a constant $C_p(\varepsilon, R)$ such that the following holds. Consider a random variable $Z \geq 0$ and a sequence $(Z^{(n)})_{n\geq 0}$ of independent copies of Z. We have then*

$$\sup\nolimits_{N\geq 1} \mathbb{P}\left\{\sup_{n\leq N} N^{-1/p} Z^{(n)} > R\right\} \leq \delta \Rightarrow \|Z\|_{p,\infty} \leq C_p(\delta, R). \tag{8.82}$$

Proof. Assume $\mathbb{P}\{N^{-1/p} \sup\limits_{n\leq N} Z^{(n)} > R\} \leq \delta$ for all $N \geq 1$. By independence of $Z^{(1)}, Z^{(2)}, \ldots$ we have

$$\mathbb{P}\left\{\sup_{n\leq N} Z^{(n)} \leq R\dot{N}^{1/p}\right\} = (\mathbb{P}\{Z \leq RN^{1/p}\})^N,$$

therefore $\mathbb{P}\{Z \leq RN^{1/p}\} \geq (1-\delta)^{1/N}$ and hence

$$\mathbb{P}\{Z > RN^{1/p}\} \leq 1 - (1-\delta)^{1/N} \leq C_1(\delta, R)N^{-1}.$$

Consider $t > 0$ such that $RN^{1/p} < t \leq R(N+1)^{1/p}$. We have

$$\mathbb{P}\{Z > t\} \leq C_1(\delta, R)N^{-1} \leq C_2(\delta, R)t^{-p}.$$

Since we trivially have $\mathbb{P}\{Z > t\} \leq 1$ if $t \leq R$, we obtain as announced

$$\|Z\|_{p,\infty} \leq (\max\{R, C_2(\delta, R)\})^{1/p}. \qquad \square$$

Corollary 8.54. *For any $0 < q < p < \infty$ there is a constant $R(p, q)$ such that for any Z as in Proposition 8.53 we have*

$$\|Z\|_{p,\infty} \leq R(p,q) \sup_{N\geq 1} \|N^{-\frac{1}{p}} \sup_{n\leq N} Z^{(n)}\|_q. \tag{8.83}$$

Proof. By homogeneity we may assume $\sup_{N\geq 1} \|N^{-1/p} \sup_{n\leq N} Z^{(n)}\|_q \leq 1$. Then $\mathbb{P}\{N^{-1/p} \sup_{n\leq N} Z^{(n)} > \delta^{-1/q}\} \leq \delta$, so by Proposition 8.53 with $R = \delta^{-1/q}$ and (say) $\delta = 1/2$ we obtain (8.83). $\square$

The following Banach space valued version of the "principle" will be used several times in the sequel. Let B be an arbitrary Banach space and let $f\colon\ \Omega \to B$ be a B-valued random variable. Regardless whether these are finite, let us denote

$$\|f\|_{L_{p,\infty}(B)} = \left(\sup_{t>0} t^p \mathbb{P}\{\|f\|_B > t\}\right)^{1/p}$$

and $\|f\|_{L_p(B)} = (\mathbb{E}\|f\|_B^p)^{1/p}$. We will denote again by $f^{(1)}, f^{(2)}, \dots$ a sequence of independent copies of the variable f.

Proposition 8.55. *For any $1 \le q < p < \infty$ there is a constant $R'(p,q)$ such that any f in $L_q(B)$ with $\mathbb{E}(f) = 0$ satisfies*

$$\|f\|_{L^{p,\infty}(B)} \le R'(p,q) \sup\nolimits_{N\ge 1} N^{-1/p} \|f^{(1)} + \cdots + f^{(N)}\|_{L_q(B)}.$$

Proof. Assume $N^{-1/p}\|f^{(1)} + \cdots + f^{(N)}\|_{L_q(B)} \le 1$ for all $N \ge 1$. By Corollary 1.23 we have

$$\left\| \sup_{1\le n\le N} N^{-1/p} \|f^{(1)} + \cdots + f^{(n)}\|_B \right\|_q \le 2^{1+1/q}$$

and hence by the triangle inequality

$$\left\| \sup_{1\le n\le N} N^{-1/p} \|f^{(n)}\|_B \right\|_q \le 2^{2+1/q}.$$

Therefore we conclude by Corollary 8.54 applied to $Z(\cdot) = \|f(\cdot)\|_B$. $\square$

Chapter 9

Martingales and metric spaces

9.1 Metric characterization of super-reflexivity: Trees

This section is based on Bourgain's [86]. By general arguments (see [4]) it was known that super-reflexivity is preserved under Lipschitz isomorphism. Therefore knowing this, one would expect there should be a characterization of super-reflexive Banach spaces using only their structure as *metric spaces*. This is precisely the content of Bourgain's characterization in Theorem 9.1 below.

Definition. Let $(T_1, d_1), (T_2, d_2)$ be metric spaces. A map $F\colon\ T_1 \to T_2$ is called Lipschitz (or Lipschitzian) if there is a constant C such that

$$\forall s, t \in T_1 \qquad d_2(F(s), F(t)) \le C d_1(s, t).$$

The smallest such constant C will be denoted by $\|F\|_{\mathrm{Lip}}$, i.e.

$$\|F\|_{\mathrm{Lip}} = \sup_{s \ne t}\{d_2(F(s), F(t))/d_1(s, t)\}.$$

Definition. Let (T_n, d_n) be a sequence of metric spaces. We say that a metric space (T, d) contains $\{T_n\}$ Lipschitz uniformly if for any n there are injective Lipschitz mappings $F_n\colon\ T_n \to T$ such that

$$\sup_n\{\|F_n\|_{\mathrm{Lip}}\|F^{-1}_{n|F_n(T_n)}\|_{\mathrm{Lip}}\} < \infty.$$

In other words, there is $\lambda \ge 1$ and positive constants a_n, b_n with $a_n b_n \le \lambda$ such that for all n

$$\forall s, t \ge T_n \qquad (1/a_n) d(s, t) \le d(F_n(s), F_n(t)) \le b_n d(s, t).$$

In the latter case we say that (T, d) (or simply T) contains the sequence $\{T_n\}$ λ-uniformly.

Let $\mathcal{T}_n$ be a finite dyadic tree with $1 + 2 + \cdots + 2^n = 2^{n+1} - 1$ vertices (or nodes). We will label these points as $\tau(\varepsilon_1 \dots \varepsilon_j)$, $1 \le j \le n$, $\varepsilon_j \in \{-1, 1\}$ and we denote by τ_ϕ the "root" of the tree.

We equip $\mathcal{T}_n$ with its natural "geodesic" distance as a graph, i.e. we set

$$d(\tau(\varepsilon'_1 \dots \varepsilon'_j), \tau(\varepsilon''_1 \dots \varepsilon''_k)) = j + k - N$$

where $N = N(\varepsilon', \varepsilon'')$ is the largest N such that $(\varepsilon'_1, \dots, \varepsilon'_N) = (\varepsilon''_1, \dots, \varepsilon''_N)$.

Theorem 9.1. *A Banach space B is super-reflexive iff B does not contain the sequence $\{\mathcal{T}_n\}$ Lipschitz uniformly.*

The if part follows from:

Lemma 9.2. *If B is not super-reflexive then B contains the sequence $\{\mathcal{T}_n\}$ Lipschitz uniformly.*

Proof. By Theorem 3.10 if B is not super-reflexive, for any $0 < \theta < 1$ and any $n \ge 1$ there are $x_0, \dots, x_{2^n}$ in B such that for any scalars α_j we have

$$(\theta/2) \sup\nolimits_j \left(\left| \sum\nolimits_{i<j} \alpha_i \right| + \left| \sum\nolimits_{i \ge j} \alpha_i \right| \right) \le \left\| \sum \alpha_j x_j \right\| \le \sum |\alpha_j|. \tag{9.1}$$

There is a natural partial order on $\mathcal{T}_n$: we say that $s < t$ ($s, t \in \mathcal{T}_n$) if s, t lie on the same branch with s closer to the root. This can also be reformulated by saying s, t are of the form $s = \tau(\varepsilon_1 \dots \varepsilon_j)$ and $t = \tau(\varepsilon_1 \dots \varepsilon_k)$ for some $k > j$ and $(\varepsilon_1, \dots, \varepsilon_k) \in \{-1, 1\}^k$.

We write $s \le t$ if either $s < t$ or $s = t$. Note that there is a bijective mapping $\varphi\colon\ \mathcal{T}_n \to [1, \dots, 2^{n+1} - 1]$ such that φ maps disjoint intervals starting at different levels in $\mathcal{T}_n$ to disjoint ones in $[1, \dots, 2^{n+1} - 1]$. The existence of φ can be proved either by looking at a picture of a tree or using the expansion of numbers in "base 2": Just set $\psi(\tau_\phi) = 0$ and $\psi(\tau(\varepsilon_1 \dots \varepsilon_k)) = \sum_1^k 2^{-j} \varepsilon_j$ and, to obtain φ, just relabel the range of ψ in increasing order as $[1, \dots, 2^{n+1} - 1]$. We can then define an "embedding" $F_n\colon\ \mathcal{T}_n \to B$ by setting

$$\forall t \in \mathcal{T}_n \qquad F_n(t) = \sum_{w \le t} x_{\varphi(w)}.$$

We claim that for all s, t

$$(\theta/2) d(s, t) \le \|F_n(s) - F_n(t)\| \le d(s, t). \tag{9.2}$$

Indeed, assume $d(s, t) = j + k$ with

$$s = \tau(\varepsilon_1 \dots \varepsilon_N \varepsilon'_{N+1} \dots \varepsilon_{N+j}) \text{ and } t = \tau(\varepsilon_1 \dots \varepsilon_N \varepsilon''_{N+1} \dots \varepsilon''_{N+k}),$$

with $\varepsilon'_{N+1} \ne \varepsilon''_{N+1}$. Let $r = \tau(\varepsilon_1 \dots \varepsilon_N)$. Then

$$F_n(s) - F_n(t) = \sum\nolimits_{k \in A'} x_k - \sum\nolimits_{k \in A''} x_k$$

where $A' \subset I', A'' \subset I''$ are disjoint subsets included in disjoint subintervals I', I'' of $[0, \dots, 2^n]$, with $|I'| = |\{w \mid r < w \le s\}| = j$ and $|I''| = |\{w \mid r < w \le t\}| = k$, and hence (9.1) yields (9.2). $\square$

The only if part of Theorem 9.1 will be deduced from:

Lemma 9.3. *If B is super-reflexive then there is a constant C and $q < \infty$ such that:*

(i) *For any $m > 1$ and any family $(x_0, \dots, x_m)$ in B we have*

$$\inf_{0\le j, j+2k\le m} k^{-1}\|x_j + x_{j+2k} - 2x_{j+k}\| \le C(\mathrm{Log}\ m)^{-\frac{1}{q}} \sup_{1\le j\le m} \|x_j - x_{j-1}\|. \tag{9.3}$$

(ii) *For any $n > 1$ and any $F\colon\ \mathcal{T}_n \to B$ we have*

$$(9.4)\quad \inf_{2k\le N+k\le n} k^{-1}\mathbb{E}\|F(\tau(\varepsilon_1\dots\varepsilon_N\varepsilon'_{N+1}\dots\varepsilon'_{N+k})) - F(\tau(\varepsilon_1\dots\varepsilon_N\varepsilon''_{N+1}\dots\varepsilon''_{N+k}))\|$$
$$\le C(\mathrm{Log}(n))^{-1/q}\|F\|_{\mathrm{Lip}}, \tag{9.5}$$

where the expectation sign denotes the (triple) average with respect to $\varepsilon, \varepsilon', \varepsilon''$ in $\{-1,1\}^{\mathbb{N}}$.

Proof. If B (and hence $L_2(B)$) is super-reflexive, we know (see (3.42)) that there is $2 \le q < \infty$ and C such that for all B-valued dyadic martingales (f_k) we have

$$\left(\sum\nolimits_1^n \|df_k\|^q_{L_2(B)}\right)^{1/q} \le C\|f_n\|_{L_2(B)}.$$

A fortiori we have

$$\inf\nolimits_{1\le k\le n} \|df_k\|_{L_1(B)} \le Cn^{-1/q}\|f_n\|_{L_\infty(B)}. \tag{9.6}$$

Let $(\mathcal{A}_k)_{k\ge0}$ denote the dyadic filtration on $[0,1[$. Recall that $\mathcal{A}_n$ is generated by the 2^n atoms $I_n(k) = [(k-1)2^{-n}, k2^{-n}[$. Let $m = 2^n + 1$ and let $f_n\colon\ [0,1] \to B$ be the $\mathcal{A}_n$-measurable function equal to

$$x_k - x_{k-1} \quad \text{on} \quad I_n(k).$$

It is easy to check that all the values of the increments $df_k(\omega) = f_k(\omega) - f_{k-1}(\omega)$ are of the form $(2k)^{-1}(x_j + x_{j+2k} - 2x_{j+k})$ for some $0 \le j, j+2k \le m$. Thus, from (9.6), we obtain (9.3) for m of the form $m = 2^n + 1$. For the general case, just choose n such that $2^n + 1 \le m < 2^{n+1} + 1$ and note that $\mathrm{Log}\ m \simeq \mathrm{Log}(2^n + 1) \simeq n$. This completes the proof of (i).

To prove (ii), we apply (i) to $L_2(B)$ in place of B. Set (for $j = 1, \cdots, n$)

$$x_j = F(\tau(\varepsilon_1\dots\varepsilon_j))$$

and we view x_j as an $\mathcal{A}_j$-measurable function of $(\varepsilon_1, \varepsilon_2, \dots)$ in $L_2(\{-1,1\}^{\mathbb{N}}, B)$. Let $N = j + k$. Let us denote $\xi = (\varepsilon_1\dots\varepsilon_N)$, $\eta' = (\varepsilon'_{N+1}\dots\varepsilon'_{N+k})$, and

$\eta'' = (\varepsilon''_{N+1} \dots \varepsilon''_{N+k})$. Note that x_{j+k} and x_j (and hence $x_j - 2x_{j+k}$) both depend only on $\xi = (\varepsilon_1 \dots \varepsilon_N)$ so that by the triangle inequality

$$\begin{aligned}
&\|x_{N+k}(\xi,\eta') - x_{N+k}(\xi,\eta'')\|_{L_2(B)} \\
&\le \|x_{N+k}(\xi,\eta') + x_j - 2x_{j+k}\|_{L_2(B)} + \|x_{N+k}(\xi,\eta'') + x_j - 2x_{j+k}\|_{L_2(B)} \\
&= 2\|x_{N+k} + x_j - 2x_{j+k}\|_{L_2(B)}.
\end{aligned}$$

Note that $\|x_j - x_{j-1}\|_{L_2(B)} \le \|x_j - x_{j-1}\|_{L_\infty(B)} \le \|F\|_{\mathrm{Lip}}$, and also that the condition $2k \le N+k \le n$ is equivalent to $0 \le j, j+2k \le n$. Therefore a fortiori we obtain (ii) from (9.3) applied to $L_2(B)$, but this time with $m = n$. □

Remark 9.4. Let $F\colon\ \mathcal{T}_n \to B$ be an injective map. If B satisfies (ii) in Lemma 9.3, for some constant $C' > 0$ independent of n we have

$$\|F\|_{\mathrm{Lip}} \|F^{-1}_{|F(\mathcal{T}_n)}\|_{\mathrm{Lip}} \ge C'(\mathrm{Log}\, n)^{1/q}. \tag{9.7}$$

Indeed, we may assume that $\|F^{-1}_{|F(\mathcal{T}_n)}\|_{\mathrm{Lip}} = 1$. Then we have

$$d(F(s), F(t)) \ge d(s,t)$$

for all s,t. From this it is easy to check that $\mathcal{T}_n$ satisfies

$$k^{-1}\mathbb{E} d(\tau(\varepsilon_1 \dots \varepsilon_N \varepsilon'_{N+1} \dots \varepsilon'_{N+k}), \tau(\varepsilon_1 \dots \varepsilon_N \varepsilon''_{N+1} \dots \varepsilon''_{N+k})) \ge 1$$

(because $\varepsilon'_{N+1} \ne \varepsilon''_{N+1}$ implies that the above distance is equal to $2k$ and this event occurs with probability $1/2$). Therefore (9.4) immediately implies (9.7).□

Proof of Theorem 9.1. The if part follows from Lemma 9.2 and the converse from the preceding remark. □

Remark 9.5. Bourgain observed in [86] that already in Hilbert space the estimate of (9.7) is sharp. See also Matoušek's [201] for more on this.

9.2 Another metric characterization of super-reflexivity: Diamonds

This section is based on [173]. Here the sequence $\{\mathcal{T}_n\}$ is replaced by the sequence $\{\Delta_n\}$ of the diamond graphs defined as follows. It will be convenient to view Δ_n as embedded in the Hamming cube, i.e. the set $\{0,1\}^{2^n}$ equipped with the Hamming distance $d(s,t) = \sum_1^{2^n} |s_j \dots t_j| = |\{j \mid s_j \ne t_j\}|$. The embedding is realized by induction as follows. We set $\Delta_0 = \{0,1\}$. Then assuming $\Delta_{n-1} \subset \{0,1\}^{2^{n-1}}$, we define $\Delta'_n \subset \{0,1\}^{2^n}$ simply by doubling each element in Δ_{n-1}, i.e. we set $\Delta'_n = \{(t,t) \mid t \in \Delta_{n-1}\} \subset \{0,1\}^{2^n}$. Then if $(s,t) \in \Delta_{n-1}$ is a pair of neighbours, i.e. s,t differ just by one digit say $s_j \ne t_j$ then there are exactly two points s' and t' in $\{0,1\}^{2^n}$ such that $d(s',(s,s)) = d(s',(t,t)) = 1$ and similarly $d(t',(s,s)) = d(t',(t,t)) = 1$. We let Δ''_n denote the collection of all the points s', t' obtained in this way and we then define $\Delta_n = \Delta'_n \cup \Delta''_n$.

Note that if $V(\Delta_n)$ (resp. $E(\Delta_n)$) denotes the set of vertices (resp. edges) in Δ_n we have by an elementary induction

$$|E(\Delta_n)| = 4^n \quad \text{and} \quad |V(\Delta_n)| = 2 + 2\sum\nolimits_0^{n-1} 4^j = 2 + 2\frac{4^n - 1}{3}. \tag{9.8}$$

Theorem 9.6. *A Banach space B is super-reflexive iff B does not contain the sequence $\{\Delta_n\}$ Lipschitz uniformly.*

The if part will follow from:

Lemma 9.7. *Fix $\theta > 0$ and $\lambda \geq 1$. Let $(x_1 \ldots, x_{2^n})$ be a λ-basic sequence in the unit ball of B such that*

$$\left\|\sum\nolimits_{j\in A} x_j\right\| \geq \theta|A| \quad \text{for any} \quad A \subset \{1, \ldots, 2^n\}.$$

Then the function $F\colon\ \Delta_n \to B$ defined (viewing Δ_n as a subset of $\{0,1\}^{2^n}$) by $F(t) = \sum_1^n t_j x_j$ satisfies

$$\forall s,t \in \Delta_n \qquad (\theta/8\lambda)d(s,t) \leq d(F(s), F(t)) \leq d(s,t).$$

For the only if part we will use:

Lemma 9.8. *Assume that B is uniformly convex. Then for any $F\colon\ \Delta_1 \to B$ such that $d(s,t) \leq \|F(s) - F(t)\| \leq Md(s,t)$ for any s,t in Δ_1 we have*

$$\|F(11) - F(00)\| \leq 2M(1 - \delta_B(2/M)).$$

Proof. Let

$$\begin{aligned} x_1 &= F(11) - F(10)\\ x_2 &= F(10) - F(00)\\ x_3 &= F(11) - F(01)\\ x_4 &= F(01) - F(00) \end{aligned}$$

so that $x_1+x_2+x_3+x_4 = 2(F(11)-F(00))$. Note $1 \leq \|x_j\| \leq M$ for $j = 1, \ldots, 4$ and also $2 \leq \|x_1 - x_3\| \leq M$ and $2 \leq \|x_2 - x_4\| \leq M$. Therefore

$$\|(x_1 + x_3)/2M\| \leq 1 - \delta_B(2/M) \quad \text{and} \quad \|(x_2 + x_4)/2M\| \leq 1 - \delta_B(2/M),$$

and hence

$$\|F(11) - F(00)\| = \|x_1 + x_3 + x_2 + x_4\|/2 \leq 2M(1 - \delta_B(2/M)). \qquad \square$$

Proof of Theorem 9.6. If B is not super-reflexive then by Remark 3.20 for any $0 < \theta < 1$, for any $\lambda > 1$ and for any n there is a sequence $x_1, \ldots, x_{2^n}$ in B satisfying the assumption of Lemma 9.7 (indeed any such finite sequence in a space f.r. in B can obviously be "copied" back in B). Thus Lemma 9.7 establishes the if part. Conversely assume B super-reflexive. Then by Theorem 3.37 we may as well assume B uniformly convex.

Let $M_n = \inf\{\|F\|_{\mathrm{Lip}}\|F^{-1}_{|F(\Delta_n)}\|_{\mathrm{Lip}}\}$ where the infimum runs over all injective $F\colon\ \Delta_n \to B$. We claim that

$$M_{n-1} \le M_n(1-\delta_B(2/M_n)). \tag{9.9}$$

Fix a number $M > M_n$. Let $F\colon\ \Delta_n \to B$ be such that $\|F^{-1}_{|F(\Delta_n)}\|_{\mathrm{Lip}} \le 1$ and $\|F\|_{\mathrm{Lip}} \le M$. We will use the observation that $\Delta_n' \subset \Delta_n$ is metrically a copy of Δ_{n-1} inside Δ_n with double distance: more precisely we have obviously $d_{\Delta_n}((s\ \ s),(t\ \ t)) = 2d_{\Delta_{n-1}}(s,t)$ for any pair $(s\ \ s),(t\ \ t)$ in Δ_n'. *Moreover*, if s and t are neighbours in Δ_{n-1} then $(s\ \ s),(t\ \ t)$ and the "new points" s',t' appearing in the definition of Δ_n'' form an isometric copy of D_1 with $(s\ \ s)(t\ \ t)$ sitting on opposite vertices. Thus for any pair of the form $(s\ \ s)(t\ \ t)$ in Δ_n' with $d_{\Delta_{n-1}}(s,t) = 1$ we must have by Lemma 9.8

$$\|F(s\ \ s) - F(t\ \ t)\| \le 2M(-\delta_B(2/M)).$$

Let $\widehat{F}(t) = F(t\ \ t)/2$. We have

$$\|\widehat{F}(s) - \widehat{F}(t)\| \le M(1-\delta_B(2/M))$$

for any pair of neighbours s,t in Δ_{n-1}. By the triangle inequality (consider a minimal path $s = t_0,\, t_1, \dots, t_N = t$ with $d(t_j, t_{j-1}) = 1$ for all j) this implies

$$\forall s,t \in \Delta_{n-1} \qquad \|\widehat{F}(s) - \widehat{F}(t)\| \le M(1-\delta_B(2/M))d_{\Delta_{n-1}}(s,t).$$

Moreover by our assumption on F, $\|\widehat{F}(s) - \widehat{F}(t)\| \ge d_{\Delta_n}((s\ \ s),(t\ \ t))/2 = d_{\Delta_{n-1}}(s,t)$. Therefore we conclude $M_{n-1} \le M(1-\delta_B(2/M))$. This proves our claim (9.9). Given this, it is easy to deduce that $M_n \to \infty$ if B is uniformly convex. Indeed, if $M_n \le M$ for all n, let $\delta = \delta_B(2/M)$, then (9.9) implies $M_n \ge M_0(1-\delta)^{-n}$ and hence $M_n \to \infty$. This completes the proof of the only if part. $\square$

9.3 Markov type p and uniform smoothness

The notion of Markov type p was introduced by K. Ball using Markov chains on (finite subsets of) the Banach space under consideration.

Let E be an arbitrary finite space and let $X_0, X_1, \dots, X_n$ be a stationary symmetric Markov chain on E with invariant probability measure μ on E. This means that $X_0, \dots, X_n$ are E-valued random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ for which there is a symmetric kernel $P\colon\ E \times E \to \mathbb{R}_+$ ("transition probability") such that for any function $f\colon\ E \to V$ with values in (say) a vector space V we have for any $0 \le k \le n$

$$\mathbb{E}^{\sigma(X_0,\dots,X_k)} f(X_n) = \int P^{n-k}(X_k,\omega) f(\omega)\, d\mu(\omega). \tag{9.10}$$

Note in particular that, since this last expression depends only on X_k, this encodes the "Markov property"

$$\mathbb{E}^{\sigma(X_0,\ldots,X_k)} f(X_n) = \mathbb{E}^{\sigma(X_k)} f(X_n).$$

The symmetry of the transition probability P implies that the chain is reversible, i.e. that $X_0, X_1, \ldots, X_n$ have the same joint distribution as $X_n, X_{n-1}, \ldots, X_0$.

Definition 9.9. A Banach space B is called of Markov type p $(1 \le p \le 2)$ if there is a constant C such that, for any n for any finite set E and any $(X_0, \ldots, X_n)$ as above we have

$$\|f(X_n) - f(X_0)\|_{L_p(B)} \le C n^{1/p} \|f(X_1) - f(X_0)\|_{L_p(B)}.$$

The smallest such C is called the Markov type p constant of B.

The next result from [212] answers a question left open by K. Ball in [70].

Theorem 9.10 ([70])**.** *Let $1 \le p \le 2$. If a Banach space B is isomorphic to a p-uniformly smooth, then B is of Markov type p.*

Proof. By our assumption on B, we know that all B-valued martingales in $L_p(B)$ satisfy (4.25). The idea of the proof is to show that $f(X_n) - f(X_0)$ can be rewritten as a sum

$$\sum d'_k + \sum d''_k + \delta$$

where (d'_k) and (d''_k) are martingale differences and we have for all $k = 1, \ldots, n$

$$\max\{\|d'_k\|_{L_p(B)}, \|d''_k\|_{L_p(B)}, \|\delta\|_{L_p(B)}\} \le 2\|f(X_1) - f(X_0)\|_{L_p(B)}.$$

Let $\mathcal{A}_k = \sigma(X_0, \ldots, X_k)$ and $\mathcal{B}_k = \sigma(X_n, \ldots, X_k)$. Let $f_n = f(X_n)$ and $\delta_k = f_k - f_{k-1}$. We have obviously

$$f_n - f_0 = \sum\nolimits_1^n \delta_k = \sum\nolimits_1^n d'_k + \sum\nolimits_1^n \mathbb{E}^{\mathcal{A}_{k-1}} \delta_k \tag{9.11}$$

where $d'_k = (\mathbb{E}^{\mathcal{A}_k} - \mathbb{E}^{\mathcal{A}_{k-1}})(\delta_k)$.

We thus obtain $f_n - f_0$ written as a sum of martingale differences $\sum d'_k$ up to another term that we will now estimate. We have

$$f_n - f_0 = \sum\nolimits_2^{n+1} \delta_{k-1} = \sum\nolimits_2^{n+1} d''_k + \sum\nolimits_2^{n+1} \mathbb{E}^{\mathcal{B}_{k-1}}(\delta_{k-1}) \tag{9.12}$$

where $d''_k = (\mathbb{E}^{\mathcal{B}_{k-2}} - \mathbb{E}^{\mathcal{B}_{k-1}})(\delta_{k-1})$. Here again (d''_k) are martingale differences.

We now claim that for any $k = 2, \ldots, n$

$$\mathbb{E}^{\mathcal{B}_{k-1}}(\delta_{k-1}) = -\mathbb{E}^{\mathcal{A}_{k-1}} \delta_k. \tag{9.13}$$

This is a simple consequence of the reversibility of the chain. Indeed, on one hand we have $\mathbb{E}^{\mathcal{A}_{k-1}} \delta_k = \mathbb{E}^{\mathcal{A}_{k-1}} f_k - f_{k-1}$ and hence by (9.10)

$$\mathbb{E}^{\mathcal{A}_{k-1}} \delta_k = \int P(X_{k-1}, t) f(t) d\mu(t) - f(X_{k-1}).$$

On the other hand, since $(X_n, \ldots, X_0)$ and $(X_0, \ldots, X_n)$ have the same distribution, we have

$$\mathbb{E}^{\mathcal{B}_{k-1}}(\delta_{k-1}) = \mathbb{E}^{X_n \ldots X_{k-1}} f_{k-1} - \mathbb{E}^{X_n \ldots X_{k-1}} f_{k-2} = f_{k-1} - \int P(X_{k-1}, t) f(t) d\mu(t),$$

and this proves our claim (9.13). Thus, adding (9.11) and (9.12) yields

$$2(f_n - f_0) = \sum\nolimits_1^n d'_k + \sum\nolimits_2^{n+1} d''_k + \mathbb{E}^{\mathcal{A}_0}\delta_1 + \mathbb{E}^{\mathcal{B}_n}\delta_n. \tag{9.14}$$

We now observe that by the triangle inequality

$$\|d'_k\|_{L_p(B)} \le 2\|\delta_k\|_{L_p(B)}$$

and since (X_k, X_{k-1}) and (X_1, X_0) have the same distribution, we have $\|\delta_k\|_{L_p(B)} = \|f_1 - f_0\|_{L_p(B)}$ for all k, so that $\|d'_k\|_{L_p(B)} \le 2\|f_1 - f_0\|_{L_p(B)}$ and similarly $\|d''_k\|_{L_p(B)} \le 2\|f_1 - f_0\|_{L_p(B)}$. Thus we obtain finally from (9.14)

$$\|f_n - f_0\|_{L_p(B)} \le \left\|\sum\nolimits_1^n d'_k\right\|_{L_p(B)} + \left\|\sum\nolimits_2^{n+1} d''_k\right\|_{L_p(B)} + 2\|f_1 - f_0\|_{L_p(B)}$$

and since by our assumption on B all B-valued martingales satisfy (4.25) we conclude that

$$\|f_n - f_0\|_{L_p(B)} \le (Cn^{1/p} + Cn^{1/p} + 2)\|f_1 - f_0\|_{L_p(B)}. \qquad \square$$

Remark. The converse to Theorem 9.10 remains an open problem: It is rather easy to show that Markov type $p > 1$ implies type p, but it is unclear whether it implies super-reflexivity.

9.4 Notes and Remarks

....

Bibliography

Books

[1] C. Ané, S. Blachère, D. Chafai, P. Fougères, Pierre, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*, (French) [Logarithmic Sobolev inequalities] Panoramas and Syntheses, 10, Société Mathématique de France, Paris, 2000, xvi+217 pp.

[2] W. Arveson, *A short course on spectral theory*, Graduate Texts in Mathematics, 209, Springer-Verlag, New York, 2002, x+135 pp.

[3] B. Beauzamy, *Introduction to Banach spaces and their geometry*, North-Holland Publishing Co., Amsterdam, 1985.

[4] Y. Benyamini and J. Lindenstrauss, *Geometric nonlinear functional analysis. Vol. 1.* American Mathematical Society, Providence, RI, 2000.

[5] J. Bergh and Löfström, *Interpolation Spaces. An Introduction*, Springer-Verlag, Berlin-New York, 1976.

[6] C. Bennett and R. Sharpley, *Interpolation of operators*, Academic Press, Inc., Boston, MA, 1988.

[7] J. Bourgain, New classes of $\mathcal{L}^p$-spaces, *Lecture Notes in Mathematics, 889*, Springer-Verlag, Berlin-New York, 1981.

[8] R. Bourgin, Geometric aspects of convex sets with the Radon–Nikodým property, *Lecture Notes in Mathematics, 993*, Springer-Verlag, Berlin, 1983.

[9] L. Breiman, *Probability*, Corrected reprint of the 1968 original, *Classics in Applied Mathematics, 7*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992, xiv+421 pp.

[10] Yu. A.Brudnyi and N. Ya. Krugljak, *Interpolation functors and interpolation spaces*, Vol. I. North-Holland Publishing Co., Amsterdam, 1991.

[11] M. Bruneau, Variation totale d'une fonction. *Lecture Notes in Mathematics, 413*, Springer-Verlag, Berlin-New York, 1974.

[12] G. Choquet, *Lectures on Analysis, Vol. II: Representation Theory*, Benjamin, Inc., New York-Amsterdam, 1969.

[13] I. Cuculescu and A.G. Oprea, *Non-commutative Probability*, Kluwer, 1994.

[14] R. Deville, G. Godefroy V. Zizler, *Smoothness and renormings in Banach spaces*, Pitman Monographs and Surveys in Pure and Applied Mathematics, 64. John Wiley & Sons, Inc., New York, 1993.

[15] C. Dellacherie and P.A. Meyer, *Probabilities and potential. B. Theory of martingales* North-Holland Publishing Co., Amsterdam, 1982.

[16] J. Diestel and J.J. Uhl, Jr., Vector measures, *Mathematical Surveys, No. 15*, American Mathematical Society, Providence, R.I., 1977, xiii+322 pp.

[17] J.L. Doob, *Stochastic Processes*, Reprint of the 1953 original, Wiley Classics Library, Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1990, viii+654 pp.

[18] R. Dudley, and R. Norvai?a, *Differentiability of six operators on nonsmooth functions and p-variation* With the collaboration of Jinghua Qian. Lecture Notes in Mathematics, 1703. Springer-Verlag, Berlin, 1999.

[19] P. Duren,

[20] R. Durrett, Brownian motion and martingales in analysis, *Wadsworth Mathematics Series*, Wadsworth International Group, Belmont, CA, 1984.

[21]

[22] R.E. Edwards and G.I. Gaudry, *Littlewood–Paley and Multiplier Theory*, Springer-Verlag, Berlin-New York, 1977.

[23] E. Effros and Z.J. Ruan. Operator Spaces. Oxford Univ. Press, Oxford, 2000.

[24] J. Garnett, *Bounded Analytic Functions*, Academic Press, 1981.

[25] A.M. Garsia, *Martingale inequalities: Seminar notes on recent progress*, Mathematics Lecture Notes Series, W.A. Benjamin, Inc., Reading, Massachusetts-London-Amsterdam, 1973, viii+184 pp.

[26] I. Gikhman and A. Skorokhod, *The theory of stochastic processes. III*, Reprint of the 1974 edition. Springer, Berlin, 2007.

[27] P. Hàjek, S. Montesinos Santaluca, J. Vanderwerff and V. Zizler, *Biorthogonal systems in Banach spaces*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 26. Springer, New York, 2008.

[28] H. Helson,

[29] K. Hoffman,

[30] R. Kadison and J. Ringrose, *Fundamentals of the Theory of Operator Algebras, Vol. II, Advanced Theory*, Academic Press, New-York, 1986.

[31] J.P. Kahane, *Some Random Series of Functions*, Second edition, Cambridge Studies in Advanced Mathematics, 5, Cambridge University Press, Cambridge, 1985.

[32] B. Kashin and A. Saakyan, *Orthogonal series.* Translations of Mathematical Monographs, 75. American Mathematical Society, Providence, RI, 1989.

[33] T. Kato, *Perturbation Theory for Linear Operators*, Reprint of the 1980 edition, *Classics in Mathematics*, Springer-Verlag, Berlin, 1995, xxii+619 pp., University Press, Cambridge-New York, 1985.

[34] Y. Katznelson,

[35] S.G. Krein, Yu. Petunin, and E.M. Semenov, *Interpolation of linear operators*, Translations of Mathematical Monographs, 54, American Mathematical Society, Providence, R.I., 1982, xii+375 pp.

[36] M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs, 89. American Mathematical Society, Providence, RI, 2001.

[37] M. Ledoux and M. Talagrand, *Probability in Banach Spaces. Isoperimetry and Processes*, Springer-Verlag, Berlin, 1991.

[38] D. Li and H. Queffélec, Introduction l'étude des espaces de Banach. Analyse et probabilités. Société Mathématique de France, Paris, 2004.

[39] R. Long, *Martingale Spaces and Inequalities*, Peking University Press, Beijing, Friedr. Vieweg & Sohn, Braunschweig, 1993.

[40] M. Métivier, *Semimartingales. A course on stochastic processes*, Walter de Gruyter & Co., Berlin-New York, 1982.

[41] M. Métivier and J. Pellaumail, *Stochastic integration*, Probability and Mathematical Statistics. Academic Press, New York-London-Toronto, Ont., 1980.

[42] P.A. Meyer, *Probabilités et potentiel*, Hermann, Paris 1966

[43] P.A. Meyer, *Un cours sur les intgrales stochastiques*, Sminaire de Probabilits, X , pp. 245–400. Lecture Notes in Math., Vol. 511, Springer, Berlin, 1976.

[44] P.A. Meyer, *Quantum Probability for Probabilists*, Springer, LN 1538.

[45] H.P. McKean, Jr., *Stochastic Integrals*, Academic Press, New York-London, 1969.

[46] V. Milman and G. Schechtman, *Asymptotic theory of finite-dimensional normed spaces*, With an appendix by M. Gromov. Lecture Notes in Mathematics, 1200.

[47] P.F.X. Müller, Isomorphisms between H^1 spaces, *Monografie Matematyczne (New Series) 66*, Birkhäuser Verlag, Basel, 2005.

[48] J. Neveu, *Discrete-Parameter Martingales*, Translated from the French by T.P. Speed, Revised edition, North-Holland Mathematical Library, Vol. 10, North-Holland Publishing Co., Amsterdam-Oxford, American Elsevier Publishing Co., Inc., New York, 1975, viii+236 pp.

[49] *K.R. Parthasarathy*, An introduction to quantum stochastic calculus. Monographs in Mathematics, 85. Birkhuser Verlag, Basel, 1992.

[50] R.R. Phelps, *Convex functions, monotone operators and differentiability*, Second edition. Lecture Notes in Mathematics, 1364. Springer-Verlag, Berlin, 1993.

[51] G. Pisier, *Introduction to operator space theory*, London Mathematical Society Lecture Note Series, 294. Cambridge University Press, Cambridge, 2003.

[52] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion, Third edition*, Springer-Verlag, Berlin, 1999, xiv+602 pp.

[53] A. W. Roberts, D. E. Varberg, *Convex functions*, Pure and Applied Mathematics, Vol. 57. Academic Press

[54] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes, and Martingales, Vol. 2. Itô Calculus*, Cambridge University Press, Cambridge, 2000.

[55] W. Rudin, *Fourier Analysis on Groups*, Interscience, New York-London, 1962.

[56] E.M. Stein, Topics in harmonic analysis related to the Littlewood–Paley theory, *Annals of Mathematics Studies, No. 63*, Princeton University Press.

[57] E. Stein, *Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton University Press, Princeton, NJ, 1993.

[58] E. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton University Press, Princeton, N.J., 1971.

[59] D.W. Stroock, *Probability Theory, An Analytic View*, Cambridge University Press, Cambridge, 1993, xvi+512 pp.

[60] M. Takesaki, *Theory of Operator Algebras I*, Springer-Verlag, New-York, 1979.

[61] A.E. Taylor, *Introduction to Functional Analysis*, John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London 1958.

[62] A.E. Taylor and D.C. Lay, *Introduction to Functional Analysis, Second edition*, John Wiley & Sons, New York-Chichester-Brisbane, 1980.

[63] D. Voiculescu, K. Dykema, and A. Nica, *Free random variables*, A non-commutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups, CRM Monograph Series, 1, American Mathematical Society, Providence, RI, 1992, vi+70 pp.

[64] P. Wojtaszczyk, *A mathematical introduction to wavelets*, London Mathematical Society Student Texts, 37. Cambridge University Press, Cambridge, 1997.

[65] A. Zygmund, *Trigonometric series* Vol. I, II. Third edition. Cambridge University Press, Cambridge, 2002.

Articles

[66] V. Aurich, Bounded holomorphic embeddings of the unit disk into Banach spaces. *Manuscripta Math.* 45 (1983), 61–67.

[67] V. Aurich, Bounded analytic sets in Banach spaces. *Ann. Inst. Fourier* (Grenoble) **36** (1986), 229–243.

[68] K. Azuma, Weighted sums of certain dependent random variables. *Thoku Math. J.* **19** (1967) 357–367.

[69] D. Bakry, L'hypercontractivité et son utilisation en théorie des semi-groupes, *Lectures on Probability Theory (Saint-Flour, 1992), 1–114*, Lecture Notes in Math., 1581, Springer, Berlin, 1994.

[70] K. Ball, Markov chains, Riesz transforms and Lipschitz maps. Geom. *Funct. Anal.* **2** (1992), 137172.

[71] K. Ball, E. Carlen, and E.H. Lieb, Sharp uniform convexity and smoothness inequalities for trace norms, *Invent. Math.* **115** (1994), 463–482.

[72] W. Beckner, Inequalities in Fourier analysis, *Ann. Math.* **102** (1975), 159–182.

[73] A. Beck, A convexity condition in Banach spaces and the strong law of large numbers. *Proc. Amer. Math. Soc.* **13** (1962) 329–334.

[74] A. Benedek, A. Calderón, and R. Panzone, Convolution operators on Banach space valued functions, *Proc. Nat. Acad. Sci. USA* **48** (1962), 356–365.

[75] J. Bergh and J. Peetre, On the spaces V_p $(0 < p \leq \infty)$. *Boll. Un. Mat. Ital.***10** (1974), 632–648.

[76] E. Berkson, T.A. Gillespie, and P.S. Muhly, Abstract spectral decompositions guaranteed by the Hilbert transform, *Proc. London Math. Soc.* **53** (1986), 489–517.

[77] P. Biane and R. Speicher, Stochastic calculus with respect to free Brownian motion and analysis on Wigner space, *Probab. Theory Related Fields* **112** (1998), 373–409.

[78] O. Blasco, Hardy spaces of vector-valued functions: duality, *Trans. Amer. Math. Soc.* **308** (1988), 495–507.

[79] O. Blasco and A. Pełczyński, Theorems of Hardy and Paley for vector-valued analytic functions and related classes of Banach spaces, *Trans. Amer. Math. Soc.* **323** (1991), no. 1, 335–367.

[80] A. Bonami, Étude des coefficients de Fourier des fonctions de $L^p(G)$, *(French) Ann. Inst. Fourier (Grenoble)* **20** (1970), 335–402; (1971).

[81] J. Bourgain, La propriété de Radon-Nikodym. Publications mathématiques de l'Université Pierre et Marie Curie No. 36 (1979)

[82] J. Bourgain, A nondentable set without the tree property, *Studia Math.* **68** (1980), 131–139.

[83] J. Bourgain, Vector valued singular integrals and the H^1-BMO duality, *Probability Theory and Harmonic Analysis*, Chao–Woyczynski (ed.), Dekker, New York, 1986, pp. 1–19.

[84] J. Bourgain, Some remarks on Banach spaces in which martingale difference sequences are unconditional, *Ark. Mat.* **21** (1983), 163–168.

[85] J. Bourgain, Extension of a result of Benedek, Calderón and Panzone, *Ark. Mat.* **22** (1984), 91–95.

[86] J. Bourgain, The metrical interpretation of superreflexivity in Banach spaces. *Israel J. Math.* **56** (1986), 222–230.

[87] J. Bourgain, Pointwise ergodic theorems for arithmetic sets (Appendix: The return time theorem), *Publications Mathématiques de l'Institut des Hautes Études Scientiques* **69** (1989), 545.

[88] J. Bourgain and W.J. Davis, Martingale transforms and complex uniform convexity, *Trans. Amer. Math. Soc.* **294** (1986), 501–515.

[89] J. Bourgain and H.P. Rosenthal, Martingales valued in certain subspaces of L_1, *Israel J. Math.* **37** (1980), 54–75.

[90] J. Bourgain and H.P. Rosenthal, Applications of the theory of semi-embeddings to Banach space theory, *J. Funct. Anal.* **52** (1983), no. 2, 149–188.

[91] J. Bourgain, H.P. Rosenthal, and G. Schechtman, An ordinal L^p-index for Banach spaces, with application to complemented subspaces of L^p, *Ann. of Math.* **114** (1981), 193–228.

[92] J. Bretagnolle, p-variation de fonctions aléatoires. I and II. Séminaire de Probabilités, VI , pp. 51–71. *Lecture Notes in Math.* Vol. 258, Springer, Berlin, 1972.

[93] S. Bu, Deux remarques sur la propriété de Radon–Nikodym analytique, *Annales de la Fac. des Sci. de Toulouse* **60** (1990), no. 2, 79–89.

[94] S. Bu, Boundary behaviour of harmonic functions with values in Banach spaces, preprint.

[95] S. Bu, The existence of radial limits of analytic functions with values in Banach spaces, *Chinese Ann. Math. Ser. B* **22(4)** (2001), 513–518.

[96] S. Bu and B. Khaoulani, Une caractérization de la propriété de Radon–Nikodym analytique pour les espaces isomorphes à leur carrés, *Math. Ann.* **288** (1990), 345–360.

[97] S. Bu and C. Le Merdy, H_p-maximal regularity and operator valued multipliers on Hardy spaces, (2006), preprint.

[98] S. Bu and W. Schachermayer, Approximation of Jensen measures by image measures under holomorphic functions and applications, *Trans. Amer. Math. Soc.* **331** (1992), no. 2, 585–608.

[99] A. V. Bukhvalov and A. A. Danilevich, Boundary properties of analytic and harmonic functions with values in a Banach space. (Russian) Mat. Zametki 31 (1982), no. 2, 203–214, 317.

[100] D.L. Burkholder, Maximal inequalities as necessary conditions for almost everywhere convergence, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **3** (1964) 75–88.

[101] D.L. Burkholder, Distribution function inequalities for martingales, *Ann. Probab.* **1** (1973), 19–42.

[102] D.L. Burkholder, Brownian motion and the Hardy spaces H^p, *Aspects of Contemporary Complex Analysis*, (Proc. NATO Adv. Study Inst., Univ. Durham, Durham, 1979), Academic Press, London-New York, 1980, pp. 97–118.

[103] D.L. Burkholder, A geometrical characterization of Banach spaces in which martingale difference sequences are unconditional, *Ann. Probab.* **9** (1981), 997–1011.

[104] D.L. Burkholder, Martingales and Fourier analysis in Banach spaces, *Probability and Analysis (Varenna, 1985), 61–108*, Lecture Notes in Math., 1206, Springer, Berlin, 1986.

[105] D.L. Burkholder, Explorations in martingale theory and its applications, *École d'Été de Probabilités de Saint-Flour XIX—1989, 1–66*, Lecture Notes in Math., 1464, Springer, Berlin, 1991.

[106] D.L. Burkholder, Martingales and singular integrals in Banach spaces, *Handbook of the Geometry of Banach Spaces, Vol. I*, North-Holland, Amsterdam, 2001, pp. 233–269.

[107] D.L. Burkholder, A geometric condition that implies the existence of certain singular integrals of Banach space-valued functions, *Conference on Harmonic Analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)*, pp. 270–286, Wadsworth, Belmont, CA, 1983.

[108] D.L. Burkholder and R.F. Gundy, Extrapolation and interpolation of quasi-linear operators on martingales, *Acta Math.* **124** (1970), 249–304.

[109] D.L. Burkholder, R.F. Gundy, and M.L. Silverstein, A maximal function characterization of the class H^p, *Trans. Amer. Math. Soc.* **157** (1971), 137–153.

[110] A. Brunel and L. Sucheston, On J-convexity and some ergodic super-properties of Banach spaces. *Trans. Amer. Math. Soc.* **204** (1975), 79–90.

[111] A.P. Calderón, Intermediate spaces and interpolation, the complex method, *Studia Math.* **24** (1964), 113–190.

[112] S.D. Chatterji, Martingale convergence and the Radon-Nikodym theorem in Banach spaces. Math. Scand. 22 (1968) 21–41.

[113] C-H. Chu, A note on scattered C^*-algebras and the Radon Nikodym property. *J. London Math. Soc.* **24** (1981), 533–536.

[114] R.R. Coifman, M. Cwikel, R. Rochberg, Y. Sagher, and G. Weiss, Complex interpolation for families of Banach spaces, *Harmonic Analysis in Euclidean Spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 2*, pp. 269–282, Proc. Sympos. Pure Math., XXXV, Part, Amer. Math. Soc., Providence, R.I., 1979.

[115] R.R. Coifman, R. Rochberg, G. Weiss, M. Cwikel, and Y. Sagher, The complex method for interpolation of operators acting on families of Banach spaces, *Euclidean Harmonic Analysis (Proc. Sem., Univ. Maryland, College Park, Md., 1979)*, pp. 123–153, Lecture Notes in Math., 779, Springer, Berlin, 1980.

[116] R.R. Coifman, M. Cwikel, R. Rochberg, Y. Sagher, and G. Weiss, A theory of complex interpolation for families of Banach spaces, *Adv. in Math.* **43** (1982), no. 3, 203–229.

[117] R.R. Coifman and S. Semmes, Interpolation of Banach spaces, Perron processes, and Yang–Mills, *Amer. J. Math.* **115** (1993), no. 2, 243–278.

[118] A. Connes, Classification of injective factors. Cases II_1, II_∞, III_λ, $\lambda \neq 1$, *Ann. of Math. (2)* **104** (1976), no. 1, 73–115.

[119] M. Cwikel, On $(L^{p_0}(A_0), L^{p_1}(A_1))_{\theta,q}$. *Proc. Amer. Math. Soc.* 44 (1974), 286–292.

[120] M. Cwikel, Complex interpolation spaces, a discrete definition and reiterationm, *Indiana Univ. Math. J.* **27** (1978), no. 6, 1005–1009.

[121] M. Cwikel and S. Janson, Interpolation of analytic families of operators, *Studia Math.* **79** (1984), no. 1, 61–71.

[122] M. Cwikel and S. Janson, Real and complex interpolation methods for finite and infinite families of Banach spaces, *Adv. in Math.* **66** (1987), no. 3, 234–290.

[123] B. Davis. On the integrability of the martingale square function. *Israel J. Math.* **8** (1970) 187–190.

[124] W. J. Davis, W. B. Johnson and J. Lindenstrauss, The ℓ_1^n problem and degrees of non-reflexivity. *Studia Math.* **55** (1976), 123–139.

[125] W. J. Davis and J. Lindenstrauss, The ℓ_1^n problem and degrees of non-reflexivity. II. *Studia Math.* **58** (1976), 179–196.

[126] M.M. Day, Uniform convexity in factor and conjugate spaces. *Ann. of Math.* **45**, (1944) 375–385.

[127] C. Demeter, M. Lacey, T. Tao and C. Thiele, Breaking the duality in the return times theorem. *Duke Math. J.* **143** (2008), 281355.

[128] G. Edgar, A non-compact Choquet theorem, *Proc. Amer. Math. Soc.* **49** (1975), 354–358.

[129] G. Edgar, Analytic martingale convergence. *J. Funct. Anal.* **69** (1986), 268–280.

[130] G. Edgar and R. F. Wheeler, Topological properties of Banach spaces. Pacic J. Math. 115 (1984), 317350.

[131] P. Enflo, Banach spaces which can be given an equivalent uniformly convex norm. *Israel J. Math.* **13** (1972), 281–288.

[132] T. Fack and H. Kosaki, Generalized s-numbers of τ-measurable operators, *Pacific J. Math.* **123** (1986), 269–300.

[133] T. Figiel, On the moduli of convexity and smoothness, *Studia Math.* **56** (1976), no. 2, 121–155.

[134] T. Figiel, Uniformly convex norms on Banach lattices, *Studia Math.* **68** (1980), 215–247.

[135] T. Figiel, J. Lindenstrauss and V. Milman, The dimension of almost spherical sections of convex bodies. *Acta Math.* **139** (1977), no. 1-2, 53–94.

[136] T. Figiel and G. Pisier, Séries aléatoires dans les espaces uniformément convexes ou uniformément lisses, *C.R. Acad. Sci. Paris Sér. A* **279** (1974), 611–614.

[137] R. Fortet and E. Mourier, Les fonctions aléatoires comme éléments aléatoires dans les espaces de Banach. *Studia Math.* **15** (1955), 62–79.

[138] S. Geiss, A counterexample concerning the relation between decoupling constants and UMD-constants, *Trans. Amer. Math. Soc.* **351** (1999), no. 4, 1355–1375.

[139] S. Geiss, Contraction principles for vector valued martingales with respect to random variables having exponential tail with exponent $2 < \alpha < \infty$, *J. Theoret. Probab.* **14** (2001), no. 1, 39–59.

[140] N. Ghoussoub and B. Maurey, Counterexamples to several problems concerning G_δ-embeddings, *Proc. Amer. Math. Soc.* **92** (1984), no. 3, 409–412.

[141] N. Ghoussoub and B. Maurey, G_δ-embeddings in Hilbert space, *J. Funct. Anal.* **61** (1985), no. 1, 72–97.

[142] N. Ghoussoub and B. Maurey, The asymptotic-norming and the Radon–Nikodým properties are equivalent in separable Banach spaces, *Proc. Amer. Math. Soc.* **94** (1985), no. 4, 665–671.

[143] N. Ghoussoub and B. Maurey, H_δ-embedding in Hilbert space and optimization on G_δ-sets, *Mem. Amer. Math. Soc.* **62** (1986), no. 349, iv+101 pp.

[144] N. Ghoussoub and B. Maurey, G_δ-embeddings in Hilbert space. II, *J. Funct. Anal.* **78** (1988), no. 2, 271–305.

[145] N. Ghoussoub and B. Maurey, Plurisubharmonic martingales and barriers in complex quasi-Banach spaces, *Ann. Inst. Fourier (Grenoble)* **39** (1989), no. 4, 1007–1060.

[146] N. Ghoussoub, B. Maurey, and W. Schachermayer, A counterexample to a problem on points of continuity in Banach spaces, *Proc. Amer. Math. Soc.* **99** (1987), no. 2, 278–282.

[147] N. Ghoussoub, B. Maurey, and W. Schachermayer, Pluriharmonically dentable complex Banach spaces, *J. Reine Angew. Math.* **402** (1989), 76–127.

[148] N. Ghoussoub, B. Maurey, and W. Schachermayer, Geometrical implications of certain infinite-dimensional decompositions, *Trans. Amer. Math. Soc.* **317** (1990), no. 2, 541–584.

[149] N. Ghoussoub, G. Godefroy, B. Maurey, and W. Schachermayer, Some topological and geometrical structures in Banach spaces, *Mem. Amer. Math. Soc.* **70** (1987), no. 378, iv+116 pp.

[150] N. Ghoussoub, J. Lindenstrauss, and B. Maurey, Analytic martingales and plurisubharmonic barriers in complex Banach spaces, *Banach Space Theory (Iowa City, IA, 1987), 111–130*, Contemp. Math., 85, Amer. Math. Soc., Providence, RI, 1989.

[151] L. Gross, Logarithmic Sobolev inequalities, *Amer. J. Math.* **97** (1975), 1061–1083.

[152] A. Grothendieck, Réarrangements de fonctions et inégalités de convexité dans les algèbres de von Neumann munies d'une trace, *Séminaire Bourbaki, Vol. 3*, exp. no. 113, 127–139, Soc. Math. France, Paris, 1955, reprinted 1995.

[153] A. Grothendieck, La théorie de Fredholm, *Bull. Soc. Math. France* **84** (1956), 319–384.

[154] R. Gundy, A decomposition for L^1-bounded martingales. *Ann. Math. Statist.* **39** (1968) 134–138.

[155] V. I. Gurarii and N. I. Gurarii, Bases in uniformly convex and uniformly smooth Banach spaces. (Russian) *Izv. Akad. Nauk SSSR Ser. Mat.* **35** (1971) 210–215.

[156] U. Haagerup, The best constants in the Khintchine inequality, *Studia Math.* **70** (1981), 231–283 (1982).

[157] U. Haagerup and G. Pisier, Factorization of analytic functions with values in non-commutative L_1-spaces and applications, *Canad. Math. J.* **41** (1989), 882–906.

[158] O. Hanner, On the uniform convexity of L^p and l^p . *Ark. Mat.* **3** (1956), 239–244.

[159] R. E. Huff and P. D. Morris, Geometric characterizations of the Radon-Nikodym property in Banach spaces. *Studia Math.***56** (1976), 157164.

[160] K. Itô and M. Nisio, On the convergence of sums of independent Banach space valued random variables, *Osaka J. Math.* **5** (1968), 35–48.

[161] N. Jain and Monrad

[162] R.C. James, Uniformly non-square Banach spaces, *Ann. Math.* **80** (1964), 542–550.

[163] R.C. James, Some self-dual properties of normed linear spaces, *Symposium on Infinite-Dimensional Topology (Louisiana State Univ., Baton Rouge, LA, 1967)*, Ann. Math. Studies, no. 69, pp. 159–175, Princeton Univ. Press, Princeton, N.J., 1972.

[164] R.C. James, Super-reflexive Banach spaces, *Canad. J. Math.* **24** (1972), 896–904.

[165] R.C. James, Super-reflexive spaces with bases, *Pacific J. Math.* **41** (1972), 409–419.

[166] R.C. James, A nonreflexive Banach space that is uniformly nonoctahedral, *Israel J. Math.* **18** (1974), 145–155.

[167] R.C. James, Nonreflexive spaces of type 2. Israel J. Math. 30 (1978), 1–13

[168] R.C. James, R. C. James and J. Lindenstrauss, The octahedral problem for Banach spaces. Proceedings of the Seminar on Random Series, Convex Sets and Geometry of Banach Spaces (Mat. Inst., Aarhus Univ., Aarhus, 1974), pp. 100–120. Various Publ. Ser., No. 24, Mat. Inst., Aarhus Univ., Aarhus, 1975.

[169] R.C. James and J.J. Schäffer. Super-reflexivity and the girth of spheres. *Israel J. Math.* **11** (1972), 398–404.

[170] R.C. James and Sundaresan,

[171] S. Janson, On complex hypercontractivity, *J. Funct. Anal.* **151** (1997), 270–280.

[172] S. Janson, On hypercontractivity for multipliers on orthogonal polynomials, *Ark. Mat.* **21** (1983), 97–110.

[173] W. B. Johnson, G. Schechtman, Diamond graphs and super-reflexivity. Preprint, 2008.

[174] R. L. Jones, J. M. Rosenblatt and M. Wierdl, Oscillation in ergodic theory: higher dimensional results. *Israel J. Math.* **135** (2003), 1–27.

[175] M. Junge, Doob's inequality for non-commutative martingales, *J. Reine Angew. Math.* **549** (2002), 149–190.

[176] M. Junge, Fubini's theorem for ultraproducts of noncommmutative L_p-spaces. *Canad. J. Math.* **56** (2004), 983–1021.

[177] M. Junge and Q. Xu, On the best constants in some non-commutative martingale inequalities, *Bull. London Math. Soc.* **37** (2005), no. 2, 243–253.

[178] N. Kalton, S. V. Konyagin, L. Vesely. Delta-semidefinite and delta-convex quadratic forms in Banach spaces. *Positivity* **12** (2008) 221–240.

[179] N.H. Katz, Matrix valued paraproducts, *J. Fourier Anal. Appl.* **300** (1997), 913–921.

[180] S. V. Kisliakov, A remark on the space of functions of bounded p-variation. *Math. Nachr.* **119** (1984), 137–140.

[181] J.L. Krivine, Sous-espaces de dimension finie des espaces de Banach rticuls. *Ann. of Math.* **104** (1976), 1–29.

[182] J.L. Krivine and B. Maurey, Espaces de Banach stables, *Israel J. Math.* **39** (1981), no. 4, 273–295.

[183] K. Kunen and H.P. Rosenthal, Martingale proofs of some geometrical results in Banach space theory. Pacific J. Math. 100 (1982), 153–175.

[184] S. Kwapień, Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients, *Studia Math.* **44** (1972), 583–595.

[185] R. Latała and K. Oleszkiewicz, On the best constant in the Khinchin-Kahane inequality. *Studia Math.* **109** (1994), 101–104.

[186] E. Lenglart, D. Lépingle and M. Pratelli, Présentation unifiée de certaines inégalités de la théorie des martingales. With an appendix by Lenglart. *Lecture Notes in Math., 784* , Seminar on Probability, XIV (Paris, 1978/1979), pp. 26–52, Springer, Berlin, 1980.

[187] D. Lépingle, Quelques inégalités concernant les martingales, *Studia Math.* **59** (1976), 63–83.

[188] D.R. Lewis and C. Stegall, Banach spaces whose duals are isomorphic to $l_1(\Gamma)$, *J. Funct. Anal.* **12** (1973), 177–187.

[189] J. Lindenstrauss, On the modulus of smoothness and divergent series in Banach spaces. *Michigan Math. J.* **10** (1963), 241–252.

[190] J. Lindenstrauss and A. Pełczyński, Absolutely summing operators in L_p-spaces and their applications. *Studia Math.* **29** (1968), 275–326.

[191] J. Lindenstrauss and H.P. Rosenthal, The $\mathcal{L}_p$ spaces, *Israel J. Math.* **7** (1969), 325–349.

[192] V. I. Liokumovich, Existence of B-spaces with a non-convex modulus of convexity, *Izv. Vyssh. Uchebn. Zaved. Mat.* **12** (1973), 4349.

[193] J. L. Lions, Une construction despaces dinterpolation, *C. R. Acad. Sci. Paris* **251** (1960) 18531855.

[194] J. L. Lions, and J. Peetre, Sur une classe d'espaces d'interpolation. *Inst. Hautes Études Sci. Publ. Math.* No. 19 (1964) 5–68.

[195] F. Lust-Piquard, Inégalités de Khintchine daus C_p $(1 < p < \infty)$, *C.R. Acad. Sci. Paris* **303** (1986), 289–292.

[196] F. Lust-Piquard and G. Pisier, Noncommutative Khintchine and Paley inequalities, *Arkiv för Mat.* **29** (1991), 241–260.

[197] M. Manstavicius, p-variation of strong Markov processes. *Ann. Probab.* **32** (2004), 20532066.

[198] M. Marcus and G. Pisier, Random Fourier series with applications to harmonic analysis, *Ann. of Math. Stud.* **101** (1981), Princeton University Press.

[199] T. Martínez and J.L. Torrea, Operator-valued martingale transforms, *Tohoku Math. J. (2)* **52** (2000), 449–474.

[200] F

[201] J. Matoušek, On embedding trees into uniformly convex Banach spaces, *Israel J. Math.* **114** (1999), 221237.

[202] B. Maurey, Systèmes de Haar, *Séminaire Maurey–Schwartz, 74–75*, Ecole Polytechnique, Paris.

[203] B. Maurey, Théorèmes de factorisation pour les opérateurs linéaires à valeurs dans les espaces L^p, with an *Astérisque*, no. 11, Société Mathématique de France, Paris, 1974.

[204] B. Maurey, Type et cotype dans les espaces munis de structures locales inconditionnelles. (Séminaire Maurey-Schwartz 1973–1974: Espaces L^p, applications radonifiantes et géométrie des espaces de Banach, Exp. Nos. 24 et 25, 25 pp. Centre de Math., École Polytech., Paris, 1974.

[205] B. Maurey, Construction de suites symétriques. *C. R. Acad. Sci. Paris Sér. A-B* **288** (1979), no. 14, A679–A681.

[206] B. Maurey, Type, cotype and K-convexity. Handbook of the geometry of Banach spaces, Vol. 2, 1299–1332, North-Holland, Amsterdam, 2003.

[207] B. Maurey and G. Pisier, Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach, *Studia Math.* **58** (1976), 45–90.

[208] T. McConnell, On Fourier multiplier transformations of Banach-valued functions, *Trans. Amer. Math. Soc.* **285** (1984), 739–757.

[209] P.W. Millar, Path behavior of processes with stationary independent increments. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **17** (1971), 53–73.

[210] I. Monroe, On the γ-variation of processes with stationary independent increments. *Ann. Math. Statist.* **43** (1972), 1213–1220.

[211] M. Musat. On the operator space UMD property and non-commutative martingale inequalities. PhD Thesis, University of Illinois at Urbana-Champaign, 2002.

[212] A. Naor, Y. Peres, O. Schramm, S. Sheffield, Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces. Duke Math. J. 134 (2006), 165–197.

[213] F. Nazarov, G. Pisier, S. Treil, and A. Volberg, Sharp estimates in vector Carleson imbedding theorem and for vector paraproducts, *J. Reine Angew. Math.* **542** (2002), 147–171.

[214] F. Nazarov and S. Treil, The weighted norm inequalities for Hilbert transform are now trivial, *C.R. Acad. Sci. Paris Sér. I Math.* **323** (1996), 717–722.

[215] F. Nazarov and S. Treil, The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis, *St. Petersburg Math. J.* **8** (1997), 721–824

[216] F. Nazarov, S. Treil, and A. Volberg, The Bellman functions and two-weight inequalities for Haar multipliers, *J. Amer. Math. Soc.* **12** (1999), 909–928.

[217] F. Nazarov, S. Treil, and A. Volberg, Counterexample to infinite dimensional Carleson embedding theorem, *C.R. Acad. Sci. Paris Sér. I Math.* **325** (1997), 383–388.

[218] E. Nelson, The free Markoff field, *J. Funct. Anal.* **12** (1973), 211–227.

[219] E. Nelson, Notes on non-commutative integration, *J. Funct. Anal.* **15** (1974), 103–116.

[220] J. Neveu, Sur l'espérance conditionnelle par rapport à un mouvement brownien, *Ann. Inst. H. Poincaré Sect. B (N.S.)* **12** (1976), no. 2, 105–109.

[221] G. Nordlander, The modulus of convexity in normed linear spaces. *Ark. Mat.* **4** (1960) 15–17.

[222] J. Parcet and N. Randrianantoanina, Non-commutative Gundy decompositions, *Proc. London Math. Soc.*, to appear.

[223] J. Peetre,

[224] S. Petermichl, Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol, *C.R. Acad. Sci. Paris Sér. I Math.* **330** (2000), 455–460.

[225] R.R. Phelps, Dentability and extreme points in Banach spaces, *J. Funct. Anal.* **17** (1974), 78–90.

[226] J. Picard, A tree approach to p-variation and to integration. *Ann. Probab.* **36** (2008), 2235–2279.

[227] G. Pisier, Martingales with values in uniformly convex spaces, *Israel J. Math.* **20** (1975), 326–350.

[228] G. Pisier, Holomorphic semigroups and the geometry of Banach spaces. *Ann. of Math.* **115** (1982), 375–392.

[229] G. Pisier, On the duality between type and cotype. Martingale theory in harmonic analysis and Banach spaces (Cleveland, Ohio, 1981), pp. 131–144, *Lecture Notes in Math.* **939**, Springer, Berlin, 1982.

[230] G. Pisier, Probabilistic methods in the geometry of Banach spaces. Probability and analysis (Varenna, 1985), 167–241, *Lecture Notes in Math.* **1206**, Springer, Berlin, 1986.

[231] G. Pisier, The dual J^* of the James space has cotype 2 and the Gordon-Lewis property. *Math. Proc. Cambridge Philos. Soc.* **103** (1988), 323–331.

[232] G. Pisier, The K_t-functional for the interpolation couple $L_1(A_0), L_\infty(A_1)$. *J. Approx. Theory* **73** (1993), no. 1, 106–117.

[233] G. Pisier, Noncommutative vector valued L_p-spaces and completely p-summing maps, *Soc. Math. France. Astérisque* **237** (1998), vi+131 pp..

[234] G. Pisier, Remarks on the non-commutative Khintchine inequalities for $0 < p < 2$. J. Funct. Anal. 256 (2009), 4128–4161.

[235] G. Pisier and Q. Xu, Random series in the real interpolation spaces between the spaces v_p. Geometrical aspects of functional analysis (1985/86), 185–209, Lecture Notes in Math., 1267, Springer, Berlin, 1987.

[236] G. Pisier and Q. Xu, The strong p-variation of martingales and orthogonal series, *Probab. Theory Related Fields* **77** (1988), no. 4, 497–514.

[237] G. Pisier and Q. Xu, Non-commutative martingale inequalities, *Comm. Math. Phys.* **189** (1997), no. 3, 667–698.

[238] G. Pisier and Q. Xu, Non-commutative L_p-spaces, *Handbook of the Geometry of Banach Spaces, Vol. II*, North-Holland, Amsterdam, 2003.

[239] V. Pták, Biorthogonal systems and reflexivity of Banach spaces, *Czechoslovak Math. J.* **9 (84)** (1959), 319–326.

[240] N. Randrianantoanina, Non-commutative martingale transforms, *J. Funct. Anal.* **194** (2002), 181–212.

[241] N. Randrianantoanina, Square function inequalities for non-commutative martingales, *Israel J. Math.* **140** (2004), 333–365.

[242] H.P. Rosenthal, On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables, *Israel J. Math.* **8** (1970), 273–303.

[243] J.L. Rubio de Francia, Fourier series and Hilbert transforms with values in UMD Banach spaces, *Studia Math.* **81** (1985), 95–105.

[244] J.L. Rubio de Francia, Martingale and integral transforms of Banach space valued functions, *Probability and Banach Spaces (Zaragoza, 1985), 195–222*, Lecture Notes in Math., 1221, Springer, Berlin, 1986.

[245] W. Schachermayer, For a Banach space isomorphic to its square the Radon–Nikodým property and the Krein–Milman property are equivalent, *Studia Math.* **81** (1985), 329–339.

[246] W. Schachermayer, The sum of two Radon–Nikodým-sets need not be a Radon–Nikodým-set, *Proc. Amer. Math. Soc.* **95** (1985), 51–57.

[247] W. Schachermayer, Some remarks concerning the Krein–Milman and the Radon–Nikodým property of Banach spaces, *Banach Spaces (Columbia, Mo., 1984), 169–176*, Lecture Notes in Math., 1166, Springer, Berlin, 1985.

[248] W. Schachermayer, The Radon–Nikodým property and the Krein–Milman property are equivalent for strongly regular sets, *Trans. Amer. Math. Soc.* **303** (1987), 673–687.

[249] J.J. Schäffer, and K. Sundaresan, Reflexivity and the girth of spheres. *Math. Ann.* **184** 1969/1970 163–168.

[250] J. Schwartz, A remark on inequalities of Calderón-Zygmund type for vector-valued functions, *Comm. Pure Appl. Math.* **14** (1961), 785–799.

[251] T. Simon, Small ball estimates in p-variation for stable processes. *J. Theoret. Probab.* **17** (2004), no. 4, 9791002.

[252] C. Stegall, The Radon–Nikodym property in conjugate Banach spaces, *Trans. Amer. Math. Soc.* **206** (1975), 213–223.

[253] C. Stegall, The duality between Asplund spaces and spaces with the Radon-Nikod?m property. *Israel J. Math.* **29** (1978), 408–412.

[254] S. Szarek, On the best constants in the Khinchin inequality. *Studia Math.* **58** (1976), 197–208.

[255] N. Tomczak-Jaegermann, The moduli of smoothness and convexity and the Rademacher averages of trace classes $S_p(1 \leq p < \infty)$. *Studia Math.* **50** (1974), 163–182.

[256] F. Watbled, Complex interpolation of a Banach space with its dual. *Math. Scand.* **87** (2000), 200–210.