INTRODUCTION TO BANACH-VALUED ANALYSIS (V5B7, WINTER SEMESTER 2020-2021)

ALEX AMENTA

Contents

Date: September 18, 2020.

1. Introduction

The one-sentence goal of this course is to study harmonic analysis in the context of functions $f\colon \mathbb{R} \to X$ of a single variable, taking values in an infinite-dimensional Banach space X.

1.1. **Conventions.** Throughout these notes we will deal with both real and complex Banach spaces. If we do not explicitly specify 'real' or 'complex', then either choice can be made, and we will write $\mathbb{K}$ to denote the scalar field (i.e. $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$). Every complex Banach space can be seen as a real Banach space by restricting scalar multiplication to the reals. On the other hand, every real Banach space can be 'complexified', a process which doubles the dimension over $\mathbb{R}$ and does what it should: for example the complexification of $\mathbb{R}^n$ is $\mathbb{C}^n$, the complexification of $L^p(S;\mathbb{R})$ is $L^p(S;\mathbb{C})$, and so on. It's a good idea not to think too hard about this.

For an exponent $p \in [1,\infty]$ we let p' denote the *Hölder conjugate*

$$p' := \begin{cases} \frac{p}{p-1} & p \in (1,\infty) \\ \infty & p = 1 \\ 1 & p = \infty, \end{cases}$$

so that $p^{-1} + (p')^{-1} = 1$ (interpreting $1/\infty$ as 0).

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

add expository intro

2.1. **Notions of measurability.** Consider a measurable space $(S,\mathcal{A})$,[1] and a Banach space X over a scalar field $\mathbb{K}$ (either $\mathbb{R}$ or $\mathbb{C}$). The topic of these notes is the analysis of functions $f\colon S \to X$, and of operators acting on such functions. Thus it will be useful for us to gain some familiarity with the idea of X-valued functions.

The simplest kind of X-valued function arises by taking a *scalar*-valued function $f\colon S \to \mathbb{K}$ and a non-zero vector $\boldsymbol{x} \in X$, and 'placing f in the direction of $\boldsymbol{x}$'. This function is denoted by $f \otimes \boldsymbol{x}$ and formally defined by

$$f \otimes \boldsymbol{x}\colon S \to X, \qquad (f \otimes \boldsymbol{x})(s) := f(s)\boldsymbol{x} \quad \text{for all } s \in S.$$

The range of $f \otimes \boldsymbol{x}$ is contained in the linear span of $\boldsymbol{x}$, and is thus 'one-dimensional'.

The second simplest kind of X-valued function are the *simple functions*. A function $f\colon S \to X$ is *simple* if there exists a finite collection of pairwise disjoint measurable subsets $S_1,\dots,S_N \subset S$ and non-zero vectors $\boldsymbol{x}_1,\dots,\boldsymbol{x}_N \in X$ such that

$$f = \sum_{n=1}^{N} \mathbb{1}_{S_n} \otimes \boldsymbol{x}_n, \tag{2.1}$$

where $\mathbb{1}_{S_n}$ is the indicator function of S_n. We denote the vector space of simple functions $S \to X$ by $\Sigma(S;X)$ or $\Sigma_{\mathcal{A}}(S;X)$. Note that the range of f is contained in $\operatorname{span}(\boldsymbol{x}_1,\dots,\boldsymbol{x}_N)$, so f can be thought of as 'finite-dimensional'.

Of course, we need more than simple functions; measure theory tells us that the most useful class of functions are the *measurable functions*. When considering Banach-valued functions, particularly when our Banach spaces are allowed to be infinite-dimensional, there is more than one notion of measurability, and these are generally inequivalent.

Definition 2.1. Consider a function $f\colon S \to X$. We say that f is

- *measurable* if for every Borel set $B \subset X$, the preimage $f^{-1}(B)$ is measurable;

[1] i.e. S is a set and $\mathcal{A}$ is a σ-algebra of subsets of S.

- *strongly measurable* (or *Bochner measurable*) if it is the pointwise limit of simple functions; that is, if there exists a sequence $(f_n)_{n=1}^{\infty}$ in $\Sigma(S;X)$ such that $f = \lim_{n\to\infty} f_n$ pointwise on S;
- *weakly measurable* if for every functional $\boldsymbol{x}^* \in X^*$, the function $\langle f, \boldsymbol{x}^*\rangle \colon S \to \mathbb{K}$ given by $s \mapsto \langle f(s), \boldsymbol{x}^*\rangle$ is measurable.

All of these notions implicitly refer to the σ-algebra $\mathcal{A}$.

With the convenient notation

$$\begin{array}{r|l} M(S;X) & \text{Measurable } f\colon S \to X \\ SM(S;X) & \text{Strongly measurable } f\colon S \to X \\ WM(S;X) & \text{Weakly measurable } f\colon S \to X, \end{array}$$

we have the containment

$$SM(S;X) \subset M(S;X) \subset WM(S;X) \tag{2.2}$$

(Exercise 1). When X is finite-dimensional these notions coincide: the derivation of strong measurability from measurability is a standard result in measure theory,[2] and weak measurability is just a convoluted rewriting of coordinatewise measurability. But in the general context of Banach spaces the inclusions (2.2) are strict.

Example 2.2 (A measurable function which is not strongly measurable)**.** Let X be a Banach space, and consider the identity map $I\colon X \to X$, which is continuous and hence measurable. If I is strongly measurable, then there exists a sequence of simple functions $(i_n)_{n\in\mathbb{N}}$ converging pointwise to I. For each $\boldsymbol{x} \in X$ we then have

$$\boldsymbol{x} = \lim_{n\to\infty} i_n(\boldsymbol{x}),$$

so that the union $U := \cup_{n\in\mathbb{N}} i_n(X)$ is dense in X. Since each i_n is simple, U is countable, which implies that X is separable. Thus if X is not separable (e.g. if $X = L^\infty(\mathbb{R})$), the identity map $I_X \colon X \to X$ is measurable (even continuous) but not strongly measurable.

Can we do this with a σ-finite measure space? Maybe an uncountable product of probability spaces? Countable generation should be the issue.

Remark 2.3. The existence of weakly measurable functions that are not measurable is not so simple, but see [4, Example 1.4.3] for an argument which shows that the σ-algebra $\sigma(X^*)$ may be strictly smaller than the Borel σ-algebra on X. The indicator function of a Borel set which is not in $\sigma(X^*)$ is then weakly measurable, but not measurable.

It turns out that the notion of strong measurability is strongly connected with that of separability.

Theorem 2.4 (Pettis measurability theorem)**.** *Let $(S,\mathcal{A})$ be a measurable space and X a Banach space. Then a function $f\colon S \to X$ is strongly measurable if and only if it is weakly measurable and* separably valued *(i.e. there exists a separable subspace $X' \subset X$ such that $f(S) \subset X'$). In particular, if X is separable, then*

$$SM(S;X) = M(S;X) = WM(S;X).$$

Proof. First suppose that f is strongly measurable. Then f is automatically weakly measurable, and we just need to show that f is separably valued. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of simple functions converging to f pointwise. Let $X_n \subset X$ be the closure of the range of f_n. Each X_n is finite-dimensional, hence separable, and the closed subspace X' of X generated by the collection $(X_n)_{n\in\mathbb{N}}$ is also separable. Since $f_n \to f$ pointwise, the range of f is contained in X', and thus f is separably valued.

[2]See for example [3, Corollary 4.2.7] in the one-dimensional case; extending this to the finite-dimensional case can be done by summing up coordinates.

Now assume that f is weakly measurable and separably valued. Without loss of generality we may simply assume that X is separable (potentially replacing X by the closure of the range of f). Let $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ be a dense sequence in X, and for each $n \in \mathbb{N}$ define a function $\varphi_n \colon X \to \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}$ such that for all $\boldsymbol{x} \in X$,

$$\|\boldsymbol{x} - \varphi_n(\boldsymbol{x})\|_X = \min_{1\leq j\leq n} \|\boldsymbol{x} - \boldsymbol{x}_j\|_X.$$

By density of $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ in X we thus have that $\varphi_n(\boldsymbol{x}) \to \boldsymbol{x}$ for all $\boldsymbol{x} \in X$. Now define functions $f_n \colon S \to X$ by

$$f_n(s) := \varphi_n(f(s)) \qquad \forall s \in S,$$

so that $f_n \to f$ pointwise. Each f_n has finite range, so to show that f_n is simple we need only show that the preimages $f_n^{-1}(\boldsymbol{x}_k)$ are measurable. For all $1 \leq k \leq n$ We have

$$f_n^{-1}(\boldsymbol{x}_k) = \{\varphi_n(f(s)) = \boldsymbol{x}_k\} = \{\|f(s) - \boldsymbol{x}_k\|_X = \min_{1\leq j\leq n} \|f(s) - \boldsymbol{x}_j\|_X\}.$$

Let $(\boldsymbol{x}_n^*)_{n\in\mathbb{N}}$ be a norming sequence of unit vectors on X^*; then since f is weakly measurable, for each $j \in \{1, \dots, n\}$ the function

$$s \mapsto \|f(s) - \boldsymbol{x}_j\|_X = \sup_{n\in\mathbb{N}} |\langle f(s) - \boldsymbol{x}_j, \boldsymbol{x}_n^*\rangle|$$

is measurable (being the countable supremum of measurable functions). Thus the function

$$\min_{1\leq j\leq n} \|f(s) - \boldsymbol{x}_j\|_X$$

is also measurable, and the representation above shows that $f_n^{-1}(\boldsymbol{x}_k)$ is measurable (being the set on which two measurable functions are equal). Hence f_n is simple, and consequently f is strongly measurable. $\square$

As a quick application of this result, we show that strong measurability is preserved under multiplication with measurable scalar-valued functions.[3]

Corollary 2.5. *Suppose that $f \colon S \to X$ is strongly measurable and the scalar-valued function $\varphi \colon S \to \mathbb{K}$ is measurable. Then the product $\varphi f \colon S \to X$, $(\varphi f)(s) := \varphi(s) f(s)$, is strongly measurable.*

Proof. By the Pettis measurability theorem, it is equivalent to show that φf is weakly measurable and separably-valued. First we show weak measurability: for each functional $\boldsymbol{x}^* \in X^*$ write for $s \in S$

$$\langle \varphi f, \boldsymbol{x}^*\rangle(s) = \varphi(s)\langle f(s), \boldsymbol{x}^*\rangle = \varphi\langle f, \boldsymbol{x}^*\rangle.$$

Since f is strongly measurable, it is also weakly measurable, and thus the product $\varphi\langle f, \boldsymbol{x}^*\rangle$ is measurable. Since this is true for all $\boldsymbol{x}^* \in X^*$, we find that φf is weakly measurable. To show that φf is separably-valued, first note that since f is separably-valued, there exists a separable closed subspace $X_0 \subset X$ such that $f(s) \in X_0$ for all $s \in S$. But then $\varphi(s) f(s) \in X_0$ as well, so φf is separably-valued. $\square$

Before moving on to Bochner spaces we note that almost-everywhere equality of strongly measurable functions is equivalent to 'coordinatewise' almost-everywhere equality. This is a surprisingly useful observation; it is often used to deduce identities for vector-valued functions from corresponding identities for scalar-valued functions.

---

[3]This can also be proven directly via pointwise approximation with simple functions.

Lemma 2.6. *Let $(S, \mathcal{A}, \mu)$ be a measure space and X a Banach space. Suppose that $f, g\colon S \to X$ are two strongly measurable functions. Then f and g are equal almost everywhere if and only if for all functionals $\boldsymbol{x}^* \in X^*$, the scalar-valued functions $\langle f, \boldsymbol{x}^* \rangle$ and $\langle g, \boldsymbol{x}^* \rangle$ are equal almost everywhere.*

Proof. The implication

$$f \overset{\text{a.e.}}{=} g \Longrightarrow \forall \boldsymbol{x}^* \in X^* \ \langle f, \boldsymbol{x}^* \rangle \overset{\text{a.e.}}{=} \langle g, \boldsymbol{x}^* \rangle$$

is straightforward, and is even true for (non-strongly) measurable functions, so we omit the proof. The converse direction is harder because although each of the sets

$$N_{\boldsymbol{x}^*} := \{s \in S : \langle f(s), \boldsymbol{x}^* \rangle \neq \langle g(s), \boldsymbol{x}^* \rangle\} \qquad \boldsymbol{x}^* \in X^*$$

has measure zero, the (uncountable!) union of these sets over all $\boldsymbol{x}^* \in X^*$ does not *a priori* have measure zero. This is where strong measurability comes into play, via the Pettis theorem. Since f and g are separably-valued, there exists a separable closed subspace $X_0 \subset X$ such that both f and g are X_0-valued. Since X_0 is separable, there is a (countable!) sequence $(\boldsymbol{x}_n^*)_{n\in\mathbb{N}}$ in X^* which separates points of X_0.[4] Now define

$$N := \bigcup_{n\in\mathbb{N}} N_{\boldsymbol{x}_n^*};$$

this set has measure zero since it is the countable union of sets with measure zero. For all $s \notin N$ we then have $\langle f(s), \boldsymbol{x}_n^* \rangle = \langle g(s), \boldsymbol{x}_n^* \rangle$ for all $n \in \mathbb{N}$, and since $(\boldsymbol{x}_n^*)_{n\in\mathbb{N}}$ separates points of X_0, it follows that $f(s) = g(s)$. Thus $f \overset{\text{a.e.}}{=} g$. $\square$

2.2. **Bochner spaces.** Given $f\colon S \to X$, we let $\|f\|_X$ denote the scalar-valued function $S \to \mathbb{K}$ defined by $s \mapsto \|f(s)\|_X$. If f is strongly measurable, then $\|f\|_X$ is also measurable, since the function $\boldsymbol{x} \mapsto \|\boldsymbol{x}\|_X$ is continuous.

Definition 2.7. Let $(S, \mathcal{A}, \mu)$ be a measure space. For $p \in [1, \infty]$, we let $L^p(S, \mu; X)$ denote the set of *strongly* measurable functions $f \in SM(S; X)$ such that $\|f\|_X \in L^p(S, \mu)$, modulo μ-a.e. equality, and we write

$$\|f\|_{L^p(S,\mu;X)} := \|\|f\|_X\|_{L^p(S,\mu)}.$$

Each $L^p(S, \mu; X)$ is a Banach space: the proof is identical to the classic proof in the scalar-valued case.[5]

Remark 2.8. It is possible for $\|f\|_X$ to be in $L^p(S)$ without f itself being strongly measurable (or even measurable). Such a function does not qualify for membership in $L^p(S; X)$.

Proposition 2.9. *Let X be a Banach space and $p \in [1, \infty)$. Then the subspace of simple functions $\Sigma(S; X) \cap L^p(S; X)$ is dense in $L^p(S; X)$.*

Proof. Fix $f \in L^p(S; X)$. Since f is strongly measurable, there exists a sequence of simple functions $f_n \in \Sigma(S; X)$ with $\lim_{n\to\infty} f_n = f$ pointwise almost everywhere. Now set

$$g_n := \mathbb{1}_{\{s\in S : \|f_n(s)\|_X \leq 2\|f\|_X\}} f_n;$$

the functions g_n are simple and they also converge to f pointwise almost everywhere. Furthermore we have

$$\|g_n\|^p_{L^p(S;X)} = \int_{\{s\in S : \|f_n(s)\|_X \leq 2\|f\|_X\}} \|f_n(s)\|^p_X \, \mathrm{d}\mu(s) \leq 2^p \|f\|^p_{L^p(S;X)},$$

[4]That is, if $\boldsymbol{x} \neq \boldsymbol{y} \in X_0$, then there exists $n \in \mathbb{N}$ such that $\langle \boldsymbol{x}, \boldsymbol{x}_n^* \rangle \neq \langle \boldsymbol{y}, \boldsymbol{x}_n^* \rangle$. See [4, Proposition B.1.11].

[5]For revision see [3, Theorem 5.2.1].

so each g_n is in $L^p(S;X)$. Since $\|f(s) - g_n(s)\|_X \leq 3\|f(s)\|_X$ for almost all s, dominated convergence yields

$$\lim_{n\to\infty} \|f - g_n\|^p_{L^p(S;X)} = \lim_{n\to\infty} \int_S \|f(s) - g_n(s)\|^p_X \, \mathrm{d}\mu(s)$$
$$= \int_S \lim_{n\to\infty} \|f(s) - g_n(s)\|^p_X \, \mathrm{d}\mu(s) = 0,$$

so that $g_n \to f$ in $L^p(S;X)$, completing the proof. □

Remark 2.10. Proposition 2.9 can be extended to more general dense subspaces of $L^p(S)$; see Exercise 3.

Note that the case $p = \infty$ is not included in the proposition above, even though the simple functions are dense in $L^\infty(S)$.

Proposition 2.11. *Let X be a Banach space. Then the simple functions are dense in $\ell^\infty(\mathbb{N};X)$ if and only if X is finite dimensional.*[6]

Proof. First suppose X is finite dimensional, and fix $f \in \ell^\infty(\mathbb{N};X)$ and $\varepsilon > 0$. Let $C = \|f\|_{\ell^\infty(\mathbb{N};X)}$, and note that the closed ball $\overline{B_C(0)} \subset X$ is compact (this uses finite dimensionality of X). Thus there exists a finite collection of vectors $(\boldsymbol{x}_i)_{i=1}^N$ in $\overline{B_C(0)}$ such that the open balls $B_\varepsilon(\boldsymbol{x}_i)$ cover $\overline{B_C(0)}$. For each $n \in \mathbb{N}$ we thus have that $f(n) \in B_\varepsilon(\boldsymbol{x}_{i(n)})$ for some $i(n) \in \{1, \ldots, N\}$. Define a function $f_\varepsilon \colon \mathbb{N} \to X$ by

$$f_\varepsilon(n) := \boldsymbol{x}_{i(n)};$$

since the range of f_ε is finite, f_ε is simple. Furthermore since $f(n) \in B_\varepsilon(\boldsymbol{x}_{i(n)})$ for each $n \in \mathbb{N}$ we have

$$\|f - f_\varepsilon\|_{\ell^\infty(\mathbb{N};X)} = \sup_{n\in\mathbb{N}} \|f(n) - \boldsymbol{x}_{i(n)}\|_X \leq \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we have established density of the simple functions in $\ell^\infty(\mathbb{N};X)$ when X is finite dimensional.

Now we prove the converse. Aiming for a contradiction, suppose that X is infinite dimensional. Then there exists a sequence $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ of unit vectors in X such that

$$\|\boldsymbol{a}_n - \boldsymbol{a}_m\|_X > 1/2 \qquad \text{for all } n \neq m.$$

Now let $f(n) = \boldsymbol{a}_n$ for all $n \in \mathbb{N}$, so that $f \in \ell^\infty(\mathbb{N};X)$, and suppose that there exists a simple function $g \in \Sigma(\mathbb{N};X)$ with $\|f - g\|_{\ell^\infty(\mathbb{N};X)} < 1/4$. Then for all $n \neq m$ we must have

$$\|\boldsymbol{a}_n - \boldsymbol{a}_m\|_X \leq \|f(n) - g(n)\|_X + \|g(n) - g(m)\|_X + \|g(m) - f(m)\|_X$$
$$\leq \frac{1}{2} + \|g(n) - g(m)\|_X,$$

so that

$$\|g(n) - g(m)\|_X \geq \|\boldsymbol{a}_n - \boldsymbol{a}_m\|_X - \frac{1}{2} > 0.$$

It follows that g has infinite range, contradicting the assumption that g is simple. □

Now we present some elementary duality results. Recall the definition of the Hölder conjugate from Section 1.1. Given a measure space $(S, \mathcal{A}, \mu)$ and a Banach

[6] This proposition can be extended to more general measure space S in place of $\mathbb{N}$, provided S contains infinitely many disjoint measurable sets of positive measure.

space X, every function $g \in L^{p'}(S;X^*)$ induces a bounded linear functional $\Phi g \in L^p(S;X)^*$ by integration of the duality pairing between X and X^*:

$$\Phi g(f) := \int_S \langle f(s), g(s)\rangle \,\mathrm{d}\mu(s) \qquad \forall f \in L^p(S;X).$$

Hölder's inequality implies that $\|\Phi g\|_{L^p(S;X)^*} \leq \|g\|_{L^{p'}(S;X^*)}$. In the scalar case $X = \mathbb{K}$, Φ is an isometric isomorphism $L^{p'}(S) \cong L^p(S)^*$: that is, every functional $\varphi \in L^p(S)^*$ is of the form $\varphi = \Phi g$ for some $g \in L^{p'}(S)$, and furthermore $\|\varphi\|_{L^p(S)^*} = \|g\|_{L^{p'}(S)}$. We will see in Section 4 that for general Banach spaces X, Φ is an isometric isomorphism if and only if X has the *Radon–Nikodym property*. For now we will establish a duality result that holds for every Banach space.

Proposition 2.12. *Let $(S,\mathcal{A},\mu)$ be a σ-finite measure space*[7] *and X a Banach space. Then for all $1 \leq p \leq \infty$, the map $\Phi\colon L^{p'}(S;X^*) \to L^p(S;X)^*$ is an isometry onto a closed subspace of $L^p(S;X)^*$ which is norming for $L^p(S;X)$: that is, for every $f \in L^p(S;X)$,*

$$\|f\|_{L^p(S;X)} = \sup_{\substack{g \in L^{p'}(S;X^*)\\ g \neq 0}} \frac{|\Phi g(f)|}{\|g\|} = \sup_{\substack{g \in L^{p'}(S;X^*)\\ \|g\|=1}} \Big| \int_S \langle f(s), g(s)\rangle \,\mathrm{d}\mu(s)\Big|.$$

Proof. To show that Φ is an isometry, it suffices to show that $\|\Phi g\|_{L^p(S;X)^*} \geq 1$ whenever $g \in L^{p'}(S;X^*)$ with $\|g\|_{L^{p'}(S;X^*)} = 1$ (the reverse estimate has already been discussed).

Mild case: $p > 1$. In this case we have $p' < \infty$, so by density of the simple functions in $L^{p'}(S;X^*)$ and continuity of Φ we may assume that g is simple, i.e.

$$g = \sum_{n=1}^N \mathbb{1}_{S_n} \otimes \boldsymbol{x}_n^*$$

for some pairwise disjoint sets $S_n \in \mathcal{A}$ with $\mu(S_n) < \infty$ and some nonzero vectors $\boldsymbol{x}_n^* \in X^*$. Let $\varepsilon > 0$, and choose unit vectors $\boldsymbol{x}_n \in X$ (depending on ε) such that

$$\langle \boldsymbol{x}_n, \boldsymbol{x}_n^*\rangle \geq (1-\varepsilon)\|\boldsymbol{x}_n^*\|_{X^*} \qquad \forall n \in \{1,\dots,N\}.$$

Using these vectors define a test function

$$f_\varepsilon := \sum_{n=1}^N \mathbb{1}_{S_n} \otimes \|\boldsymbol{x}_n^*\|_{X^*}^{p'-1} \boldsymbol{x}_n.$$

Then we can compute

$$\begin{aligned}\|f_\varepsilon\|_{L^p(S;X)}^p &= \sum_{n=1}^N \mu(S_n)\|\boldsymbol{x}_n^*\|_{X^*}^{p(p'-1)}\|\boldsymbol{x}_n\|_X^p \\ &= \sum_{n=1}^N \mu(S_n)\|\boldsymbol{x}_n^*\|_{X^*}^{p'} = \|g\|_{L^{p'}(S;X^*)}^{p'} = 1\end{aligned}$$

as the $\boldsymbol{x}_n$ are unit vectors and $p(p'-1) = 1$. Testing Φg against f_ε yields

$$\begin{aligned}\Phi g(f_\varepsilon) = \sum_{n=1}^N \mu(S_n)\|\boldsymbol{x}_n^*\|_{X^*}^{p'-1}\langle \boldsymbol{x}_n, \boldsymbol{x}_n^*\rangle &\geq (1-\varepsilon)\sum_{n=1}^N \mu(S_n)\|\boldsymbol{x}_n^*\|_{X^*}^{p'} \\ &= (1-\varepsilon)\|g\|_{L^{p'}(S;X^*)} = 1-\varepsilon.\end{aligned}$$

This proves that $\|\Phi g\|_{L^p(S;X)^*} \geq 1-\varepsilon$. Since $\varepsilon > 0$ was arbitrary, we find that $\|\Phi g\|_{L^p(S;X)^*} \geq 1$ as intended.

[7]The σ-finiteness assumption is only needed for $p = 1$.

Spicy case: $p = 1$. Fix $\varepsilon > 0$ and define

$$A_\varepsilon := \{s \in S : \|g(s)\|_{X^*} > 1 - \varepsilon\}. \tag{2.3}$$

Then $\mu(A_\varepsilon) > 0$, but we could run into the problem that $\mu(A_\varepsilon) = \infty$. Since S is σ-finite, we can write S as an increasing union of sets of finite measure

$$S = \bigcup_{n\in\mathbb{N}} S_n, \qquad S_n \subset S_{n+1},\ \mu(S_n) < \infty \quad \forall n \in \mathbb{N},$$

and thus for sufficiently large n the set

$$A_\varepsilon^n := \{s \in S_n : \|g(s)\|_{X^*} > 1 - \varepsilon\}$$

satisfies $0 < \mu(A_\varepsilon) < \infty$.[8] Let $B_\varepsilon = A_\varepsilon^n$ for such a large n.

Since g is strongly measurable, the Pettis measurability theorem says that $g(B_\varepsilon)$ is separable, and thus there exists a sequence $(\boldsymbol{x}_k^*)_{k\in\mathbb{N}}$ in X^* such that

$$g(B_\varepsilon) \subset \bigcup_{k\in\mathbb{N}} B_\varepsilon(\boldsymbol{x}_k^*)$$

and thus

$$B_\varepsilon \subset \bigcup_{k\in\mathbb{N}} g^{-1}(B_\varepsilon(\boldsymbol{x}_k^*)).$$

Since $\mu(B_\varepsilon) > 0$, there exists a vector $\boldsymbol{x}^* \in X^*$ (i.e. one of the vectors $\boldsymbol{x}_k^*$) such that the set

$$B_{\varepsilon,\boldsymbol{x}^*} := B_\varepsilon \cap g^{-1}(B_\varepsilon(\boldsymbol{x}^*)) = \{s \in B_\varepsilon : \|g(s) - \boldsymbol{x}^*\|_{X^*} < \varepsilon\}$$

has positive measure. Picking a point $s_0 \in B_{\varepsilon,\boldsymbol{x}^*}$ and using the definition of B_ε, we see that

$$\|\boldsymbol{x}^*\|_{X^*} \geq \|g(s_0)\|_{X^*} - \|g(s_0) - \boldsymbol{x}^*\|_{X^*} > 1 - 2\varepsilon.$$

Now fix a unit vector $\boldsymbol{x} \in X$ such that $\langle \boldsymbol{x}, \boldsymbol{x}^*\rangle \geq \|\boldsymbol{x}^*\|_{X^*} - \varepsilon$, and consider the test function

$$f_\varepsilon := \mathbb{1}_{B_{\varepsilon,\boldsymbol{x}^*}} \otimes \mu(B_{\varepsilon,\boldsymbol{x}^*})^{-1}\boldsymbol{x}.$$

Then $\|f_\varepsilon\|_{L^1(S;X)} = 1$, and

$$\begin{aligned}
|\Phi g(f)| &= \Big| \fint_{B_{\varepsilon,\boldsymbol{x}^*}} \langle \boldsymbol{x}, g(s)\rangle \,\mathrm{d}\mu(s)\Big| \\
&\geq \Big| \fint_{B_{\varepsilon,\boldsymbol{x}^*}} \langle \boldsymbol{x}, \boldsymbol{x}^*\rangle \,\mathrm{d}\mu(s)\Big| - \Big| \fint_{B_{\varepsilon,\boldsymbol{x}^*}} \langle \boldsymbol{x}, g(s) - \boldsymbol{x}\rangle \,\mathrm{d}\mu(s)\Big| \\
&\geq (\|\boldsymbol{x}^*\|_{X^*} - \varepsilon) - \varepsilon \geq 1 - 4\varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we get $\|\Phi g\|_{L^1(S;X)^*} \geq 1$, as we wanted.

Norming property: Since X embeds isometrically into its double dual X^{**}, we can identify X-valued functions with X^{**}-valued functions which take values in

[8] The main issue is that A_ε^n could have zero measure, but if this were true for all n, then $\mu(A_\varepsilon) = \sup_{n\in\mathbb{N}} \mu(A_\varepsilon^n) = 0$, which is a contradiction.

X. We exploit this and the previous results to show

$$\begin{aligned}\|f\|_{L^p(S;X)} = \|f\|_{L^p(S;(X^*)^*)} &= \|\Phi f\|_{L^{p'}(S;X^*)^*} \\ &= \sup_{\substack{g\in L^{p'}(S;X^*)\\ \|g\|=1}} \int_S |\langle g(s), f(s)\rangle_{X^*,X^{**}}| \,\mathrm{d}\mu(s) \\ &= \sup_{\substack{g\in L^{p'}(S;X^*)\\ \|g\|=1}} \int_S |\langle f(s), g(s)\rangle_{X,X^*}| \,\mathrm{d}\mu(s) \\ &= \sup_{\substack{g\in L^{p'}(S;X^*)\\ g\neq 0}} \frac{|\Phi g(f)|}{\|g\|},\end{aligned}$$

completing the proof. □

2.3. **The Bochner integral.** We turn to defining integrals of vector-valued functions. As in the scalar-valued theory we start with simple functions, for which there is only one reasonable definition. Let $(S, \mathcal{A}, \mu)$ be a measure space and X a Banach space. If $f \in \Sigma(S; X)$ is a simple function represented as

$$f = \sum_{n=1}^{N} \mathbb{1}_{S_n} \otimes \boldsymbol{x}_n,$$

and if in addition $f \in L^1(S, \mu; X)$, we define the *Bochner integral*

$$\int_S f \,\mathrm{d}\mu = \int_S f(s) \,\mathrm{d}\mu(s) := \sum_{n=1}^{N} \mu(S_n)\boldsymbol{x}_n \in X.$$

Note that the assumption $f \in L^1(S, \mu; X)$ is equivalent to having $\mu(S_n) < \infty$ for all n. The Bochner integral is a linear map $\Sigma(S; X) \cap L^1(S, \mu; X) \to X$, and for f as above it satisfies

$$\Big\| \int_S f(s) \,\mathrm{d}\mu(s) \Big\|_X \leq |\mu(S_n)| \|\boldsymbol{x}_n\|_X = \|f\|_{L^1(S,\mu;X)}.$$

Thus by density of $\Sigma(S; X) \cap L^1(S, \mu; X)$ in $L^1(S, \mu; X)$ (Proposition 2.9), the Bochner integral extends to a bounded linear map $L^1(S, \mu; X) \to X$ which we continue to call the Bochner integral and denote by the same symbol. Thus the Bochner integral of $g \in L^1(S, \mu; X)$ is given by

$$\int_S g \,\mathrm{d}\mu := \lim_{n\to\infty} \int_S f_n \,\mathrm{d}\mu \in X$$

where $f_n \in \Sigma(S; X) \cap L^1(S, \mu; X)$ for all $n \in \mathbb{N}$ and $f_n \to g$ in $L^1(S, \mu; X)$.

The Bochner integral satisfies the following properties.

Proposition 2.13. *Let $(S, \mathcal{A}, \mu)$ be a measure space and X a Banach space. The Bochner integral satisfies the following properties:*

Linearity: *For $f_1, f_2 \in L^1(S, \mu; X)$ and $\lambda_1, \lambda_2 \in \mathbb{K}$,*

$$\int_S (\lambda_1 f_1 + \lambda_2 f_2)(s) \,\mathrm{d}\mu(s) = \lambda_1 \int_S f_1(s)\mathrm{d}\mu(s) + \lambda_2 \int_S f_2(s)\mathrm{d}\mu(s).$$

Commutation with linear maps: *If $f \in L^1(S, \mu; X)$ and $T \in \mathcal{L}(X, Y)$ is a bounded linear map into a Banach space Y,*

$$T\Big(\int_S f \,\mathrm{d}\mu \Big) = \int_S T f(s) \,\mathrm{d}\mu(s) \in Y$$

where $Tf \in L^1(S,\mu;Y)$ is given by $(Tf)(s) = T(f(s))$ for almost all $s \in S$. In particular, if $\boldsymbol{x}^* \in X^* = \mathcal{L}(X,\mathbb{K})$, then

$$\Big\langle \int_S f \,\mathrm{d}\mu, \boldsymbol{x}^* \Big\rangle = \int_S \langle f(s), \boldsymbol{x}^* \rangle \,\mathrm{d}\mu(s) \in \mathbb{K}.$$

Closure: *If $f \in L^1(S,\mu;X)$ and X_0 is a closed subspace of X such that $f(s) \in X_0$ for almost all $s \in S$, then $\int_S f \,\mathrm{d}\mu \in X_0$.*

Dominated convergence: *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $L^1(S,\mu;X)$, $f\colon S \to X$, and suppose that $\lim_{n\to\infty} f_n = f$ almost everywhere. Suppose that there exists a non-negative $g \in L^1(S,\mu)$ such that $\|f_n\|_X \leq g$ almost everywhere. Then $f \in L^1(S,\mu;X)$ and*

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

Substitution / Change of Variables: *Let $(T,\mathcal{B})$ be a measurable space and $\varphi\colon S \to T$ a measurable function, and let $\nu = \mu \circ \varphi^{-1}$ denote the pushforward measure. Suppose $g \in L^1(T,\nu;X)$. Then $g \circ \varphi \in L^1(S,\mu;X)$, and*

$$\int_S g \circ \varphi \,\mathrm{d}\mu = \int_T g \,\mathrm{d}\nu.$$

Proof. Linearity follows from the definition. The remaining properties are proved as follows:

Commutation with linear maps: by continuity it suffices to prove this for simple $f \in \Sigma(S;X) \cap L^1(S,\mu;X)$. Writing f as in (2.1) we have

$$T\Big(\int_S \sum_{n=1}^N \mathbb{1}_{S_n} \otimes \boldsymbol{x}_n \,\mathrm{d}\mu\Big) = \sum_{n=1}^N \mu(S_n) \otimes T(\boldsymbol{x}_n) = \int_S Tf \,\mathrm{d}\mu.$$

Closure: We may assume that X_0 is a proper subspace of X, otherwise there is nothing to show. Let $\boldsymbol{y} \in X \setminus X_0$, and by Hahn–Banach[9] choose a functional $\boldsymbol{x}^* \in X^*$ such that $\langle \boldsymbol{y}, \boldsymbol{x}^* \rangle = 1$ and $X_0 \subset \ker \boldsymbol{x}^*$. Then by the commutation property above we have

$$\Big\langle \int_S f \,\mathrm{d}\mu, \boldsymbol{x}^* \Big\rangle = \int_S \langle f(s), \boldsymbol{x}^* \rangle \,\mathrm{d}\mu(s) = 0$$

since $f(s) \in X_0$ for almost all $s \in S$. Thus $\int_S f \,\mathrm{d}\mu \neq y$. Since $y \in X \setminus X_0$ was arbitrary, we conclude that $\int_S f \,\mathrm{d}\mu \in X_0$.

Dominated convergence: By continuity of the Bochner integral it suffices to show that $f \in L^1(S,\mu;X)$ and $f_n \to f$ in $L^1(S,\mu;X)$. The first fact follows from $\|f\|_{L^1(S,\mu;X)} \leq \|g\|_{L^1(S,\mu)}$. For the second, since $\|(f_n - f)(s)\|_X \leq 2g(s)$ almost everywhere, we have

$$\lim_{n\to\infty} \int_S \|(f_n - f)(s)\|_X \,\mathrm{d}\mu(s) = 0$$

by scalar dominated convergence.

Substitution: First we need to show that $g \circ \varphi$ is strongly measurable. Let

$$g_i = \sum_{n=1}^{N_i} \mathbb{1}_{S_{n,i}} \otimes \boldsymbol{x}_{n,i}$$

[9] If you are philosophically opposed to Hahn–Banach, then see [4, Corollary 1.1.22] for a proof that avoids it, and promptly stop reading these notes to avoid further frustration.

be a sequence of simple functions converging to g μ-almost everywhere. Then

$$g_i \circ \varphi = \sum_{n=1}^{N_i} (\mathbb{1}_{S_{n,i}} \circ \varphi) \otimes \boldsymbol{x}_{n,i} = \sum_{n=1}^{N_i} \mathbb{1}_{\varphi^{-1}(S_{n,i})} \otimes \boldsymbol{x}_{n,i}$$

is a sequence of simple functions converging to $g \circ \varphi$ ν-almost everywhere, so $g \circ \varphi$ is also strongly measurable. The identity for scalar-valued functions

$$\int_S \|f \circ \varphi(s)\|_X \, \mathrm{d}\mu(s) = \int_T \|f(s)\|_X \, \mathrm{d}\nu(t)$$

implies that $f \circ \varphi \in L^1(S, \mu; X)$. Finally, for all $\boldsymbol{x}^* \in X^*$ we have by the commutation property and the substitution identity for scalar-valued functions

$$\begin{aligned}\Big\langle \int_S f \circ \varphi \, \mathrm{d}\mu, \boldsymbol{x}^* \Big\rangle = \int_S \langle f(\varphi(s)), \boldsymbol{x}^* \rangle \, \mathrm{d}\mu(s) &= \int_T \langle f(s), \boldsymbol{x}^* \rangle \, \mathrm{d}\nu(t) \\ &= \Big\langle \int_T f \, \mathrm{d}\nu, \boldsymbol{x}^* \Big\rangle,\end{aligned}$$

which proves the result.

□

There is also a Fubini theorem for Banach-valued functions (but no Tonelli theorem, as we do not have access to the notion of a non-negative vector-valued function).

Proposition 2.14 (Fubini)**.** *Let $(S, \mathcal{A}, \mu)$ and $(T, \mathcal{B}, \nu)$ be σ-finite measure spaces*[10] *and let $f \in L^1(S \times T; X)$. Then*

- *for almost all $s \in S$ the function $f(s, \cdot)$ is in $L^1(T; X)$,*
- *for almost all $t \in T$ the function $f(\cdot, t)$ is in $L^1(S; X)$,*
- *the functions $\int_T f(\cdot, t) \, \mathrm{d}\nu(t)$ and $\int_S f(s, \cdot) \, \mathrm{d}\mu(s)$ are in $L^1(T; X)$ and $L^1(S; X)$ respectively, and*

$$\int_{S \times T} f \, \mathrm{d}(\mu \times \nu) = \int_T \Big(\int_S f(s,t) \, \mathrm{d}\mu(s) \Big) \mathrm{d}\nu(t) = \int_S \Big(\int_T f(s,t) \, \mathrm{d}\nu(t) \Big) \mathrm{d}\mu(s). \tag{2.4}$$

Proof. Consider an everywhere-defined representative of f. Since f is strongly measurable, by the Pettis measurability theorem (Theorem 2.4), it is weakly measurable and separably valued. Thus the functions $f(s, \cdot)$ and $f(\cdot, t)$ are separably valued for all $s \in S$ and $t \in T$, and by the corresponding scalar-valued result, they are both weakly measurable. Thus $f(s, \cdot)$ and $f(\cdot, t)$ are strongly measurable. Now since the function $(s, t) \mapsto \|f(s, t)\|_X$ is integrable, the scalar Fubini theorem implies all of the integrability claims. The equalities (2.4) are proven by scalarisation: for

[10] The scalar Fubini theorem fails for non-σ-finite spaces!

$\boldsymbol{x}^* \in X^*$ we have

$$\begin{aligned}\Big\langle \int\limits_{S\times T} f \,\mathrm{d}(\mu\times\nu), \boldsymbol{x}^* \Big\rangle &= \int\limits_{S\times T} \langle f(s,t), \boldsymbol{x}^*\rangle \,\mathrm{d}(\mu\times\nu) \\ &= \int\limits_S \int\limits_T \langle f(s,t), \boldsymbol{x}^*\rangle \,\mathrm{d}\nu(t)\,\mathrm{d}\mu(s) \\ &= \int\limits_S \Big\langle \int\limits_T f(s,t)\,\mathrm{d}\nu(t), \boldsymbol{x}^* \Big\rangle \,\mathrm{d}\mu(s) \\ &= \Big\langle \int\limits_S \Big(\int\limits_T f(s,t)\,\mathrm{d}\nu(t)\Big)\,\mathrm{d}\mu(s), \boldsymbol{x}^* \Big\rangle\end{aligned}$$

by the scalar Fubini theorem, and likewise with the order of S and T reversed. □

Let's move away from the abstract stuff for a moment and define the most important operators in analysis.

Definition 2.15. Let X be a complex Banach space. For a Bochner integrable function $f \in L^1(\mathbb{R}^d; X)$ we define the *Fourier transform* as the Bochner integral

$$\hat{f}(\xi) = \mathcal{F}(f)(\xi) := \int\limits_{\mathbb{R}^d} f(t) e^{-2\pi i t\cdot\xi}\,\mathrm{d}t \qquad \forall \xi \in \mathbb{R}^d.$$

Note that $\hat{f}(\xi) \in X$ for all $x \in \mathbb{R}^d$. We also define the *inverse Fourier transform* on $g \in L^1(\mathbb{R}^d; X)$:

$$g^\vee(t) = \mathcal{F}^{-1}(g)(x) := \int\limits_{\mathbb{R}^d} g(\xi) e^{2\pi i t\cdot\xi}\,\mathrm{d}\xi \qquad \forall t \in \mathbb{R}^d.$$

For functions $f \in L^1(\mathbb{T}^d; X)$ on the d-torus $\mathbb{T}^d = [0,1]^d$, we use the same notation for the Fourier transform (and its inverse on $g \in L^1(\mathbb{Z}^d; X)$)

$$\begin{aligned}\hat{f}(n) = \mathcal{F}(f)(n) &:= \int\limits_{\mathbb{T}^d} f(t) e^{-2\pi i t\cdot n}\,\mathrm{d}t \qquad \forall n \in \mathbb{Z}^d, \\ g^\vee(t) = \mathcal{F}^{-1}(g)(t) &:= \sum_{n\in\mathbb{Z}^d} g(n) e^{2\pi i t\cdot n} X \qquad \forall t \in \mathbb{R}^d.\end{aligned}$$

Note that if $f \in L^1(\mathbb{R}^d; X)$, then the function $x \mapsto f(x)e^{-2\pi i x\cdot\xi}$ is strongly integrable for each $\xi \in \mathbb{R}^d$ (see Lemma 2.5), so the definitions above make sense.[11] Furthermore

$$\|\hat{f}(\xi)\|_X \le \int\limits_{\mathbb{R}^d} \|f(x)e^{-2\pi i x\cdot\xi}\|_X \,\mathrm{d}x = \|f\|_{L^1(\mathbb{R}^d;X)}.$$

In fact, $\hat{f}$ is itself strongly integrable, and the Fourier transform is bounded from $L^1(\mathbb{R}^d; X)$ to $L^\infty(\mathbb{R}^d; X)$ (see Exercise 7). Formally, the Fourier transform and inverse Fourier transform are mutually inverse operators, but to make this statement rigourous we have to restrict to appropriate classes of functions or distributions, which for now we will not do.

[11] Analogous statements hold for $\mathbb{T}^d$ and $\mathbb{Z}^d$ of course.

2.4. **Extensions of operators to Bochner spaces.**

Definition 2.16. For a measurable space $(S, \mathcal{A})$ and a set $V \subset M(S;\mathbb{K})$ of measurable scalar-valued functions on S, we define the *algebraic tensor product*

$$V \otimes X := \operatorname{span}\{f \otimes \boldsymbol{x} : f \in V, \boldsymbol{x} \in X\} \subset SM(S;X).^{12}$$

That is, $V \otimes X$ is the set of finite linear combinations of X-valued functions of the form $f \otimes \boldsymbol{x}$, where f is a scalar-valued function in V and $\boldsymbol{x} \in X$.

For example, when V is the set of characteristic functions of measurable sets, $V \otimes X = \Sigma(S;X)$ is the set of X-valued simple functions. Another fundamental example is $V = L^p(S)$ for some $p \in [1,\infty]$.

Proposition 2.17. *Let $(S, \mathcal{A}, \mu)$ be a measure space, X a Banach space, and $p \in [1,\infty)$. Then $L^p(S) \otimes X$ is a dense subspace of $L^p(S;X)$.*

Proof. For $f \in L^p(S)$ and $\boldsymbol{x} \in X$ we compute

$$\|f \otimes \boldsymbol{x}\|^p_{L^p(S,\mu;X)} = \int_S \|f(s)\boldsymbol{x}\|^p_X \,\mathrm{d}\mu(s) = \|\boldsymbol{x}\|^p_X \|f\|^p_{L^p(S)},$$

so that $f \otimes \boldsymbol{x} \in L^p(S;X)$. By linearity, this implies $L^p(S) \otimes X$ is contained in $L^p(S;X)$. For density, note that $L^p(S) \otimes X$ contains $(\Sigma(S;\mathbb{K}) \cap L^p(S)) \otimes X$, and that

$$(\Sigma(S;\mathbb{K}) \cap L^p(S)) \otimes X = \Sigma(S;X) \cap L^p(S;X),$$

as both spaces are equal to the set of simple functions with finite measure support. By Proposition 2.9, this space is dense in $L^p(S;X)$, and thus the same is true of $L^p(S) \otimes X$. ☐

Definition 2.18. Let $(S_i, \mathcal{A}_i, \mu_i)$ $(i \in \{1,2\})$ be measure spaces, $p_1 \in [1,\infty)$ and $p_2 \in [1,\infty]$, and consider a bounded linear operator

$$T \colon L^{p_1}(S_1,\mu_1) \to L^{p_2}(S_2,\mu_2)$$

acting on scalar-valued functions. Let X be a Banach space. The *tensor extension* of T with the identity $I \colon X \to X$ is the map between algebraic tensor products

$$T \otimes I \colon L^{p_1}(S_1,\mu_1) \otimes X \to L^{p_2}(S_2,\mu_2) \otimes X$$

satisfying $(T \otimes I)(f \otimes \boldsymbol{x}) = (Tf) \otimes \boldsymbol{x}$ for all $f \in L^{p_1}(S_1,\mu_1)$ and $\boldsymbol{x} \in X$.

The tensor extension is a well-defined map between algebraic tensor products $L^{p_1}(S_1) \otimes X \to L^{p_2}(S_2) \otimes X$. By Proposition 2.17, $L^{p_1}(S_1) \otimes X$ is a dense subspace of $L^{p_1}(S_1;X)$, while $L^{p_2}(S_2) \otimes X$ is a subspace of $L^{p_2}(S_2;X)$ (possibly non-dense if $p_2 = \infty$), so if there exists $C < \infty$ such that

$$\|(T \otimes I)f\|_{L^{p_2}(S_2;X)} \le C\|f\|_{L^{p_1}(S_1;X)} \qquad \forall f \in L^{p_1}(S_1) \otimes X, \tag{2.5}$$

then $T \otimes I$ may be extended to a bounded linear operator $L^{p_1}(S_1;X) \to L^{p_2}(S_2;X)$.

Definition 2.19. With the notation above, if the estimate (2.5) holds, we say that T *admits a bounded X-valued extension*, and we denote the continuous extension of $T \otimes I$ by $\widetilde{T}_X$, $\widetilde{T}$, or even just T.

Writing out a general element $f \in L^p(S) \otimes X$ as a linear combination of elementary tensors, we see that T admits a bounded X-valued extension if and only if there exists a constant $C < \infty$ such that

$$\Big\| \sum_{n=1}^N (Tf_n) \otimes \boldsymbol{x}_n \Big\|_{L^{p_2}(S_2;X)} \le C \Big\| \sum_{n=1}^N f_n \otimes \boldsymbol{x}_n \Big\|_{L^{p_1}(S_1;X)} \tag{2.6}$$

[12] Functions in $V \otimes X$ are strongly measurable since they have finite-dimensional range.

for all functions $f_n \in L^{p_1}(S_1)$ and vectors $\boldsymbol{x}_n \in X$. This estimate does not simply follow from boundedness of T. It turns out to rely on potentially subtle interactions between the operator T and the Banach space X.

We have already seen one fundamental example.

Example 2.20. Fix a measure space $(S, \mathcal{A}, \mu)$ and let $T \colon L^1(S) \to \mathbb{K}$ denote the Lebesgue integral.[13] Let X be a Banach space. Then for all $f \in \Sigma(S; \mathbb{K}) \otimes X$ we have

$$(T \otimes I)f = (T \otimes I)\Big(\sum_{n=1}^N \mathbb{1}_{S_n} \otimes \boldsymbol{x}_n\Big) = \sum_{n=1}^N T(\mathbb{1}_{S_n}) \otimes \boldsymbol{x}_n = \sum_{n=1}^N \mu(S_n) \otimes \boldsymbol{x}_n = \int_S f \, \mathrm{d}\mu,$$

so that the tensor extension of the Lebesgue integral agrees with the Bochner integral, which we have already shown maps $L^1(S; X)$ to X. Thus the Lebesgue integral admits a bounded X-valued extension, namely the Bochner integral.

In this example, the Banach space X plays no real role; we will see in Theorem 2.22 that this phenomenon occurs for all positive operators. Before that we record a simple observation: tensor extensions can do no better than the original operator.

Proposition 2.21. *Fix measure spaces $(S_i, \mathcal{A}_i, \mu_i)$ $(i \in \{1,2\})$ and exponents $p_1 \in [1,\infty)$, $q \in [1,\infty]$. Let $T \in \mathcal{L}(L^p(S), L^q(S))$ be a bounded linear operator, and let X be a Banach space. Then the tensor extension $T \otimes I$ satisfies*

$$\|T \otimes I\|_{L^{p_1}(S_1;X) \to L^{p_2}(S_2;X)} \geq \|T\|_{L^{p_1}(S_1) \to L^{p_2}(S_2)}.$$

Proof. Fix a nonzero vector $\boldsymbol{x} \in X$. Then for all nonzero $f \in L^{p_1}(S_1)$ we have

$$\begin{aligned}\|(T \otimes I)(f \otimes \boldsymbol{x})\|_{L^{p_2}(S_2;X)} = \|Tf \otimes \boldsymbol{x}\|_{L^{p_2}(S_2;X)} &= \|Tf\|_{L^{p_2}(S_2)} \|\boldsymbol{x}\|_X \\ &= \frac{\|Tf\|_{L^{p_2}(S_2)}}{\|f\|_{L^{p_1}(S_1)}} \|f \otimes \boldsymbol{x}\|_{L^{p_1}(S_1;X)}.\end{aligned}$$

Taking the supremum over all nonzero $f \in L^p(S)$ completes the proof. ☐

Theorem 2.22. *Fix measure spaces $(S_i, \mathcal{A}_i, \mu_i)$ $(i \in \{1,2\})$, $p_1 \in [1,\infty)$, and $p_2 \in [1,\infty]$. Let $T \in \mathcal{L}(L^{p_1}(S_1), L^{p_2}(S_2))$ be a bounded linear operator which is* positive*, meaning that for all a.e. non-negative $f \in L^{p_1}(S_1)$, $Tf \in L^{p_2}(S_2)$ is also a.e. non-negative.*[14] *Then T admits a bounded X-valued extension for every Banach space X, and in fact*

$$\|\widetilde{T}\|_{L^{p_1}(S_1;X) \to L^{p_2}(S_2;X)} = \|T\|_{L^{p_1}(S_1) \to L^{p_2}(S_2)}.$$

this is fundamental enough that it should be isolated as a lemma

Proof. We will show the estimate

$$\|\widetilde{T}f(s)\|_X \leq T(\|f\|_X)(s) \tag{2.7}$$

for all $f \in \Sigma(S; X) \cap L^{p_1}(S_1; X)$ and almost all $s \in S$. This will imply

$$\begin{aligned}\|\widetilde{T}f\|_{L^{p_2}(S_2;X)} &= \Big(\int_S \|\widetilde{T}f(s)\|_X^{p_2} \, \mathrm{d}\mu(s)\Big)^{1/p_2} \\ &\leq \Big(\int_S T(\|f(s)\|_X)^{p_2} \, \mathrm{d}\mu(s)\Big)^{1/p_2} \\ &\leq \|T\|_{L^{p_1}(S_1) \to L^{p_2}(S_2)} \Big(\int_S \|f(s)\|_X^{p_1} \, \mathrm{d}\mu(s)\Big)^{1/p_1} \\ &= \|T\|_{L^{p_1}(S_1) \to L^{p_2}(S_2)} \|f\|_{L^{p_1}(S_1;X)}\end{aligned}$$

[13] This fits in the scope of Definition 2.18 by considering $\mathbb{C}$ as a Lebesgue space $L^1(\mathrm{pt})$ over a single point, with counting measure. Then X may be identified with the Bochner space $L^1(\mathrm{pt}; X)$.

[14] When the scalar field $\mathbb{K}$ is $\mathbb{C}$, 'non-negative' simply means 'real-valued and non-negative'.

which implies the result by density of $\Sigma(S;X) \cap L^{p_1}(S_1;X)$ in $L^{p_1}(S_1;X)$.

Now let's prove (2.7). Consider a simple function

$$f = \sum_{n=1}^{N} \mathbb{1}_{S_n} \otimes \boldsymbol{x}_n$$

and note that $|T(\mathbb{1}_{S_n})| = T(\mathbb{1}_{S_n})$ by positivity of T. Then

$$\begin{aligned}
\|\widetilde{T} f(s)\|_X &= \Big\| \sum_{n=1}^{N} T(\mathbb{1}_{S_n})(s) \boldsymbol{x}_n \Big\|_X \\
&\leq \sum_{n=1}^{N} |T(\mathbb{1}_{S_n})(s)| \|\boldsymbol{x}_n\|_X \\
&= \sum_{n=1}^{N} T(\mathbb{1}_{S_n})(s) \|\boldsymbol{x}_n\|_X = T\Big(\sum_{n=1}^{N} \mathbb{1}_{S_n} \|\boldsymbol{x}_n\|_X \Big)(s) = T(\|f\|_X)(s),
\end{aligned}$$

proving (2.7) and completing the proof. □

Theorem 2.22 shows that the 'extension problem' for positive operators is not much of a problem: positive operators extend automatically. Of course, most interesting operators are not positive.

Example 2.23.

Fourier transform. counterexample for ℓ^∞

write this up

Exercises.

Exercise 1. Let X be a Banach space and $(S, \mathcal{A})$ a measurable space. Prove the containments

$$SM(S;X) \subset M(S;X) \subset WM(S;X).$$

Exercise 2. Let X be a Banach space, let A be a topological space, and let μ be a Borel measure on A.

- If X is separable or A is separable, show that $C(A;X)$ is contained in $L^\infty(A, \mu; X)$.
- Give an example of a topological space A and a Banach space X such that $C(A;X)$ is not contained in $L^\infty(A, \mu; X)$.

Exercise 3. Let X be a Banach space and $(S, \mathcal{A}, \mu)$ a measure space and $p \in [1, \infty)$. Let $V \subset L^p(S; \mathbb{K})$ be a dense subspace. Show that $V \otimes X$ is dense in $L^p(S;X)$.

Exercise 4. Let X be a Banach space and $(S, \mathcal{A}, \mu)$ a measure space such that the σ-algebra $\mathcal{A}$ is finite. Show that the isometric embedding

$$\Phi \colon L^{p'}(S, \mathcal{A}, \mu; X^*) \to L^p(S, \mathcal{A}, \mu; X)^*, \qquad \Phi g(f) = \int_S \langle f(x), g(x) \rangle \, \mathrm{d}\mu(x)$$

is an isomorphism for all $p \in [1, \infty]$.

Exercise 5. Let $(S_i, \mathcal{A}_i, \mu_i)$ $(i \in \{1,2\})$ be measure spaces, let $p_1 \in [1, \infty)$, and let $p_2 \in [1, \infty]$. Suppose that $T \in \mathcal{L}(L^{p_1}(S_1), L^{p_2}(S_2))$ is a bounded linear operator.

- Show that T admits a bounded X-valued extension for all finite-dimensional Banach spaces X.
- Let X be any Banach space and suppose $\boldsymbol{x}^* \in X^*$. Show that for all $f \in L^{p_1}(S_1) \otimes X$,

$$\langle (T \otimes I) f, \boldsymbol{x}^* \rangle = T(\langle f, \boldsymbol{x}^* \rangle).$$

Exercise 6. Let $(S, \mathcal{A}, \mu)$ be a measure space and $p \in (1, \infty)$, let T be a bounded linear operator on $L^p(S, \mu)$, and let X be a Banach space. Let T^* denote the adjoint of T. Show that T admits a bounded X-valued extension if and only if the adjoint $T^* \in \mathcal{L}(L^{p'}(S, \mu))$ admits a bounded X^*-valued extension, and show that

$$(\widetilde{T}_X)^* \Phi g = \widetilde{(T^*)}_{X^*} g$$

for all $g \in L^{p'}(S; X^*)$, where $\Phi \colon L^{p'}(S; X^*) \to L^p(S; X)^*$ is as in Proposition 2.12. Conclude that for all $f \in L^p(S; X)$ and all $g \in L^{p'}(S; X)$,

$$\langle \widetilde{T}_X f, g \rangle = \langle f, \widetilde{(T^*)}_{X^*} g \rangle. \tag{2.8}$$

Exercise 7. Let X be a complex Banach space. Show that $\hat{f} \in C(\mathbb{R}^d; X)$ for all $f \in L^1(\mathbb{R}^d; X)$. Conclude that the Fourier transform is bounded from $L^1(\mathbb{R}^d; X)$ to $L^\infty(\mathbb{R}^d; X)$.

Exercise 8. Let X and Y be Banach spaces and consider an operator-valued function $M \colon \mathbb{R}^d \to \mathcal{L}(X, Y)$, where $\mathcal{L}(X, Y)$ is the Banach space of bounded linear operators from X to Y. Suppose that M is continuous with respect to the strong operator topology on $\mathcal{L}(X, Y)$: that is, suppose that for all vectors $\boldsymbol{x} \in X$, the map

$$M(\cdot)\boldsymbol{x} \colon \mathbb{R}^d \to Y, \qquad \xi \mapsto M(\xi)\boldsymbol{x}$$

is continuous.

- Let $g \colon \mathbb{R}^d \to X$ be strongly measurable. Show that the function $Mg \colon \mathbb{R}^d \to Y$ defined by $(Mg)(\xi) := M(\xi)g(\xi)$ is strongly measurable.
- Suppose in addition that the function $\xi \mapsto \|M(\xi)\|_{\mathcal{L}(X,Y)}$ is measurable,[15] and that
$$\int_{\mathbb{R}^d} \|M(\xi)\|_{\mathcal{L}(X,Y)} \, \mathrm{d}\xi < \infty.$$
Show that the operator $T_M f := (M\hat{f})^\vee$ is well-defined and bounded from $L^1(\mathbb{R}^d; X)$ to $C(\mathbb{R}^d; Y)$.

Exercise 9. Let H be an infinite-dimensional separable Hilbert space with inner product $(\cdot, \cdot)$, and let $(S, \mathcal{A}, \mu)$ be a measure space.

- Show that $L^2(S; H)$ is a Hilbert space with respect to the inner product
$$(f, g) := \int_S (f(s), g(s)) \, \mathrm{d}\mu(s) \qquad (f, g \in L^2(S; H)).$$
- Let $(\boldsymbol{h}_n)_{n \in \mathbb{N}}$ be an orthonormal basis of H, and $(f_n)_{n \in \mathbb{N}}$ an orthonormal basis of $L^2(S, \mu)$. Show that the elementary tensors $\{f_n \otimes \boldsymbol{h}_m : n, m \in \mathbb{N}\}$ are an orthonormal basis of $L^2(S; H)$.
- Suppose that the Hilbert space H, as above, is defined over $\mathbb{C}$. Show that the Fourier transform on the torus, initially defined as a bounded operator $\mathcal{F} \colon L^1(\mathbb{T}^d; H) \to \ell^\infty(\mathbb{Z}^d; H)$, extends to an isometry from $L^2(\mathbb{T}^d; H)$ to $\ell^2(\mathbb{Z}^d; H)$.

Exercise 10. Give an example of a measure space $(S, \mathcal{A}, \mu)$, a Banach space X, and a measurable function $f \colon S \to X$ such that

- $\langle f, \boldsymbol{x}^* \rangle \overset{\text{a.e.}}{=} 0$ for all $\boldsymbol{x}^* \in X^*$,
- f is everywhere non-zero.

I don't even know the answer to this one, maybe too hard...

3. Probability in Banach spaces

Write expository introduction

[15] This can be proven under various assumptions, but don't worry about that now.

3.1. **Gambling in Banach spaces.** To motivate the theory in this section, we're going to imagine a betting game. At each turn, you bet on the outcome of a coin toss. The quantities that you can bet are taken from a Banach space X. The initial state of your wallet, $\boldsymbol{s}_{-1}$, is the zero vector

$$\boldsymbol{s}_{-1} = 0 \in X.$$

At each time $n \in \mathbb{N} = \{0, 1, \ldots\}$, you choose a vector $\boldsymbol{x}_n \in X$ to wager. I then flip a fair coin, which shows either Heads or Tails, and the state of your wallet becomes

$$\boldsymbol{s}_n = \begin{cases} \boldsymbol{s}_{n-1} + \boldsymbol{x}_n & \text{if the coin shows Heads} \\ \boldsymbol{s}_{n-1} - \boldsymbol{x}_n & \text{if the coin shows Tails.} \end{cases}$$

The Banach space X is not ordered, so there is no canonical notion of $\boldsymbol{s}_n$ being 'more' or 'less' than $\boldsymbol{s}_{n-1}$. Thus the game is not about winning or losing (the true winner of the game is Functional Analysis). In this chapter we will discuss various probabilistic concepts that can be well-understood in the context of this game.

3.2. **Filtrations and stochastic processes.**

Definition 3.1. A *filtration* on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is a monotone increasing (i.e. nondecreasing) sequence of σ-subalgebras

$$\mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \subseteq \mathcal{A}.$$

Example 3.2. Consider the unit interval $[0, 1) \subset \mathbb{R}$ with Borel σ-algebra and Lebesgue measure. For each $n \in \mathbb{N}$, let $\mathcal{F}_n$ be the σ-algebra generated by the *dyadic intervals of length* 2^{-n}, i.e. intervals of the form

$$[2^{-n}k, 2^{-n}(k + 1)) \qquad k = 0, 1, 2, \ldots, 2^n - 1.$$

Then $(\mathcal{F}_n)_{n\in\mathbb{N}}$ is a filtration, which we call the *(standard) dyadic filtration*. A Banach-valued function $f : [0, 1) \to X$ is $\mathcal{F}_n$ measurable if and only if it is constant on each dyadic interval of length 2^{-n}.

Example 3.3. Let $\{-1, 1\}$ be a two-point space with uniform probability measure, and consider the infinite product

$$\Omega := \prod_{n\in\mathbb{N}} \{-1, 1\} = \{-1, 1\}^{\mathbb{N}},$$

which (equipped with the product σ-algebra and measure) is a probability space. Elements of Ω are sequences $\omega = (\omega_n)_{n\in\mathbb{N}}$ where each $\omega_n = \pm 1$. Suppose $n \in \mathbb{N}$, fix a vector $\eta = (\eta_0, \eta_1, \ldots \eta_n) \in \{-1, 1\}$ of length $n + 1$, and define the set

$$A_\nu := \{\omega \in \Omega : \omega_k = \eta_k \ \forall k = 0, 1, \ldots, n\};$$

that is, a point $\omega \in \Omega$ belongs to A_ν if its first $n + 1$ components are given by ν. For each $n \in \mathbb{N}$, let $\mathcal{F}_n$ be the σ-algebra generated by all sets A_ν with $\nu \in \{-1, 1\}^n$. Then $(\mathcal{F}_n)_{n\in\mathbb{N}}$ is a filtration, which we call the *coordinate filtration*. A Banach-valued function $f : \Omega \to X$ is $\mathcal{F}_n$-measurable if and only if $f(\omega) = f(\omega_0, \ldots, \omega_n)$ only depends on the first $n + 1$ coordinates of its input variable.

Remark 3.4. Example 3.3 encodes the same information as Example 3.2. Each dyadic interval $I \subset [0, 1]$ has exactly two dyadic subintervals, and every vector $\nu \in \{-1, 1\}^n$ can be extended in exactly two ways to a vector in $\{-1, 1\}^{n+1}$. Equivalently, each infinite sequence $\omega \in \{0, 1\}^{\mathbb{N}}$ corresponds to the binary expansion of a number $t \in [0, 1)$, and this correspondence is bijective up to a measure zero subset (corresponding to the negligible non-uniqueness of binary expansions). The set of sequences $\omega' \in \{0, 1\}^{\mathbb{N}}$ whose first $n+1$ entries coincide with those of ω then corresponds to the set $A_{(\omega_0,\ldots,\omega_n)}$, which corresponds to the unique dyadic interval of length $2^{-(n+1)}$ containing t.

Filtrations are closely linked with stochastic processes. While we don't plan on saying anything really serious about these in this course, it will be useful to keep the core concept in mind, as it guides a lot of probabilistic intuition.

Definition 3.5. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X a Banach space. A *(discrete-time) X-valued stochastic process* on $(\Omega, \mathcal{A}, \mathbb{P})$ is a sequence of $\mathcal{A}$-measurable random variables $f_n \colon \Omega \to X$, $n \in \mathbb{N}$. Given a filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$, a stochastic process $(f_n)_{n\in\mathbb{N}}$ is called *predictable* (with respect to the filtration) if each f_n is $\mathcal{A}_{n-1}$-measurable (with the convention that $\mathcal{A}_{-1} = \{\varnothing, \Omega\}$).

Remark 3.6. With obvious modifications one can talk about filtrations and stochastic processes starting at an arbitrary index, finite filtrations/processes, or filtrations/processes with respect to arbitrary (total or partial) orders, for example with a continuous time index. In this course we will only consider discrete indexing sets contained in $\mathbb{N}$.

One should think of a filtration $(\mathcal{A}_n)_{n=0}^\infty$ as representing the progression of available information over time, usually in relation to a stochastic process. Each σ-subalgebra $\mathcal{A}_n \subset \mathcal{A}$ represents the information available at time n. There are two equivalent ways of thinking about the availability of information: one is that at time n one has access to all $\mathcal{A}_n$-measurable subsets; the other is that at time n one has access to all $\mathcal{A}_n$-measurable functions. The monotonicity assumption says that no information is lost as time progresses. Predictability of a stochastic process $(f_n)_{n\in\mathbb{N}}$ with respect to the filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$ thus says the following: if the available information is represented by $(\mathcal{A}_n)_{n\in\mathbb{N}}$, then at each time n, one already has access to the $\mathcal{A}_n$-measurable function f_{n+1}.

Example 3.7. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, X a Banach space, and let $(f_n)_{n\in\mathbb{N}}$ be an X-valued stochastic process on $(\Omega, \mathcal{A}, \mathbb{P})$. The *filtration generated by the process* $(f_n)_{n\in\mathbb{N}}$ is given by

$$\mathcal{F}_n := \sigma(f_0, f_1, \dots, f_n) \qquad \forall n \in \mathbb{N}.$$

The information-theoretic intuition says that at time $n \in \mathbb{N}$, one 'knows' the functions $f_0, f_1, \dots f_n$, as these are in $\mathcal{F}_n$, and one also knows all functions of the form

$$g \circ (f_0, f_1, \dots, f_n) \colon \omega \mapsto g(f_0(\omega), f_1(\omega), \dots, f_n(\omega))$$

where $g \colon X^{n+1} \to \mathbb{C}$ is measurable (as such compositions are automatically measurable). In fact, all $\mathcal{F}_n$-measurable functions $\Omega \to \mathbb{C}$ are of this form.

cite a reference

Example 3.8. Consider the game we introduced in Section 3.1. At each time $n \in \mathbb{N}$ I flip a fair coin, which comes up Heads (H) or Tails (T) with equal probability. The natural probability space on which to base this game is the infinite product $\Omega = \{-1, +1\}^{\mathbb{N}}$ (see Example 3.3). The value -1 represents Tails, while $+1$ represents Heads. For each $n \in \mathbb{N}$ let $\pi_n \colon \Omega \to \{-1, +1\}$ be the n-th coordinate function, which represents the outcome of the n-th coin toss. The sequence $(\pi_n)_{n\in\mathbb{N}}$ is a scalar-valued stochastic process, and the filtration it generated is precisely the coordinate filtration discussed in Example 3.3.

Your bet at time n, the vector $\boldsymbol{x}_n \in X$, is allowed to depend on the outcomes $\pi_0, \pi_1, \dots, \pi_{n-1}$: you do not need to register all your bets in advance. In probabilistic language, $\boldsymbol{x}_n \colon \Omega \to X$ is $\mathcal{F}_{n-1}$-measurable, i.e. the sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ is a stochastic process which is predictable with respect to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$.

Now consider the stochastic process $(\boldsymbol{s}_n)_{n\in\mathbb{N}}$, representing the evolution of the state of your wallet. By definition we have

$$\boldsymbol{s}_{n+1} = \boldsymbol{s}_n + \pi_{n+1}\boldsymbol{x}_{n+1} \qquad \forall n \in \mathbb{N};$$

keep in mind that this is an equality of X-valued random variables, i.e. functions $\Omega \to X$. Since $\boldsymbol{s}_n$, π_{n+1}, and $\boldsymbol{x}_{n+1}$ are all $\mathcal{F}_{n+1}$-measurable, we find that $\boldsymbol{s}_{n+1}$ is $\mathcal{F}_{n+1}$-measurable (i.e. we know the state of our wallet $\boldsymbol{s}_{n+1}$ at time $n+1$). Heuristically, $\boldsymbol{s}_{n+1}$ should not be $\mathcal{F}_n$-measurable unless $\boldsymbol{x}_{n+1} \equiv 0$, as this would amount to predicting the future (which can only be done by wagering nothing). You should prove this rigourously (Exercise 12).

Definition 3.9. Given a filtration $(\mathcal{A}_n)_{n=0}^{\infty}$ on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, a random variable $T\colon \Omega \to \mathbb{N} \cup \{\infty\}$ is called a *stopping time* (with respect to $(\mathcal{A}_n)$) if

$$\{\omega \in \Omega : T(\omega) \leq n\} \in \mathcal{A}_n \qquad \forall n \geq 0.$$

The stopping time T is *finite* if T is almost surely finite.

Generally stopping times T are defined in terms of some kind of stochastic *stopping condition*. Interpreting the filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$ as modelling the available information at time n, T being a stopping time says precisely that at time n, one 'knows' the set of points $\omega \in \Omega$ for which $T(\omega) \leq n$. Said less precisely, if T is a stopping time, then at time n, one can determine whether or not $T \leq n$.

Example 3.10. We return to the betting game of Section 3.1, elaborated upon in Example 3.8. Let's suppose that our goal is to get the state of our wallet $\boldsymbol{s} \in X$ into a fixed Borel measurable set $K \subset X$, and that we intend to stop betting once this condition holds (i.e. from that point on we only wager the zero vector).

Let

$$T_K(\omega) := \inf\{n \in \mathbb{N} : \boldsymbol{s}_n(\omega) \in K\}$$

with the usual convention that $T_K(\omega) = \infty$ if $\boldsymbol{s}_n(\omega) \notin K$ for all $n \in \mathbb{N}$. That is, T_K is the first time n at which $\boldsymbol{s}_n \in K$. At time n we heuristically know whether or not our wallet satisfied $\boldsymbol{s}_m \in K$ for some $m \leq n$, which indicates that T_K should be a stopping time with respect to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ associated with the stochastic process $(\pi_n)_{n\in\mathbb{N}}$. Rigourously, one shows this by writing for all $n \in \mathbb{N}$

$$\begin{aligned}\{\omega \in \Omega : T_K(\omega) \leq n\} &= \big\{\omega : \inf\{m : \boldsymbol{s}_m(\omega) \in K\} \leq n\big\} \\ &= \{\omega : \boldsymbol{s}_m(\omega) \in K \text{ for some } m \leq n\} \\ &= \bigcup_{m=0}^{n} \boldsymbol{s}_m^{-1}(K),\end{aligned}$$

and noting that since each $\boldsymbol{s}_m$ is $\mathcal{F}_m$-measurable, the set above is $\mathcal{F}_n$-measurable. Thus T_K is a stopping time. Of course, whether T_K is a finite stopping time depends on the set $K \subset X$, the wager vectors $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$, and potentially even the geometry of X (see Exercise 13).

The proof of the following proposition was already done in the previous exercise for the a particular stochastic process, but the proof is identical for a general stochastic process.

Proposition 3.11. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X a Banach space. Let $(f_n)_{n\in\mathbb{N}}$ be an X-valued stochastic process and $K \subset X$ a Borel measurable set. Then the function $T_K\colon \Omega \to \mathbb{N} \cup \{\infty\}$ defined by*

$$T_K(\omega) := \inf\{n \in \mathbb{N} : f_n(\omega) \in K\}$$

is a stopping time.

The stopping time T_K defined above is called the *first hitting time of K*.

3.3. **Conditional expectations.**

a bit of exposition; include defn of $\mathbb{E}$ and talk about 'best approximation with given information'

Definition 3.12. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X a Banach space. Let $f \in L^1(\Omega, \mathcal{A}; X)$ be a integrable X-valued random variable. Given a σ-subalgebra $\mathcal{B} \subset \mathcal{A}$, a *conditional expectation of f given $\mathcal{B}$* is a $\mathcal{B}$-measurable random variable $\mathbb{E}^{\mathcal{B}} f \in L^1(\Omega, \mathcal{B}; X) \subset L^1(\Omega, \mathcal{A}; X)$ such that

$$\int_B \mathbb{E}^{\mathcal{B}} f \, d\mathbb{P} = \int_B f \, d\mathbb{P} \qquad \text{for all } B \in \mathcal{B}. \tag{3.1}$$

Example 3.13. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $\mathcal{B} \subset \mathcal{A}$ be a sub-σ-algebra which is *atomic*, in the sense that there is a collection of pairwise disjoint subsets $(B_\lambda)_{\lambda \in \Lambda}$ of $\mathcal{B}$ which generate $\mathcal{B}$, such that $\mathbb{P}(B_\lambda) > 0$ for all λ, and such that if B_λ can be written as a disjoint union $B_\lambda = C \cup D$ for some sets $C, D \in \mathcal{B}$, then $\mathbb{P}(C) = 0$ or $\mathbb{P}(D) = 0$ (i.e. the sets B_λ are *atoms*). Let's compute *the* conditional expectation $\mathbb{E}^{\mathcal{B}} f$ of an integrable random variable $f \in L^1(\mathcal{A}; X)$ (it turns out there is only one). Since the atoms $(B_\lambda)_\lambda$ generate $\mathcal{B}$ and are pairwise disjoint, and since $\mathbb{E}^{\mathcal{B}} f$ is $\mathcal{B}$-measurable, $\mathbb{E}^{\mathcal{B}} f$ must be constant on each B_λ, so that

$$\mathbb{E}^{\mathcal{B}} f = \sum_{\lambda \in \Lambda} \mathbb{1}_{B_\lambda} \otimes \boldsymbol{x}_\lambda$$

for some vectors $\boldsymbol{x}_\lambda \in X$. Averaging over one of the atoms B_λ and using (3.1) tells us that

$$\boldsymbol{x}_\lambda = \frac{1}{\mathbb{P}(B_\lambda)} \int_{B_\lambda} \mathbb{E}^{\mathcal{B}} f \, d\mathbb{P} = \frac{1}{\mathbb{P}(B_\lambda)} \int_{B_\lambda} f \, d\mathbb{P}.$$

In probabilistic terms, the quantity on the right hand side is the conditional expectation of f given B_λ, which exists since $\mathbb{P}(B_\lambda) > 0$.

The previous example shows that conditional expectations with respect to atomic σ-algebras exist and are unique. The same is true for general σ-algebras, but proving this will take a few steps. First we establish the uniqueness of conditional expectations.

Proposition 3.14. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X a Banach space. For any $f \in L^1(\mathcal{A}; X)$ and any σ-subalgebra $\mathcal{B} \subset \mathcal{A}$, if $\mathbb{E}^{\mathcal{B}} f$ and $\widetilde{\mathbb{E}}^{\mathcal{B}} f$ are two conditional expectations of f given $\mathcal{B}$, then $\mathbb{E}^{\mathcal{B}} f = \widetilde{\mathbb{E}}^{\mathcal{B}} f$ almost surely.*

Proof. First we consider the real one-dimensional case. Fix $f \in L^1(\mathcal{A}; \mathbb{R})$. For all $B \in \mathcal{B}$ we have

$$\int_B \mathbb{E}^{\mathcal{B}} f - \widetilde{\mathbb{E}}^{\mathcal{B}} f \, d\mathbb{P} = \int_B f \, d\mathbb{P} - \int_B f \, d\mathbb{P} = 0.$$

Since $\mathbb{E}^{\mathcal{B}} f - \widetilde{\mathbb{E}}^{\mathcal{B}} f$ is $\mathcal{B}$-measurable, the subsets

$$B_+ := \{\mathbb{E}^{\mathcal{B}} f - \widetilde{\mathbb{E}}^{\mathcal{B}} f > 0\} \quad \text{and} \quad B_- := \{\mathbb{E}^{\mathcal{B}} f - \widetilde{\mathbb{E}}^{\mathcal{B}} f < 0\}$$

are both in $\mathcal{B}$, so we get

$$\int_\Omega |\mathbb{E}^{\mathcal{B}} f - \widetilde{\mathbb{E}}^{\mathcal{B}} f| \, d\mathbb{P} = \int_{B_+} \mathbb{E}^{\mathcal{B}} f - \widetilde{\mathbb{E}}^{\mathcal{B}} f \, d\mathbb{P} - \int_{B_-} \mathbb{E}^{\mathcal{B}} f - \widetilde{\mathbb{E}}^{\mathcal{B}} f \, d\mathbb{P} = 0,$$

establishing that $\mathbb{E}^{\mathcal{B}} f = \widetilde{\mathbb{E}}^{\mathcal{B}} f$ almost surely.

Now let X be any real Banach space and suppose $f \in L^1(\mathcal{A}; X)$. We aim to show that $\mathbb{E}^{\mathcal{B}} f = \widetilde{\mathbb{E}}^{\mathcal{B}} f$ almost surely. By Lemma 2.6 it suffices to show that

$$\langle \mathbb{E}^{\mathcal{B}} f, \boldsymbol{x}^* \rangle \overset{\text{a.s.}}{=} \langle \widetilde{\mathbb{E}}^{\mathcal{B}} f, \boldsymbol{x}^* \rangle \qquad \text{for all } \boldsymbol{x}^* \in X^*. \tag{3.2}$$

This will follow from showing that $\langle \mathbb{E}^{\mathcal{B}} f, \boldsymbol{x}^* \rangle$ and $\langle \widetilde{\mathbb{E}}^{\mathcal{B}} f, \boldsymbol{x}^* \rangle$ are both conditional expectations of the $\mathbb{R}$-valued function $\langle f, \boldsymbol{x}^* \rangle$ given $\mathcal{B}$, thanks to the uniqueness result already established in the one-dimensional case. For all $B \in \mathcal{B}$ we have

$$\int_B \langle \mathbb{E}^{\mathcal{B}} f, \boldsymbol{x}^* \rangle \,\mathrm{d}\mathbb{P} = \Big\langle \int_B \mathbb{E}^{\mathcal{B}} f \,\mathrm{d}\mathbb{P}, \boldsymbol{x}^* \Big\rangle = \Big\langle \int_B f \,\mathrm{d}\mathbb{P}, \boldsymbol{x}^* \Big\rangle = \int_B \langle f, \boldsymbol{x}^* \rangle \,\mathrm{d}\mathbb{P}$$

using that $\mathbb{E}^{\mathcal{B}} f$ is a conditional expectation of f, and the same argument shows that

$$\int_B \langle \widetilde{\mathbb{E}}^{\mathcal{B}} f, \boldsymbol{x}^* \rangle \,\mathrm{d}\mathbb{P} = \int_B \langle f, \boldsymbol{x}^* \rangle \,\mathrm{d}\mathbb{P}.$$

Thus $\langle \mathbb{E}^{\mathcal{B}} f, \boldsymbol{x}^* \rangle$ and $\langle \widetilde{\mathbb{E}}^{\mathcal{B}} f, \boldsymbol{x}^* \rangle$ are conditional expectations of of $\langle f, \boldsymbol{x}^* \rangle$ given $\mathcal{B}$, establishing (3.2), and thus proving that $\mathbb{E}^{\mathcal{B}} f \overset{\text{a.s.}}{=} \widetilde{\mathbb{E}}^{\mathcal{B}} f$.

Finally, if X is a complex Banach space, the result follows by considering real and imaginary parts separately. ☐

Next we will establish existence, positivity, and L^p-contractivity of conditional expectations in the scalar-valued case.

Theorem 3.15. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. For any $f \in L^1(\mathcal{A})$ and any σ-subalgebra $\mathcal{B} \subset \mathcal{A}$, a conditional expectation $\mathbb{E}^{\mathcal{B}} f$ exists. The operator $f \mapsto \mathbb{E}^{\mathcal{B}} f$ is linear, and for all $p \in [1, \infty]$, $\mathbb{E}^{\mathcal{B}}$ is a positive contraction on $L^p(\mathcal{A})$: that is, if $f \in L^p(\mathcal{A})$, then*

$$\|\mathbb{E}^{\mathcal{B}} f\|_p \leq \|f\|_p,$$

and if f is a.s. nonnegative then so is $\mathbb{E}^{\mathcal{B}} f$.

Proof. We can prove positivity from the defining property (3.1), before we establish existence. Suppose $f \in L^1(\mathcal{A})$ is a.s. nonnegative. Then for all $B \in \mathcal{B}$ we have

$$\int_B \mathbb{E}^{\mathcal{B}} f \,\mathrm{d}\mathbb{P} = \int_B f \,\mathrm{d}\mathbb{P} \geq 0,$$

which implies that the $\mathcal{B}$-measurable function $\mathbb{E}^{\mathcal{B}} f$ is a.s. nonnegative.[16]

Now fix $p \in [1, \infty]$ and let $f \in L^p(\mathcal{A})$; we will construct a linear contractive conditional expectation operator $\mathbb{E}^{\mathcal{B}}$ on $L^p(\mathcal{B}; \mathbb{K})$ directly.

Mild case: $p > 1$, so most importantly $p' < \infty$. The inclusion map $\iota \colon L^{p'}(\mathcal{B}) \to L^{p'}(\mathcal{A})$ is contractive, so (using that L^p is the dual of $(L^{p'})^*$, which requires $p' < \infty$) its adjoint $\mathbb{E}^{\mathcal{B}} := \iota^* \colon L^p(\mathcal{A}) \to L^p(\mathcal{B})$ is also contractive. For all $f \in L^p(\mathcal{A})$ and $B \in \mathcal{B}$ we have

$$\int_B \mathbb{E}^{\mathcal{B}} f \,\mathrm{d}\mathbb{P} = \langle \iota^* f, \mathbb{1}_B \rangle = \langle f, \iota \mathbb{1}_B \rangle = \langle f, \mathbb{1}_B \rangle = \int_B f \,\mathrm{d}\mathbb{P},$$

so $\mathbb{E}^{\mathcal{B}} f \in L^p(\mathcal{B}) \subset L^1(\mathcal{B})$ is a conditional expectation of f given $\mathcal{B}$.

(German) spicy case: $p = 1$**.** The difficulty here is that L^1 is *strictly* contained in the dual of L^∞, so taking an adjoint of the inclusion $L^\infty(\mathcal{B}) \to L^\infty(\mathcal{A})$ is not so straightforward.[17] Instead we argue by density. We have that $L^2(\mathcal{A})$ is dense in $L^1(\mathcal{A})$, so we aim to extend the conditional expectation defined above (in the case $p = 2$) by continuity. For $f \in L^2(\mathcal{A})$ and $g \in L^\infty(\mathcal{B})$ we have

$$|\langle \mathbb{E}^{\mathcal{B}} f, g \rangle| = |\langle f, \iota g \rangle| = |\langle f, g \rangle| \leq \|f\|_1 \|g\|_\infty$$

[16] This uses an exercise from measure theory: if g is $\mathcal{B}$-measurable and $\int_B g \geq 0$ for all $\mathcal{B}$-measurable sets, then $g \geq 0$ a.e.. Proof: the set $N := \{g(\omega) < 0\}$ is $\mathcal{B}$-measurable, and assuming it has positive measure leads to the contradiction $0 \leq \int_B g < 0$.

[17] See Exercise 15.

using that $L^\infty(\mathcal{B}) \subset L^2(\mathcal{B})$. Taking the supremum over all nonzero $g \in L^\infty(\mathcal{B})$ proves that

$$\|\mathbb{E}^{\mathcal{B}} f\|_1 \leq \|f\|_1,$$

so $\mathbb{E}^{\mathcal{B}}$ extends to a contraction $L^1(\mathcal{A}) \to L^1(\mathcal{B})$. For $f \in L^1(\mathcal{A})$ and $B \in \mathcal{B}$, using that integration on B is a continuous linear functional on L^1, we have

$$\int_B \mathbb{E}^{\mathcal{B}} f \,\mathrm{d}\mathbb{P} = \lim_{n\to\infty} \int_B \mathbb{E}^{\mathcal{B}} f_n \,\mathrm{d}\mathbb{P} = \lim_{n\to\infty} \int_B f_n \,\mathrm{d}\mathbb{P} = \int_B f \,\mathrm{d}\mathbb{P}$$

where f_n is a sequence in $L^2(\mathcal{A})$ converging to f in $L^1(\mathcal{A})$. Thus $\mathbb{E}^{\mathcal{B}} f$ is a conditional expectation of f given $\mathcal{B}$, and we are done. □

Note that the proof of the previous result also establishes the following adjoint relation, which can also be proven directly from the defining property (3.1) (Exercise 14).

Proposition 3.16. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $\mathcal{B}$ a σ-subalgebra of $\mathcal{A}$. For all $p \in (1, \infty]$, the conditional expectation $\mathbb{E}^{\mathcal{B}}$ on $L^p(\mathcal{A})$ is the adjoint of the corresponding conditional expectation on $L^{p'}(\mathcal{A})$.*

Now we can use the extension theorem for positive operators to show the existence of conditional expectations of Banach-valued random variables.

Proposition 3.17. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let $\mathcal{B}$ be a σ-subalgebra of $\mathcal{A}$, and let X be a Banach space. Then for any $f \in L^1(\mathcal{A}; X)$, a conditional expectation $\mathbb{E}^{\mathcal{B}}_X f$ of f given $\mathcal{B}$ exists. Furthermore, for all $p \in [1, \infty]$, $\mathbb{E}^{\mathcal{B}}$ is a contraction on $L^p(\mathcal{A}; X)$.*

Proof. First fix $p \in [1, \infty)$. Since the conditional expectation $\mathbb{E}^{\mathcal{B}}$ is a positive operator on $L^p(\mathcal{A})$, by Theorem 2.22 it admits a bounded X-valued extension, which we denote by $\mathbb{E}^{\mathcal{B}}_X$. Since $\mathbb{E}^{\mathcal{B}}$ is contractive, so is $\mathbb{E}^{\mathcal{B}}_X$. We just need to show that for all $f \in L^p(\mathcal{A}; X)$, $\mathbb{E}^{\mathcal{B}}_X f$ is a conditional expectation of f given $\mathcal{B}$; we will do this by scalarisation. For all $B \in \mathcal{B}$ and all functionals $\boldsymbol{x}^* \in X^*$, since the function $\langle f, \boldsymbol{x}^* \rangle$ is in $L^1(\mathcal{A})$, we have

$$\begin{aligned}\Big\langle \int_B \mathbb{E}^{\mathcal{B}}_X f \,\mathrm{d}\mathbb{P}, \boldsymbol{x}^* \Big\rangle &= \int_B \langle \mathbb{E}^{\mathcal{B}}_X f, \boldsymbol{x}^* \rangle \,\mathrm{d}\mathbb{P} \\ &\overset{(*)}{=} \int_B \mathbb{E}^{\mathcal{B}}(\langle f, \boldsymbol{x}^* \rangle) \,\mathrm{d}\mathbb{P} = \int_B \langle f, \boldsymbol{x}^* \rangle \,\mathrm{d}\mathbb{P} = \Big\langle \int_B f \,\mathrm{d}\mathbb{P}, \boldsymbol{x}^* \Big\rangle.\end{aligned}$$

(see Exercise 5 for the starred equality). Since this holds for all $\boldsymbol{x}^* \in X^*$, we have

$$\int_B \mathbb{E}^{\mathcal{B}}_X f \,\mathrm{d}\mathbb{P} = \int_B f \,\mathrm{d}\mathbb{P},$$

which shows that $\mathbb{E}^{\mathcal{B}}_X$ is a conditional expectation of f given $\mathcal{B}$.

Now we establish the result for $p = \infty$: let $f \in L^\infty(\mathcal{A}; X) \subset L^2(\mathcal{A}; X)$. Then a conditional expectation $\mathbb{E}^{\mathcal{B}}_X f$ of f given $\mathcal{B}$ is defined as an element of $L^2(\mathcal{B}; X)$: we just need to show that $\|\mathbb{E}^{\mathcal{B}}_X f\|_\infty \leq \|f\|_\infty$. We can test this by duality using Proposition 2.12 and that $L^2(\mathcal{B}; X^*)$ is dense in $L^1(\mathcal{B}; X^*)$.

For all $g \in L^2(\mathcal{B}; X^*)$, since the operator $\mathbb{E}^{\mathcal{B}} \in \mathcal{L}(L^2(\mathcal{A}))$ is self-adjoint, we have

$$\begin{aligned}|\langle \mathbb{E}^{\mathcal{B}}_X f, g \rangle| = |\langle \widetilde{\mathbb{E}^{\mathcal{B}}} f, g \rangle| &\overset{(*)}{=} |\langle f, \widetilde{\mathbb{E}^{\mathcal{B}}} g \rangle| \\ &\leq \|f\|_{L^\infty(\mathcal{A}; X)} \|\mathbb{E}^{\mathcal{B}}_X g\|_{L^1(\mathcal{A}; X^*)} \\ &\leq \|f\|_{L^\infty(\mathcal{A}; X)} \|g\|_{L^1(\mathcal{A}; X^*)}.\end{aligned}$$

For the starred equality see Exercise 6 in the previous chapter, particularly the identity (2.8). Taking the supremum over all nonzero $g \in L^2(\mathcal{B}; X^*)$ completes the proof. □

further properties of vector-valued conditional expectations, as needed

Proposition 3.18. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathcal{B}$ a σ-subalgebra of $\mathcal{A}$, and let X and Y be Banach spaces. Consider an operator-valued function $T \in L^\infty(\Omega, \mathcal{B}; \mathcal{L}(X, Y))$. For $f \in L^1(\Omega, \mathcal{A}; X)$ define the function $Tf \in L^1(\Omega, \mathcal{A}; Y)$ by*

$$Tf(\omega) := T(\omega)f(\omega).$$

Then the identity

$$\mathbb{E}^{\mathcal{B}}(Tf) = T\mathbb{E}^{\mathcal{B}} f$$

holds.

Proof. Exercise 16. □

Remark 3.19. We have only considered conditional expectations $\mathbb{E}^{\mathcal{B}}$ on probability spaces, but the concept can be extended to general measure spaces $(S, \mathcal{A}, \mu)$ provided that the measure μ is σ-finite on the σ-subalgebra $\mathcal{B} \subset \mathcal{A}$ (although the arguments require a fair bit of modification). This approach is taken in [4].

3.4. Martingales and martingale transforms.

Definition 3.20. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let $(\mathcal{A}_n)_{n\in\mathbb{N}}$ be a filtration on this space, and let X be a Banach space.

- A stochastic process $(M_n)_{n\in\mathbb{N}}$ with each $M_n \in L^1(\Omega; X)$ is called a *martingale* with respect to $(\mathcal{A}_n)_{n\in\mathbb{N}}$ if

$$M_n = \mathbb{E}^{\mathcal{A}_n} M_{n+1} \qquad \forall n \in \mathbb{N}. \tag{3.3}$$

  Note that in particular each M_n is $\mathcal{A}_n$-measurable.
- Let $(M_n)_{n\in\mathbb{N}}$ be a martingale as above. The associated *martingale difference sequence* is the process $(dM_n)_{n\in\mathbb{N}}$ in $L^1(\Omega; X)$ defined by

$$dM_0 := M_0, \qquad dM_k := M_k - M_{k-1} \quad (k \geq 1)$$

  (the difference equation can be made to hold for all $k \geq 0$ by defining $M_{-1} := 0$).

Remark 3.21. Now is a good time to complete Exercise 17, establishing a few elementary properties of martingales.

Martingales are stochastic processes that are 'balanced': at time n, the best estimate of the state of the process at time $n+1$ is precisely the current state of the process. They are intricately linked with many topics in harmonic and functional analysis.

Example 3.22. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\mathcal{A}_n)_{n\in\mathbb{N}}$ a filtration. Let X be a Banach space and $f \in L^1(\Omega; X)$. For each $n \in \mathbb{N}$ define $f_n := \mathbb{E}^{\mathcal{A}_n} f \in L^1(\Omega; X)$. Then by the monotonicity property of conditional expectations,

$$\mathbb{E}^{\mathcal{A}_n} f_{n+1} = \mathbb{E}^{\mathcal{A}_n}\mathbb{E}^{\mathcal{A}_{n+1}} f = \mathbb{E}^{\mathcal{A}_n} f = f_n,$$

so $(f_n)_{n\in\mathbb{N}}$ is a martingale. This is called the *martingale associated with* f.

When $\Omega = [0, 1)$ is the unit interval with Borel σ-algebra and Lebesgue measure, and when we consider the dyadic filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ as in Example 3.2, the martingale associated with a function $f \in L^1([0, 1); X)$ is given by

$$f_n = \sum_{I \in \mathcal{D}_n} \mathbb{1}_I \otimes \langle f \rangle_I$$

where $\mathcal{D}_n$ is the set of dyadic intervals $I \subset [0,1)$ of length 2^{-n} and $\langle f \rangle_I = \fint_I f(t)\,\mathrm{d}t$ is the average of f on I (these conditional expectations were computed in Example 3.13). Each dyadic interval $I \in \mathcal{D}_{n-1}$ can be 'halved', i.e. $I = I_- \cup I_+$, where $I_\pm \in \mathcal{D}_n$ and I_- is to the left of I_+ (i.e. $\sup_{I_-} = \inf_{I_+}$). Let's compute the difference df_n on an interval $I \in \mathcal{D}_{n-1}$:

$$\begin{aligned}
\mathbb{1}_I(df_n) &= \mathbb{1}_I(f_n - f_{n-1}) \\
&= \mathbb{1}_{I_-} \otimes \langle f \rangle_{I_-} + \mathbb{1}_{I_+} \otimes \langle f \rangle_{I_+} - (\mathbb{1}_- + \mathbb{1}_+) \otimes \langle f \rangle_I \\
&= \mathbb{1}_{I_-} \otimes \Big(\frac{2}{|I|} \int\limits_{I_-} f - \frac{1}{|I|} \int\limits_{I} f \Big) + \mathbb{1}_{I_+} \otimes \Big(\frac{2}{|I|} \int\limits_{I_+} f - \frac{1}{|I|} \int\limits_{I} f \Big) \\
&= \mathbb{1}_{I_-} \otimes \Big(\frac{1}{|I|} \int\limits_{I_-} f - \frac{1}{|I|} \int\limits_{I_+} f \Big) - \mathbb{1}_{I_+} \otimes \Big(\frac{1}{|I|} \int\limits_{I_-} f - \frac{1}{|I|} \int\limits_{I_+} f \Big) \\
&= h_I \otimes \langle f, h_I \rangle
\end{aligned}$$

where

$$h_I := \frac{1}{|I|^{1/2}} (\mathbb{1}_{I_-} - \mathbb{1}_{I_+})$$

is the (L^2-normalised) *Haar function* associated with $I \in \mathcal{D}_n$. Thus the representation of f in terms of martingale differences corresponds to its Haar expansion (ignoring the issue of whether the sums converge):

$$f = f_0 + \sum_{n \geq 1} df_n = \langle f \rangle_{[0,1)} + \sum_{n \geq 1} \sum_{I \in \mathcal{D}_{n-1}} h_I \otimes \langle f, h_I \rangle.$$

The martingale associated with an L^p function has the following fundamental convergence property. This will be extended to almost sure convergence in Theorem 3.31.

Theorem 3.23. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\mathcal{A}_n)_{n\in\mathbb{N}}$ a filtration. Let $\mathcal{A}_\infty \subset \mathcal{A}$ be the sub-σ-algebra generated by $\cup_{n\in\mathbb{N}} \mathcal{A}_n$. Let X be a Banach space and $p \in [1,\infty)$. Then for all $f \in L^p(\Omega, \mathbb{P}; X)$ we have $\mathbb{E}^{\mathcal{A}_n} f \to \mathbb{E}^{\mathcal{A}_\infty} f$ with convergence in L^p.*

Proof. Since $\mathbb{E}^{\mathcal{A}_n} f = \mathbb{E}^{\mathcal{A}_n}\mathbb{E}^{\mathcal{A}_\infty} f$ for all $n \in \mathbb{N}$, it suffices to assume that $f = \mathbb{E}^{\mathcal{A}_\infty} f$, i.e. that f is $\mathcal{A}_\infty$-measurable. We will reduce to showing that

$$\bigcup_{n\in\mathbb{N}} L^p(\Omega, \mathcal{A}_n, \mathbb{P}; X)$$

is dense in $L^p(\Omega, \mathcal{A}_\infty, \mathbb{P}; X)$. Assuming this is true for the moment, given $\varepsilon > 0$, there exists $n \in \mathbb{N}$ and $g \in L^p(\mathcal{A}_n; X)$ such that $\|f - g\|_p < \varepsilon$. We also have $\mathbb{E}^{\mathcal{A}_m} g = g$ for all $m > n$, so for all such m we have

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

Taking $m \to \infty$ and noting that ε was arbitrary, we find that $\mathbb{E}^{\mathcal{A}_m} f \to f$ in L^p.

It remains to prove the density statement, and by Exercise 3 it suffices to do this in the scalar case $X = \mathbb{K}$. Consider the collection of sets

$$\mathcal{C} := \Big\{ A \in \mathcal{A}_\infty : \mathbb{1}_A \in \overline{\bigcup_{n\in\mathbb{N}} L^p(\mathcal{A}_n)} \Big\}$$

where the closure is in $L^p(\mathcal{A}_\infty)$. Then $\mathcal{C}$ is a σ-algebra which contains $\mathcal{A}_n$ for each $n \in \mathbb{N}$, so that $\mathcal{A}_\infty = \mathcal{C}$, which implies that all $\mathcal{A}_\infty$-simple functions are contained in $\overline{\bigcup_{n\in\mathbb{N}} L^p(\mathcal{A}_n)}$, and thus that this closure is $L^p(\mathcal{A}_\infty)$. $\square$

Example 3.24. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, X a Banach space, and let $(g_n)_{n\in\mathbb{N}}$ be a sequence of random variables in $L^1(\Omega; X)$ which are mutually independent. Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be the filtration generated by the process $(g_n)_{n\in\mathbb{N}}$, and for each $n \in \mathbb{N}$ let $\sigma_n := \sum_{m=0}^n g_n$ be the sum of the first $n+1$ random variables. Then we have

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_n}\sigma_{n+1} &= \mathbb{E}^{\mathcal{F}_n}\Big(\sum_{m=0}^n g_m\Big) + \mathbb{E}^{\mathcal{F}_n} g_{n+1} \\ &= \sum_{m=0}^n g_m = \sigma_n\end{aligned}$$

since the random variables $(g_m)_{m=0}^n$ are $\mathcal{F}_n$-measurable and g_{n+1} is independent of $\mathcal{F}_n$. Thus the sum process $(\sigma_n)_{n\in\mathbb{N}}$ is a martingale.

this needs to be an exercise or a note

Proposition 3.25. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, X a Banach space, and let $(M_n)_{n\in\mathbb{N}}$ be a martingale (valued in X) with respect to a filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$. Let Y be another Banach space, and let $(T_n)_{n\in\mathbb{N}}$ be a sequence of operators in $L^\infty(\Omega; \mathcal{L}(X, Y))$ which is predictable with respect to $(\mathcal{A}_n)_{n\in\mathbb{N}}$. For each $n \in \mathbb{N}$ and $\omega \in \Omega$ define a function $T_n M_n \in L^1(\Omega; Y)$ by*

$$(T \cdot M)_n := \sum_{m=0}^n T_m dM_m \in Y.$$

Then $(T \cdot M)_{n\in\mathbb{N}}$ is a (Y-valued) martingale with respect to $(\mathcal{A}_n)_{n\in\mathbb{N}}$.

Proof. Integrability of each $(T \cdot M)_n$ follows from that of each M_m and the a.s. uniform boundedness of each T_m. To see that $(T \cdot M)_{n\in\mathbb{N}}$ is a martingale with respect to $(\mathcal{A}_n)_{n\in\mathbb{N}}$, it suffices to show that

pur a prop. allowing for this

$$\mathbb{E}^{\mathcal{A}_n} d(T \cdot M)_{n+1} = 0 \tag{3.4}$$

for all $n \in \mathbb{N}$. By Proposition 3.18, since T_{n+1} is $\mathcal{A}_n$-measurable by the predictability assumption and dM_{n+1} is independent of $\mathcal{A}_n$ (Exercise **??**), we have

$$\mathbb{E}^{\mathcal{A}_n} d(T \cdot M)_{n+1} = \mathbb{E}^{\mathcal{A}_n}(T_{n+1} dM_{n+1}) = T_{n+1}\mathbb{E}^{\mathcal{A}_n} dM_{n+1} = 0,$$

proving (3.4) as required. □

Example 3.26. We return once more to our betting game, with notation given in Example 3.8. Consider the stochastic process $(\boldsymbol{s}_n)_{n\in\mathbb{N}}$ representing the evolution of the state of your wallet: recall that

$$d\boldsymbol{s}_{n+1} = \boldsymbol{s}_{n+1} - \boldsymbol{s}_n = \pi_{n+1}\boldsymbol{x}_{n+1} \qquad \forall n \in \mathbb{N},$$

where π_{n+1} is the outcome of the coin toss and $\boldsymbol{x}_{n+1}$ is the vector wagered at time $n+1$. The process $(\pi_n)_{n\in\mathbb{N}}$ generates the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$. Since the random variables $(\pi_n)_{n\in\mathbb{N}}$ are mutually independent, the sum process $(\sigma_n)_{n\in\mathbb{N}}$ given by

$$\sigma_n := \sum_{m=0}^n \pi_n$$

is a martingale (see Example 3.24). Now suppose that the wager vectors $\boldsymbol{x}_n \in L^1(\Omega, X)$ are integrable. We assumed that this sequence is predictable with respect to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ (meaning that we place the n-th bet before seeing the outcome of the n-th coin toss), so we can construct the martingale transform $\boldsymbol{x}\cdot\sigma$.[18] By definition, the difference sequence of $\boldsymbol{x}\cdot\sigma$ is

$$d(\boldsymbol{x}\cdot\sigma)_{n+1} = \boldsymbol{x}_{n+1} d\sigma_{n+1} = \pi_{n+1}\boldsymbol{x}_{n+1} = d\boldsymbol{s}_{n+1},$$

[18]Technically we are identifying X with $\mathcal{L}(\mathbb{C}; X)$ here. Given a vector $\boldsymbol{y} \in X$, the associated linear operator $\mathbb{C} \to X$ maps $\lambda \in \mathbb{C}$ to $\lambda\boldsymbol{y} \in X$.

and the initial term is

$$(\boldsymbol{x} \cdot \sigma)_0 = \boldsymbol{x}_0 \pi_0 = \boldsymbol{s}_0,$$

so we have the equality of martingales $(\boldsymbol{s}_n)_{n\in\mathbb{N}} = ((\boldsymbol{x} \cdot \sigma)_n)_{n\in\mathbb{N}}$. That is, the state of your wallet $(\boldsymbol{s}_n)_{n\in\mathbb{N}}$ is a martingale, and it is given by the martingale transform of the sum of coin flips $(\sigma_n)_{n\in\mathbb{N}}$ by the wager vectors $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$.

3.5. **Maximal inequalities and pointwise convergence.**

Definition 3.27. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let $(\mathcal{A}_n)_{n\in\mathbb{N}}$ be a filtration on this space. A real-valued stochastic process $(M_n)_{n\in\mathbb{N}}$ with each $M_n \in L^1(\Omega; \mathbb{R})$ is called a *submartingale* with respect to $(\mathcal{A}_n)_{n\in\mathbb{N}}$ if

$$M_n \overset{\text{a.s.}}{\leq} \mathbb{E}^{\mathcal{A}_n} M_{n+1} \qquad \forall n \in \mathbb{N}. \tag{3.5}$$

As an example, consider a martingale $(f_n)_{n\in\mathbb{N}}$ taking values in a Banach space X, with respect to a filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$. Then by the pointwise estimate (2.7) established in the proof of Theorem 2.22 we have for all $n \in \mathbb{N}$

$$\|f_n\|_X = \|\mathbb{E}^{\mathcal{A}_n} f_{n+1}\|_X \overset{\text{a.s.}}{\leq} \mathbb{E}^{\mathcal{A}_n} \|f_{n+1}\|_X,$$

so that the sequence $(\|f_n\|_X)_{n\in\mathbb{N}}$ is a submartingale.

Definition 3.28. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X a Banach space. Given an X-valued stochastic process $(f_n)_{n\in\mathbb{N}}$ we define the *maximal function*

$$f^*(\omega) := \sup_{n\in\mathbb{N}} \|f_n(\omega)\|_X \qquad \forall \omega \in \Omega$$

Theorem 3.29 (Doob's maximal inequalities)**.** *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $(f_n)_{n\in\mathbb{N}}$ be a real-valued submartingale with respect to a filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$. Then for all $t > 0$ we have*

$$t\mathbb{P}(\{f^* > t\}) \leq \sup_{n\in\mathbb{N}} \int_{\{f^*>t\}} |f_n| \, \mathrm{d}\mathbb{P},$$

and if $f^* \overset{\text{a.s.}}{\geq} 0$ then for all $p \in (1, \infty)$ we have

$$\|f^*\|_{L^p(\Omega)} \leq p' \sup_{n\in\mathbb{N}} \|f_n\|_{L^p(\Omega)}.$$

Corollary 3.30. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X a Banach space, and let $(f_n)_{n\in\mathbb{N}}$ be an X-valued martingale. Then for all $t > 0$ we have*

$$t\mathbb{P}(\{f^* > t\}) \leq \sup_{n\in\mathbb{N}} \|f_n\|_{L^1(\Omega;X)},$$

and for all $p \in (1, \infty)$ we have

$$\|f^*\|_{L^p(\Omega;X)} \leq p' \sup_{n\in\mathbb{N}} \|f_n\|_{L^p(\Omega;X)}.$$

Proof of Theorem 3.29. Fix $t > 0$ and define the stopping time (relative to the filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$)

$$T := \inf\{k \in \mathbb{N} : f_k > t\}.$$

Then we have

$$\begin{aligned} t\mathbb{P}(\{f^* > t\}) = t\mathbb{P}(\{T < \infty\}) &= \lim_{N\to\infty} \sum_{k=0}^{N} t\mathbb{P}(\{T = k\}) \\ &\leq \lim_{N\to\infty} \Big| \sum_{k=0}^{N} \int_{\{T=k\}} f_k \, \mathrm{d}\mathbb{P} \Big| \leq \lim_{N\to\infty} \Big| \sum_{k=0}^{N} \int_{\{T=k\}} \mathbb{E}^{\mathcal{A}_k} f_N \, \mathrm{d}\mathbb{P} \Big| \end{aligned}$$

using that $f_k > t$ on the set $\{T = k\}$ and that $(f_n)_{n\in\mathbb{N}}$ is a submartingale. Since T is a stopping time we have $\{T = k\} \in \mathcal{A}_k$, so for all $N \in \mathbb{N}$ we have

$$\Big|\sum_{k=0}^{N} \int_{\{T=k\}} \mathbb{E}^{\mathcal{A}_k} f_N \,\mathrm{d}\mathbb{P}\Big| = \Big|\sum_{k=0}^{N} \int_{\{T=k\}} f_N \,\mathrm{d}\mathbb{P}\Big| = \Big|\int_{\{T\le N\}} f_N \,\mathrm{d}\mathbb{P}\Big| \le \sup_{N\in\mathbb{N}} \int_{\{f^*>t\}} |f_N| \,\mathrm{d}\mathbb{P}$$

proving the first inequality.

For the second inequality, using that $f^* \overset{\text{a.s.}}{\ge} 0$ and supposing $\varepsilon > 0$, for N sufficiently large we have

$$\begin{aligned}
\|f^*\|_p^p &= \int_0^\infty p t^{p-1} \mathbb{P}(\{f^* > t\}) \,\mathrm{d}t \\
&\le (1+\varepsilon) \int_0^\infty p t^{p-2} \int_{\{f^*>t\}} f_N(\omega) \,\mathrm{d}\mathbb{P}(\omega) \,\mathrm{d}t \\
&= (1+\varepsilon) \int_\Omega f_N(\omega) \Big(\int_0^{f^*(\omega)} p t^{p-2} \,\mathrm{d}t \Big) \,\mathrm{d}\mathbb{P}(\omega) \\
&= (1+\varepsilon) \int_\Omega \frac{p}{p-1} f_N(\omega) f^*(\omega)^{p-1} \,\mathrm{d}\mathbb{P}(\omega) \\
&\le (1+\varepsilon) p' \|f_N\|_p \|f^*\|_p^{p-1}
\end{aligned}$$

using Hölder's inequality in the last step. Dividing through by $\|f^*\|_p^{p-1}$ yields

$$\|f^*\|_p \le (1+\varepsilon) p' \sup_{n\in\mathbb{N}} \|f_n\|_p$$

for all $\varepsilon > 0$, which completes the proof. □

A general principle says that convergence in L^p plus L^p-boundedness of the appropriate maximal function implies almost everywhere convergence. Here is the consequence for martingales.

Theorem 3.31. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\mathcal{A}_n)_{n\in\mathbb{N}}$ a filtration. Let $\mathcal{A}_\infty \subset \mathcal{A}$ be the sub-σ-algebra generated by $\cup_{n\in\mathbb{N}} \mathcal{A}_n$. Let X be a Banach space and $p \in [1, \infty)$. Then for all $f \in L^p(\Omega, \mathbb{P}; X)$ we have $\mathbb{E}^{\mathcal{A}_n} f \to \mathbb{E}^{\mathcal{A}_\infty} f$ almost surely.*

Proof. See Exercise 18. □

Note that this theorem says that the martingale $(f_n)_{n\in\mathbb{N}} = (\mathbb{E}^{\mathcal{A}_n} f)_{n\in\mathbb{N}}$ associated with an $\mathcal{A}_\infty$-measurable function $f \in L^p(\Omega; X)$ converges almost surely to f.

Of course, this martingale is L^p-bounded in the following sense:

Definition 3.32. An X-valued stochastic process $(f_n)_{n\in\mathbb{N}}$ on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is called *L^p-bounded* if

$$\sup_{n\in\mathbb{N}} \|f_n\|_{L^p(\Omega;X)} < \infty.$$

This raises an important question: given a general L^p-bounded X-valued martingale $(f_n)_{n\in\mathbb{N}}$, does it automatically hold that $f_n = \mathbb{E}^{\mathcal{A}_n} f$ for some $f \in L^p(\Omega; X)$? The answer turns out to depend on the geometry of X, and we will discuss this in the next section. For now we will quickly settle the scalar case. When $p = 1$ we need an additional condition to guarantee relative weak compactness.

Definition 3.33. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. A bounded subset $\mathcal{F} \subset L^1(\Omega)$ is *uniformly integrable* (or *equi-integrable*) if for all $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\mathbb{P}(A) < \delta \Rightarrow \sup_{f \in \mathcal{F}} \int_A \|f(\omega)\|_X \, \mathrm{d}\mathbb{P}(\omega) < \varepsilon \qquad \forall A \in \mathcal{A}.$$

For a Banach space X, we say that an X-valued stochastic process $(f_n)_{n\in\mathbb{N}}$ is uniformly integrable if the set $\{\|f_n\|_X : n \in \mathbb{N}\} \subset L^1(\Omega)$ is uniformly integrable.

A bounded subset $\mathcal{F} \subset L^1(\Omega)$ is uniformly integrable if and only if it is weakly relatively compact (see [1, Theorem 5.2.9]). For $p \in (1, \infty)$, since $L^p(\Omega)$ is reflexive, every bounded subset $\mathcal{F} \subset L^p(\Omega)$ is weakly relatively compact (see Corollary 9.5 of the Banach–Alaoglu theorem). In both cases, the Eberlein–Smulian theorem (Theorem 9.6) says that every bounded sequence in $L^p(\Omega)$ (with the additional assumption of uniform integrability if $p = 1$) has a convergent subsequence. We use this to prove the following theorem.

Theorem 3.34. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and fix $p \in [1, \infty)$. Let $(f_n)_{n\in\mathbb{N}}$ be an L^p-bounded scalar-valued martingale with respect to a filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$. If $p = 1$, suppose furthermore that $(f_n)_{n\in\mathbb{N}}$ is uniformly integrable. Then there exists a function $f_\infty \in L^p(\Omega, \mathcal{A}, \mathbb{P})$ such that $f_n = \mathbb{E}^{\mathcal{A}_n} f_\infty$ for all $n \in \mathbb{N}$.*

Note that by Theorems 3.23 and 3.31 we have $f_n \to f_\infty$ almost surely and in L^p, so a corollary is that the martingale (f_n) is a.s. convergent.

Proof. By the discussion above, there is a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ which converges weakly to a limit $f_\infty \in L^p(\Omega)$. For all $n \in \mathbb{N}$ and $A \in \mathcal{A}_n$,

$$\int_A f_\infty \, \mathrm{d}\mathbb{P} = \lim_{k\to\infty} \int_A f_{n_k} \, \mathrm{d}\mathbb{P},$$

and whenever k is so large that $n_k \geq n$ we have

$$\int_A f_{n_k} \, \mathrm{d}\mathbb{P} = \int_A \mathbb{E}^{\mathcal{A}_n} f_{n_k} \, \mathrm{d}\mathbb{P} = \int_A f_n \, \mathrm{d}\mathbb{P}$$

by the martingale property. Thus for all $A \in \mathcal{A}_n$ we have

$$\int_A f_\infty \, \mathrm{d}\mathbb{P} = \int_A f_n,$$

which implies that $f_n = \mathbb{E}^{\mathcal{F}_n} f_\infty$. □

Remark 3.35. In fact, when $p = 1$, the assumption of uniform integrability can be removed, and the following is true: every scalar-valued martingale with $\sup_{n\in\mathbb{N}} \|f_n\|_1 < \infty$ converges almost everywhere (but not necessarily in L^1). We'll skip the proof of this result so as to spend more time with Banach spaces, but if you're interested see [6, Theorem 1.34]. The proof isn't particularly difficult.

Later on we are going to need a corresponding result for submartingales.

Theorem 3.36. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(f_n)_{n\in\mathbb{N}}$ a (real-valued) submartingale on Ω which is L^1-bounded and uniformly integrable. Then f_n converges almost surely and in L^1.*

come back to this proof. Pisier Theorem 1.36

Proof. □

3.6. **Martingale convergence as a Banach space property.** With Theorem 3.34 as inspiration, we make the following definition.

Definition 3.37. For $p \in [1, \infty]$, we say that a Banach space X has the *p-martingale convergence property* (or *p-MCP*) if every X-valued martingale which is uniformly bounded in L^p (and uniformly integrable, when $p = 1$) converges almost surely. We say that X has this property with respect to a given probability space $(\Omega, \mathcal{A}, \mathbb{P})$ if it holds for every X-valued martingale on this space.

Remark 3.38. This is not a standard definition, because it turns out to be equivalent to the more familiar *Radon–Nikodym property*, which we will discuss in the next chapter. But for now, it will help to give the property a more martingaley name.

Formally, the 1-MCP is the strongest of these properties, and the ∞-MCP is the weakest. As stated in the remark above, it will turn out that these properties are equivalent, but we don't know that yet. For the moment we will investigate the p-MCP naïvely, without invoking this equivalence.

Proposition 3.39. *Let $1 \leq p < q \leq \infty$. If a Banach space X has the p-MCP, then it also has the q-MCP.*

Proof. For $p > 1$ this follows from the continuous inclusion $L^q(\Omega) \subset L^p(\Omega)$ for probability spaces Ω: an L^q-bounded martingale is also L^p-bounded, and one can then invoke the p-MCP to derive the q-MCP. For $p = 1$ the same argument applies once we show that a bounded subset $\mathcal{F} \subset L^q(\Omega)$ is uniformly integrable. To see this, for all $f \in \mathcal{F}$ and measurable $A \subset \Omega$ use Hölder's inequality to estimate

$$\int_A |f(\omega)| \, \mathrm{d}\mathbb{P}(\omega) \leq \|f\|_p \mathbb{P}(A)^{1/p'}.$$

Thus for all $\varepsilon > 0$, if $\mathbb{P}(A) < (\varepsilon / \sup\{\|f\|_p : f \in \mathcal{F}\})^{p'}$ then

$$\int_A |f(\omega)| \, \mathrm{d}\mathbb{P}(\omega) < \varepsilon,$$

so $\mathcal{F}$ is uniformly integrable.

□

Remark 3.35 says that the scalar field $\mathbb{K}$ has the 1-MCP, and arguing coordinatewise shows that every finite-dimensioal Banach space also has this property. In the following two examples we will show that the Banach spaces c_0 and $L^1(\Omega)$ do not have the ∞-MCP (and hence do not have the p-MCP for any $p \in [1, \infty]$).

Example 3.40. Let $\{-1, 1\}$ be equipped with the uniform probability measure and let $\Omega = \{-1, 1\}^{\mathbb{N}}$ have the product measure. Let $\pi_n \colon \Omega \to \{-1, 1\}$ be the n-th coordinate function. Consider the Banach space c_0 of scalar-valued sequences $(a_n)_{n \in \mathbb{N}}$ satisfying $\lim_{n \to \infty} a_n = 0$, equipped with the ℓ^∞-norm. Let $(\boldsymbol{e}_n)_{n \in \mathbb{N}}$ be the canonical basis of c_0, i.e.

$$\boldsymbol{e}_n(m) = \begin{cases} 1 & m = n \\ 0 & m \neq n. \end{cases}$$

Define a c_0-valued martingale $(f_n)_{n \in \mathbb{N}}$ with respect to the filtration generated by $(\pi_n)_{n \in \mathbb{N}}$ by

$$f_n := \sum_{k=0}^{n} \pi_k \otimes \boldsymbol{e}_k.$$

(As shown in Example 3.26, this is a martingale; in fact, it is just a special case of our 'betting game', in which one wagers the vector $\boldsymbol{e}_n$ at time n, regardless of the previous outcomes.) Then we have

$$\|f_n\|_{L^\infty(\Omega;c_0)} = \operatorname*{ess\,sup}_{\omega\in\Omega} \|(\pi_k(\omega))_{k=0}^n\|_{c_0} = 1,$$

so the martingale $(f_n)_{n\in\mathbb{N}}$ is uniformly bounded in L^∞. But for all ω in Ω and all $n < m$, we have

$$\|f_n(\omega) - f_m(\omega)\|_{c_0} = \max_{n<k\leq m} |\pi_k(\omega)| = 1,$$

so the sequence $(f_n(\omega))_{n\in\mathbb{N}}$ is not Cauchy in c_0 and hence not convergent.

Example 3.41. With the notation of the previous example, let $X = L^1(\Omega)$ and consider the X-valued martingale

$$f_n(\omega) := \prod_{k\leq n} (1 + \pi_k(\omega)\pi_k) \tag{3.6}$$

(Exercise 20 asks you to show that this is indeed a martingale). For all $\omega \in \Omega$ and $n \in \mathbb{N}$ we have

$$\begin{aligned} \|f_n(\omega)\|_X &= \int_\Omega \prod_{k\leq n} (1 + \pi_k(\omega)\pi_k(\eta))\,\mathrm{d}\mathbb{P}(\eta) \\ &= \prod_{k\leq n} \int_\Omega (1 + \pi_k(\omega)\pi_k(\eta))\,\mathrm{d}\mathbb{P}(\eta) = 1 \end{aligned}$$

by mutual independence of the variables π_k, so the martingale (f_n) is uniformly bounded in L^∞. However, we also have

$$\begin{aligned} &\|f_n(\omega) - f_{n+1}(\omega)\|_X \\ &= \int_\Omega \Big| \Big(1 - (1 + \pi_{n+1}(\omega)\pi_{n+1}(\eta))\Big) \prod_{k\leq n} (1 + \pi_k(\omega)\pi_k(\eta)) \Big| \,\mathrm{d}\mathbb{P}(\eta) \\ &= \int_\Omega |\pi_{n+1}(\omega)\pi_{n+1}(\eta)|\,\mathrm{d}\mathbb{P}(\eta) \prod_{k\leq n} \int_\Omega 1 + \pi_k(\omega)\pi_k(\eta)\,\mathrm{d}\mathbb{P}(\eta) = 1, \end{aligned}$$

which shows that the sequence $(f_n(\omega))_{n\in\mathbb{N}}$ cannot be Cauchy in X, and hence is not convergent.

It is not surprising that the Banach spaces c_0 and $L^1(\Omega)$ fail the ∞-MCP, as these are classically 'bad' spaces. Most spaces that arise in practise are better behaved.

Theorem 3.42. *If X is a separable dual space (i.e. X is separable and $X = Y^*$ for some Banach space Y), then X has the ∞-MCP.*

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be an X-valued martingale (on a probability space Ω) which is uniformly bounded in L^∞. By homogeneity we may assume without loss of generality that each f_n is valued in the closed unit ball of X, which by Banach–Alaoglu is weak-* compact. For each $\omega \in \Omega$, let $f(\omega)$ be a weak-* limit point of the sequence $(f_n(\omega))_{n\in\mathbb{N}}$.

Since $X = Y^*$ is separable, so is Y, and we can choose a countable dense subset $D \subset \overline{B_Y}$ of the unit ball of Y. For each $\boldsymbol{y} \in D$, the scalar-valued martingale $(\langle \boldsymbol{y}, f_n\rangle)_{n\in\mathbb{N}}$ is uniformly bounded in L^∞ and thus converges a.s. to a limit (Theorem 3.34), which must be $\langle \boldsymbol{y}, f\rangle$. For each $n \in \mathbb{N}$ and $\boldsymbol{y} \in D$ let $N_{\boldsymbol{y}} \subset \Omega$ denote the null set on which this convergence fails, let $N = \bigcup_{\boldsymbol{y}\in D} N_{\boldsymbol{y}}$ so that $\mathbb{P}(N) = 0$ (since D is countable), and observe that

$$\langle \boldsymbol{y}, f_n(\omega)\rangle \to \langle \boldsymbol{y}, f(\omega)\rangle \qquad \forall \omega \in \Omega \setminus N \quad \forall \boldsymbol{y} \in D.$$

Since D is dense in the unit ball of Y we have this convergence (away from N) for all $\boldsymbol{y} \in Y$, and since Y is weak-* dense in Y^{**} we have

add Goldstine's theorem to the appendix

$$\langle f_n(\omega), \boldsymbol{y}^{**} \rangle \to \langle f(\omega), \boldsymbol{y}^{**} \rangle \qquad \forall \omega \in \Omega \setminus N \quad \forall \boldsymbol{y}^{**} \in Y^{**} = X^*.$$

Since the functions f_n are all measurable, this shows that $f \colon \Omega \setminus N \to X$ is weakly measurable, and since X is separable, Pettis tells us that f is strongly measurable.

It remains to show that $f_n \to f$ on $\Omega \setminus N$ in the norm topology on X. For all $\omega \in \Omega \setminus N$ and $\boldsymbol{x} \in X$ we have, using that for each $\boldsymbol{y} \in D$ the scalar-valued martingale $\langle \boldsymbol{y}, \boldsymbol{x} - f_n \rangle$ converges to $\langle \boldsymbol{y}, \boldsymbol{x} - f \rangle$ on $\Omega \setminus N$,

$$\|\boldsymbol{x} - f_n(\omega)\|_X = \sup_{\boldsymbol{y} \in D} |\langle \boldsymbol{y}, \boldsymbol{x} - f_n(\omega) \rangle| = \sup_{\boldsymbol{y} \in D} |\mathbb{E}^{\mathcal{A}_n}(\langle \boldsymbol{y}, \boldsymbol{x} - f \rangle)(\omega)| \leq \mathbb{E}^{\mathcal{A}_n}(\|\boldsymbol{x} - f\|_X)(\omega)$$

using (2.7). Since the function $\|\boldsymbol{x} - f\|_X$ is bounded and measurable, we have

$$\limsup_{n\to\infty} \|\boldsymbol{x} - f_n(\omega)\|_X \leq \limsup_{n\to\infty} \mathbb{E}^{\mathcal{A}_n}(\|\boldsymbol{x} - f\|_X)(\omega) = \|\boldsymbol{x} - f(\omega)\|_X.$$

Now taking $\boldsymbol{x} = f(\omega)$ shows that $f_n(\omega) \to f(\omega)$ in X. The proof is complete. ☐

So separable dual spaces have the ∞-MCP: this includes the spaces $L^p([0,1])$ for $p \in (1,\infty)$ as well as the sequence space ℓ^1 (which is the dual of c_0). The sequence space ℓ^∞ contains c_0, which does not have p-MCP for any p as shown above, so ℓ^∞ doesn't have any of these properties either. Of course, ℓ^∞ is a dual space, but it isn't separable.

Lemma 3.43. *For all $p \in [1,\infty]$, the p-MCP is separably determined: that is, a Banach space X has the p-MCP if and only if every separable closed subspace $Y \subset X$ has the p-MCP.*

Proof. The 'only if' direction is immediate, as a Y-valued martingale can be seen as an X-valued martingale, and if a martingale converges a.s. in X, then since Y is closed it must also converge a.s. in Y.

On the other hand, suppose that every separable closed subspace $Y \subset X$ has the p-MCP, and let (f_n) be an X-valued martingale which is uniformly bounded in L^p. Each f_n is strongly measurable and hence separably-valued by the Pettis theorem (Theorem 2.4), so there is a sequence of separable closed subspaces $Y_n \subset X$ such that f_n takes values in Y_n. The union of these spaces generates a separable closed subspace Y. The martingale $(f_n)_{n\in\mathbb{N}}$ then takes values in Y, and since Y has the p-MCP by assumption, $(f_n)_{n\in\mathbb{N}}$ is a.s. convergent in Y, hence also in X. ☐

Corollary 3.44. *If X is reflexive, then X has the ∞-MCP.*

Proof. By the previous lemma, it suffices to show that every separable closed subspace $Y \subset X$ has the ∞-MCP, and by Theorem 3.42 we just need to show that every such Y is a dual space.

Consider the annihilator

$$Y^\perp := \{\boldsymbol{x}^* \in X^* : \langle \boldsymbol{y}, \boldsymbol{x}^* \rangle = 0 \text{ for all } y \in Y\},$$

and the double annihilator

$$Y^{\perp\perp} = (Y^\perp)^\perp = \{\boldsymbol{x}^{**} \in X^{**} : \langle \boldsymbol{z}, \boldsymbol{x}^{**} \rangle = 0 \text{ for all } z \in Y^\perp\}.$$

By Proposition 9.3 we know that $Y^{\perp\perp}$ is isometrically isomorphic to the dual space $X^*/Y^\perp$, so it suffices to show that $j(Y) = Y^{\perp\perp}$, where $j \colon X \to X^{**}$ is the canonical inclusion. The containment $j(Y) \subset Y^{\perp\perp}$ is a direct consequence of the definition. To show the reverse inclusion, suppose that $\boldsymbol{x}^{**} \notin j(Y)$; we will conclude that $\boldsymbol{x}^{**} \notin Y^{\perp\perp}$. To do this we need to find a functional $\boldsymbol{x}^* \in Y^\perp$ such that $\langle \boldsymbol{x}^*, \boldsymbol{x}^{**} \rangle \neq 0$. Since X is reflexive, $\boldsymbol{x}^{**} = j(\boldsymbol{x})$ for some $\boldsymbol{x} \in X \setminus Y$. By Hahn–Banach there exists a functional $\boldsymbol{x}^* \in X^*$ such that $\langle \boldsymbol{x}, \boldsymbol{x}^* \rangle = 1$ and $\boldsymbol{x}^* \in Y^\perp$. Since $\langle \boldsymbol{x}, \boldsymbol{x}^* \rangle = \langle \boldsymbol{x}^*, j(\boldsymbol{x}) \rangle = \langle \boldsymbol{x}^*, \boldsymbol{x}^{**} \rangle$, the functional $\boldsymbol{x}^*$ does exactly what we want. ☐

Thus reflexive spaces have the ∞-MCP, even if they are not separable.

Exercises.

Exercise 11. Let $(f_n)_{n\in\mathbb{N}}$ be a stochastic process on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Suppose that (f_n) is predictable with respect to the filtration generated by (f_n) (see Example 3.7). Show that the process is deterministic, in the sense that each f_n is constant.

Exercise 12. In the setting of Example 3.8, show that the random variable $\boldsymbol{s}_{n+1}$ is $\mathcal{F}_n$-measurable if and only if $\boldsymbol{x}_{n+1} \equiv 0$.

Exercise 13. This exercise takes place in the setting of Example 3.10.

- Let $X = \ell^\infty(\mathbb{N})$. Suppose that the wager vectors $\boldsymbol{x}_n \colon \Omega \to \ell^\infty(\mathbb{N})$ are such that for all $\omega \in \Omega$, the vectors $(\boldsymbol{x}_n(\omega))_{n\in\mathbb{N}}$ are pairwise distinct standard basis vectors (i.e. $\{0,1\}$-valued sequences, zero for all but one index). Fix $\lambda > 0$ and let $K = \{\boldsymbol{a} \in \ell^\infty(\mathbb{N}) : \|\boldsymbol{a}\|_\infty \geq \lambda\} = \ell^\infty(\mathbb{N}) \setminus B_\lambda(0)$. Show that the stopping time
$$T_K(\omega) := \inf\{n \in \mathbb{N} : \boldsymbol{s}_n(\omega) \in K\}$$
is finite if and only if $\lambda \leq 1$.
- As above, but now let $X = \ell^2(\mathbb{N})$, and show that the stopping time T_K is finite for all $\lambda > 0$.

Exercise 14. Use the defining property (3.1) of conditional expectations (i.e. do not use details of its construction) to prove Proposition 3.16.

Exercise 15. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $\mathcal{B}$ a σ-subalgebra of $\mathcal{A}$. Using that $L^1(\mathcal{A}) \subsetneq L^\infty(\mathcal{A})^*$, show that the adjoint of the inclusion map $\iota\colon L^\infty(\mathcal{B}) \to L^\infty(\mathcal{A})$, which *a priori* maps $L^1(\mathcal{A}) \to L^\infty(\mathcal{B})^* \supsetneq L^1(\mathcal{B})$, actually maps into $L^1(\mathcal{B})$ *without invoking the existence of a conditional expectation operator on* L^1.

Exercise 16. Prove Proposition 3.18.

Exercise 17. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, X a Banach space, and let $(M_n)_{n\in\mathbb{N}}$ be a martingale with respect to some filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$.

- Show that $M_n = \mathbb{E}^{\mathcal{A}_n} M_m$ for all $n, m \in \mathbb{N}$ with $m > n$,
- Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be the filtration generated by $(M_n)_{n\in\mathbb{N}}$, i.e.
$$\mathcal{F}_n := \sigma(M_0, M_1, \dots, M_n).$$
Show that $(M_n)_{n\in\mathbb{N}}$ is a martingale with respect to $(\mathcal{F}_n)_{n\in\mathbb{N}}$.
- For all $p \in [1, \infty]$, show that the sequence $\|M_n\|_{L^p(\Omega,\mathbb{P};X)}$ is monotonically increasing in n.

add: martingales in bijective correspondence with their difference sequences

add: martingales = independent increments

Exercise 18. Prove the pointwise convergence theorem for martingales, Theorem 3.31, as a consequence of Doob's maximal inequality and the L^p-convergence theorem (Theorem 3.23).

Exercise 19. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Show that a bounded subset $\mathcal{F} \subset L^1(\Omega)$ is uniformly integrable if and only if
$$\lim_{t>0} \sup_{f\in\mathcal{F}} \int_{|f|>t} |f(\omega)| \, \mathrm{d}\mathbb{P}(\omega) = 0.$$

Exercise 20. Show that the $L^1(\Omega; P)$-valued stochastic process defined in (3.6) is a martingale.

Exercise 21. Modify Example 3.41 to show that $L^1([0,1])$ does not have the ∞-martingale convergence property. (Do not simply use that $L^1([0,1])$ is isometrically isomorphic to $L^1(\Omega)$—construct a 'bad' martingale directly.)

4. The Radon–Nikodym property

We now move from Banach-valued analysis and probability to Banach-valued *measure theory*, and finally to the *geometry* of Banach spaces. We will tie these concepts together via the Radon–Nikodym property, which is ostensibly a measure-theoretic property but has equivalent characterisations in terms of Bochner spaces, martingales, and convex sets.

4.1. **Vector measures and the Radon–Nikodym property.**

Definition 4.1. Let X be a Banach space and $(S, \mathcal{A})$ a measurable space (recall: $\mathcal{A}$ is a σ-algebra on S). An X-valued *vector measure* is a function $\mu \colon \mathcal{A} \to X$ which is countably additive, in the sense that for all sequences $(E_n)_{n\in\mathbb{N}}$ of pairwise disjoint sets in $\mathcal{A}$,

$$\mu\Big(\bigcup_{n\in\mathbb{N}} E_n\Big) = \sum_{n\in\mathbb{N}} \mu(E_n).$$

Note that this condition includes convergence of the series on the right hand side in X.

Vector measures are just like measures, except the measure of a set $E \subset S$ is now a vector $\mu(E) \in X$ rather than a scalar. We are most interested in vector measures with the following boundedness condition.

Definition 4.2. Let X be a Banach space and μ an X-valued vector measure on a maesurable space $(S, \mathcal{A})$. The *variation* of μ is the scalar-valued measure $|\mu| \colon \mathcal{A} \to [0, \infty]$ defined by

$$|\mu|(E) := \sup_{\pi} \sum_{A\in\pi} \|\mu(A)\|_X,$$

where the supremum ranges over all partitions π of S into $\mathcal{A}$-measurable sets. We define the *total variation norm* $\|\mu\|_{\mathrm{var}} := |\mu|(S)$, and we say that μ has *bounded variation* if $\|\mu\|_{\mathrm{var}} < \infty$. Equivalently, μ has bounded variation if there exists a finite scalar-valued measure ν on $\mathcal{A}$ such that $\|\mu(A)\|_X \leq \nu(A)$ for all $A \in \mathcal{A}$ (as the minimal measure with this property is $|\mu|$).

It is not particularly difficult to define integrals of scalar-valued functions with respect to vector measures.

Proposition 4.3. *Let X be a Banach space and μ an X-valued vector measure of bounded variation on a measurable space $(S, \mathcal{A})$. Then there is a unique continuous linear map $[\mu] \colon L^1(S, \mathcal{A}, |\mu|) \to X$ such that $[\mu](\mathbb{1}_A) = \mu(A)$ for all $A \in \mathcal{A}$. We use integral notation to denote this map, i.e. we write*

$$\int_S f(s)\,\mathrm{d}\mu(s) := [\mu](f) \qquad \forall f \in L^1(S, \mathcal{A}, |\mu|).$$

Proof. We skip the verification that the prescription $[\mu](\mathbb{1}_A)$ extends by linearity to a consistent map on integrable simple functions.[19] We just need to show boundedness, and the conclusion will follow by density. Consider a simple function $g \in L^1(S, \mathcal{A}, |\mu|)$ of the form

$$g = \sum_{n=1}^{N} c_n \mathbb{1}_{S_n}$$

[19] "It is dreadfully boring to show that this formula defines a linear map... from the space of simple functions of the above form into X and we leave this as an exercise for masochists." [2, pp5-6]

with scalars $c_n \in \mathbb{K}$. Then

$$\|[\mu](g)\|_X \leq \sum_{n=1}^{N} |c_n| \|\mu(S_n)\|_X \leq \sum_{n=1}^{N} |c_n| |\mu|(S_n) = \|g\|_{L^1(|\mu|)}.$$

That's all. □

Example 4.4. Let $(S, \mathcal{A})$ be a measurable space and X a Banach space. Suppose ν is a finite scalar-valued measure on $(S, \mathcal{A})$ and $f \in L^1(S, \mathcal{A}, \nu; X)$. Then we can define an X-valued vector measure μ (sometimes denoted $\mu = f\nu$) by Bochner integration:

$$\mu(A) = \int_A f(s)\, \mathrm{d}\nu.$$

This vector measure has bounded variation: given a partition $S = \bigcup_{n\in\mathbb{N}} S_n$, we compute

$$\sum_{n\in\mathbb{N}} \|\mu(S_n)\|_X = \sum_{n\in\mathbb{N}} \left\| \int_{S_n} f(s)\, \mathrm{d}\nu \right\|_X \leq \int_S \|f(s)\|_X \, \mathrm{d}\nu$$

so that $\|\mu\|_{\mathrm{var}} \leq \|f\|_{L^1(\nu;X)}$.[20]

Now let's revise some measure theory. Recall that if μ and ν are two scalar-valued signed measures on a measurable space $(S, \mathcal{A})$, then we say ν *is absolutely continuous with respect to* μ, written $\nu \ll \mu$, if $A \in \mathcal{A}$ and $\mu(A) = 0$ implies $\nu(A) = 0$.

Theorem 4.5 (Radon–Nikodym)**.** *Let $(S, \mathcal{A})$ be a measurable space, and let μ be a σ-finite measure on $\mathcal{A}$. Let ν be a finite signed measure on $\mathcal{A}$ such that $\nu \ll \mu$. Then there exists a unique $h \in L^1(S, \mathcal{A}, \mu)$ such that*

$$\nu(A) = \int_A h(s)\, \mathrm{d}\mu(s) \qquad \forall A \in \mathcal{A}.$$

See [3, Theorem 5.5.4] for a proof. One might expect that an analogous theorem holds for vector measures, but it turns out to depend on the geometry of the target Banach space.

Definition 4.6. Let $(S, \mathcal{A}, \mu)$ be a σ-finite measure space. A Banach space X is said to have the *Radon–Nikodym property (RNP) with respect to* $(S, \mathcal{A}, \mu)$ if for every X-valued vector measure ν on $(S, \mathcal{A})$ such that $\|\nu\|_{\mathrm{var}} < \infty$ and $|\nu| \ll \mu$, there is a function $f \in L^1(S, \mathcal{A}, \mu; X)$ such that $\nu = f\mu$ (as defined in Example 4.4). We say X has the *Radon–Nikodym property* if it has the property above with respect to every σ-finite measure space $(S, \mathcal{A}, \mu)$.

The classical Radon–Nikodym theorem says that the scalar fields $\mathbb{R}$ and $\mathbb{C}$ have the RNP. We will investigate this property for other Banach spaces by considering its relationship with martingales and with properties of convex sets. We will also connect it with the duality of Bochner spaces $L^p(S; X)$, answering a question left open in Chapter 2. Before moving on we record a simple reduction.

sufficient to consider finite measure

Proposition 4.7.

[20] In fact, this is an equality. See [6, pp43].

4.2. **The RNP and martingale convergence.** First we connect the Radon–Nikodym property to the martingale convergence properties introduced in the previous chapter. This is done by means of the following vector measures associated with uniformly integrable martingales.

Theorem 4.8. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X a Banach space. Let $(f_n)_{n\in\mathbb{N}}$ be an X-valued L^1-bounded uniformly integrable martingale with respect to a filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$. Then there exists an X-valued vector measure on $\mathcal{A}$ with the following properties:*

- $\mu(A) = \int_A f_n \,\mathrm{d}\mathbb{P}$ *for all* $A \in \mathcal{A}_n$,
- $\|\mu\|_{\mathrm{var}} \le \sup_n \|f_n\|_{L^1(\Omega;X)}$,
- μ *is absolutely continuous with respect to* $\mathbb{P}$.

Proof. For all $A \in \mathcal{A}$ we wish to define

$$\mu(A) := \lim_{k\to\infty} \int_A f_k \,\mathrm{d}\mathbb{P}, \tag{4.1}$$

but it is not immediate that this limit exists. If $A \in \mathcal{A}_n$ then for $k \ge n$ we have

$$\int_A f_k \,\mathrm{d}\mathbb{P} = \int_A f_n \,\mathrm{d}\mathbb{P}$$

by the martingale property, so at least for $A \in \mathcal{A}_n$ the limit exists and equals $\int_A f_n \,\mathrm{d}\mathbb{P}$, establishing the first desired property. For a general $A \in \mathcal{A}$, for each $k \in \mathbb{N}$ we have

$$\int_A f_k \,\mathrm{d}\mathbb{P} = \mathbb{E}(f_k \mathbb{1}_A) = \mathbb{E}(\mathbb{E}^{\mathcal{A}_k}(f_k \mathbb{1}_A)) = \mathbb{E}(f_k \mathbb{E}^{\mathcal{A}_k}(\mathbb{1}_A)).$$

make sure this property is listed somewhere

Thus to show that the limit in (4.1) exists, we need to show that the sequence $(\mathbb{E}(f_k \mathbb{E}^{\mathcal{A}_k}(\mathbb{1}_A)))_{k\in\mathbb{N}}$ is Cauchy. Let $\varphi_{k,\ell} = \mathbb{E}^{\mathcal{A}_k}(\mathbb{1}_A) - \mathbb{E}^{\mathcal{A}_\ell}(\mathbb{1}_A)$. For $k < \ell$ we have, using conditional expectation magic,

$$\begin{aligned}
\mathbb{E}(f_k \mathbb{E}^{\mathcal{A}_k}(\mathbb{1}_A)) - \mathbb{E}(f_\ell \mathbb{E}^{\mathcal{A}_\ell}(\mathbb{1}_A)) &= \mathbb{E}\Big(f_k \mathbb{E}^{\mathcal{A}_k}(\mathbb{1}_A) - f_\ell \mathbb{E}^{\mathcal{A}_\ell}(\mathbb{1}_A)\Big) \\
&= \mathbb{E}\Big(\mathbb{E}^{\mathcal{A}_k}(f_\ell \mathbb{E}^{\mathcal{A}_k}(\mathbb{1}_A)) - \mathbb{E}^{\mathcal{A}_k}(f_\ell \mathbb{E}^{\mathcal{A}_\ell}(\mathbb{1}_A))\Big) \\
&= \mathbb{E}\Big(f_\ell \mathbb{E}^{\mathcal{A}_k}(\mathbb{1}_A) - f_\ell \mathbb{E}^{\mathcal{A}_\ell}(\mathbb{1}_A)\Big) = \mathbb{E}(f_\ell \varphi_{k,\ell}).
\end{aligned}$$

Thus we get

$$\|\mathbb{E}(f_k \mathbb{E}^{\mathcal{A}_k}(\mathbb{1}_A)) - \mathbb{E}(f_\ell \mathbb{E}^{\mathcal{A}_\ell}(\mathbb{1}_A))\|_X \le \|f_\ell \varphi_{k,\ell}\|_{L^1(\Omega;X)}.$$

Since $\|\varphi_{k,\ell}\|_\infty \le 2$ for all k and ℓ, for all $t > 0$ we have

$$\begin{aligned}
\|f_\ell \varphi_{k,\ell}\|_{L^1(\Omega;X)} &\le \Big(\int_{\|f_\ell\|_X > t} + \int_{\|f_\ell\|_X \le t} \Big) \|f_\ell(\omega)\|_X |\varphi_{k,\ell}(\omega)| \,\mathrm{d}\mathbb{P}(\omega) \\
&\le \Big(2 \int_{\|f_\ell\|_X > t} \|f_\ell(\omega)\|_X \mathrm{d}\mathbb{P}(\omega) + t\mathbb{E}|\varphi_{k,\ell}|\Big)
\end{aligned}$$

so that

$$\begin{aligned}
\|f_\ell \varphi_{k,\ell}\|_{L^1(\Omega;X)} &\le 2 \limsup_{t\to 0} \int_{\|f_\ell\|_X > t} \|f_\ell(\omega)\|_X \mathrm{d}\mathbb{P}(\omega) \\
&\le 2 \limsup_{t\to 0} \sup_{j\in\mathbb{N}} \int_{\|f_j\|_X > t} \|f_j(\omega)\|_X \mathrm{d}\mathbb{P}(\omega).
\end{aligned}$$

Uniform integrability of (f_j) says exactly that the the right hand side here is zero (see Exercise 19), which establishes that the limit in (4.1) exists.

We still need to show that μ is actually a vector measure. It is clear from the definition that it is finitely additive, but we need countable additivity. Consider the submartingale $(\|f_n\|_X)_{n\in\mathbb{N}}$: this is L^1-bounded and uniformly integrable, so by Theorem 3.36 it has an L^1-limit $g \in L^1(\Omega)$. Thus for all $A \in \mathcal{A}$

$$\|\mu(A)\|_X \leq \lim_{n\to\infty} \int_A \|f_n(\omega)\| \, \mathrm{d}\mathbb{P}(\omega) = \int_A g(\omega) \, \mathrm{d}\mathbb{P}(\omega).$$

By Exercise 24, this implies that μ is countably additive with

$$\|\mu\|_{\mathrm{var}} \leq \|g\mathbb{P}\|_{\mathrm{var}} = \|g\|_{L^1(\Omega)} = \sup_{n\in\mathbb{N}} \|f_n\|_{L^1(\Omega;X)},$$

and that

$$|\mu| \ll g\mathbb{P} \ll \mathbb{P},$$

as required. □

Now we can show the claimed connection between the RNP and martingale convergence.

Theorem 4.9. *Let X be a Banach space which has the Radon–Nikodym property with respect to a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Then X has the 1-martingale convergence property with respect to $(\Omega, \mathcal{A}, \mathbb{P})$.*

Proof. Fix an L^1-bounded uniformly integrable X-valued martingale $(f_n)_{n\in\mathbb{N}}$, with respect to a filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$. By Theorem 4.8, there exists an X-valued vector measure μ on $\mathcal{A}$ such that

$$\mu(A) = \int_A f_n \, \mathrm{d}\mathbb{P} \qquad \forall A \in \mathcal{A}_n$$

and $\mu \ll \mathbb{P}$. Since X has the RNP with respect to $(\Omega, \mathcal{A}, \mathbb{P})$ there exists a function $f \in L^1(\Omega, \mathbb{P}; X)$ such that

$$\int_A f \, \mathrm{d}\mathbb{P} = \mu(A) = \int_A f_n \, \mathrm{d}\mathbb{P}$$

for all $A \in \mathcal{A}_n$. Equivalently stated, we have

$$\mathbb{E}^{\mathcal{A}_n} f = f_n$$

for all $n \in \mathbb{N}$, and thus by Theorem 3.31 f_n is almost surely convergent to f. Thus X has the 1-martingale convergence property, and the proof is complete. □

Corollary 4.10. *The spaces c_0 and $L^1([0,1])$ do not have the RNP.*

Proof. In Examples 3.40 and 3.41 we showed that these spaces do not have the ∞-MCP, and hence they do not have the 1-MCP. □

In summary, for all $p \in (1, \infty]$ we currently have the implications

$$\mathrm{RNP} \Longrightarrow 1-\mathrm{MCP} \Longrightarrow p-\mathrm{MCP} \Longrightarrow \infty-\mathrm{MCP}$$

where these properties are taken either universally or with respect to a given probability space. In the next section we will add more properties and 'complete the loop'.

4.3. **Trees and dentability.**

Definition 4.11. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X a Banach space. Given $\delta > 0$, an X-valued L_1-bounded martingale $(f_n)_{n\in\mathbb{N}}$ on Ω is called *δ-separated* if the following properties hold:

- f_0 is constant,
- each f_n takes only finitely many values,
- for all $n \in \mathbb{N}$ and $\omega \in \Omega$, $\|f_n(\omega) - f_{n+1}(\omega)\|_X \geq \delta$.

The set $S := \{f_n(\omega) : n \in \mathbb{N}, \omega \in \Omega\}$ is called a *δ-separated tree.*

The martingales described in Examples 3.40 and 3.41 (see also Exercise 21) are 1-separated, and thus yield 1-separated trees in c_0 and L_1. When one tries to draw a δ-separated tree on a piece of paper, one quickly starts to run out of space. This is because pieces of paper model finite-dimensional Banach spaces, which have good martingale convergence properties.

Proposition 4.12. *If a Banach space X has the ∞-MCP, then for all $\delta > 0$, X does not contain a bounded δ-separated tree.*

Proof. Bounded δ-separated trees correspond to L^∞-bounded martingales such that

$$\|f_n(\omega) - f_{n+1}(\omega)\|_X,$$

for all $\omega \in \Omega$ and $n \in \mathbb{N}$, which directly obstructs convergence of (f_n) everywhere in Ω. □

Thus our chain of implications now takes the form

$$\mathrm{RNP} \Longrightarrow 1 - \mathrm{MCP} \Longrightarrow p - \mathrm{MCP} \Longrightarrow \infty - \mathrm{MCP} \Longrightarrow \mathrm{NBST}$$

where NBST stands for 'no bounded separated trees'.[21] The connection between (nonexistence of) bounded separated trees and the Radon–Nikodym property will go through the concept of *dentable sets.*

Definition 4.13. A subset $D \subset X$ of a Banach space X is called *dentable* if for all $\varepsilon > 0$ there exists $\boldsymbol{x} \in D$ such that

$$\boldsymbol{x} \notin \overline{\mathrm{conv}}(D \setminus B_\varepsilon(\boldsymbol{x}))$$

where $\overline{\mathrm{conv}}$ denotes the closure of the convex hull.

Before going further let us prove a lemma which relates 'non-dentability at scale ε' to a corresponding property of an enlarged set which does not involve closures.

Lemma 4.14. *Let X be a Banach space. Fix $\varepsilon > 0$ and let $D \subset B$ be a subset such that for all $\boldsymbol{x} \in D$,*

$$\boldsymbol{x} \in \overline{\mathrm{conv}}(D \setminus B_\varepsilon(\boldsymbol{x})). \tag{4.2}$$

Then for all $\boldsymbol{x} \in \tilde{D} := D + B_{\varepsilon/2}(0)$,

$$\boldsymbol{x} \in \mathrm{conv}(\tilde{D} \setminus B_{\varepsilon/2}(\boldsymbol{x})) \tag{4.3}$$

(note that no closure is taken here).

Proof. Fix $\boldsymbol{x} = \boldsymbol{x}' + \boldsymbol{y} \in \tilde{D}$, where $\boldsymbol{x}' \in D$ and $\|\boldsymbol{y}\|_X < \varepsilon/2$. Choose $\delta > 0$ so small that $\delta + \|\boldsymbol{y}\|_X < \varepsilon/2$. By the assumed property (4.2), there exists $n \in \mathbb{N}$, scalars $\alpha_i \in [0, 1]$, $i = 1, \ldots, n$, and vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in D \setminus B_\varepsilon(\boldsymbol{x})$ such that

$$\boldsymbol{x}' = \boldsymbol{z} + \sum_{i=1}^n \alpha_i \boldsymbol{x}_i.$$

[21] As with MCP, this is not standard terminology: ultimately it's just equivalent to RNP.

for some $\boldsymbol{z} \in B_\delta(0)$. We then have

$$\boldsymbol{x} = \boldsymbol{z} + \boldsymbol{y} + \sum_{i=1}^n \alpha_i \boldsymbol{x}_i = \sum_{i=1}^n \alpha_i (\boldsymbol{z} + \boldsymbol{y} + \boldsymbol{x}_i).$$

The points $\boldsymbol{z} + \boldsymbol{y} + \boldsymbol{x}_i$ are in $\tilde{D} \setminus B_{\varepsilon/2}(\boldsymbol{x})$: indeed, we have

$$\|\boldsymbol{z} + \boldsymbol{y}\|_X \leq \delta + \|\boldsymbol{y}\|_X < \varepsilon/2$$

by the choice of δ, and

$$\|\boldsymbol{x} - (\boldsymbol{z} + \boldsymbol{y} + \boldsymbol{x}_i)\|_X = \|\boldsymbol{x}' - \boldsymbol{z} - \boldsymbol{x}_i\|_X \geq \|\boldsymbol{x}' - \boldsymbol{x}_i\|_X - \|\boldsymbol{z}\|_X > \varepsilon - \delta > \varepsilon/2,$$

finishing the job. □

Theorem 4.15. *Let X be a Banach space, and suppose that for all $\delta > 0$, X does not contain a bounded δ-separated tree. Then every bounded subset of X is dentable.*

Proof. We prove the contrapositive: we suppose that there exists a bounded non-dentable set $D \subset X$, and for given $\delta > 0$ we will construct a bounded δ-separated tree. Since D is non-dentable, there exists $\varepsilon > 0$ such that for all $\boldsymbol{x} \in D$,

$$\boldsymbol{x} \in \overline{\mathrm{conv}}(D \setminus B_{2\varepsilon}(\boldsymbol{x})).$$

By Lemma 4.14, for all $\boldsymbol{x} \in \tilde{D} := D + B_\varepsilon(0)$,

$$\boldsymbol{x} \in \mathrm{conv}(\tilde{D} \setminus B_\varepsilon(\boldsymbol{x})).$$

We will construct a ε-separated tree in the bounded set $\tilde{D}$, and by rescaling this will yield the result.

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be the interval $[0,1]$ with Borel σ-algebra and Lebesgue measure. We construct a ε-separated martingale inductively. Let $\boldsymbol{x}_0 \in \tilde{D}$ be arbitrary and $f_0 \equiv \boldsymbol{x}_0$. Since $\boldsymbol{x}_0 \in \mathrm{conv}(\tilde{D} \setminus B_\varepsilon(\boldsymbol{x}_0))$, there exist numbers $\alpha_1, \ldots, \alpha_n \in (0,1)$ summing to 1 and vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \tilde{D}$ such that

$$\boldsymbol{x}_0 = \sum_{i=1}^n \alpha_i \boldsymbol{x}_i \qquad \text{and} \qquad \|\boldsymbol{x}_i - \boldsymbol{x}_0\|_X \geq \varepsilon.$$

Partition the unit interval $[0,1]$ into intervals $(I_i)_{i=1}^n$ with length $|I_i| = \alpha_i$, let $\mathcal{A}_0$ be the trivial σ-algebra on $[0,1]$, and let $\mathcal{A}_1$ be the σ-algebra generated by the intervals $(I_i)_{i=1}^n$. Define

$$f_1 = \sum_{i=1}^n \alpha_i \mathbb{1}_{I_i} \otimes \boldsymbol{x}_i.$$

Then $\mathbb{E}^{\mathcal{A}_0} f_1 = f_0$, and $\|f_1(\omega) - f_0(\omega)\|_X \geq \varepsilon$ for all $\omega \in [0,1]$. Since each point $\boldsymbol{x}_i$ is in $\tilde{D}$, we can repeat this process inductively, representing each $\boldsymbol{x}_i$ as a convex combination of vectors in $\tilde{D} \setminus B_\varepsilon(\boldsymbol{x}_i)$, using these vectors to define f_2 on I_i, and so on, to construct a ε-separated martingale valued in $\tilde{D}$, and thus a bounded ε-separated tree. □

Finally we will 'complete the loop' in our discussion of the Radon–Nikodym property, martingale convergence, and dentability.

Theorem 4.16. *Let X be a Banach space such that every bounded subset of X is dentable. Then X has the Radon–Nikodym property.*

Proof. Let $(S, \mathcal{A}, \mu)$ be a finite measure space, and let $\boldsymbol{\nu} \colon \mathcal{A} \to X$ be an X-valued vector measure with $\|\boldsymbol{\nu}\|_X \leq \mu$ and $|\boldsymbol{\nu}| \ll \mu$. Our task is to find a function $f \in L^1(S, \mathcal{A}, \mu; X)$ such that $\boldsymbol{\nu} = f\mu$. By Proposition 4.7 this is sufficient to prove that X has the RNP.

For all sets $\alpha \in \mathcal{A}$ let

$$\mathcal{A}_+(\alpha) := \{\beta \in \mathcal{A} : \beta \subset \alpha, \mu(\beta) > 0\},$$
$$\mathcal{A}_+ := \mathcal{A}_+(S)$$

and for all $\alpha \in \mathcal{A}_+$ define

$$\boldsymbol{x}_\alpha := \mu(\alpha)^{-1}\boldsymbol{\nu}(\alpha) \in X, \qquad C_\alpha := \{\boldsymbol{x}_\beta : \beta \in \mathcal{A}_+(\alpha)\} \subset X.$$

Note that $\|\boldsymbol{x}_\alpha\|_X \leq 1$ for all $\alpha \in \mathcal{A}_+$ by the assumptions on $\boldsymbol{\nu}$, so every C_α is bounded and hence (by assumption) dentable.

We make the following claim: *for all* $\varepsilon > 0$ *and* $\alpha \in \mathcal{A}_+$, *there exists* $\alpha' \in \mathcal{A}_+(\alpha)$ *such that* $\operatorname{diam}(C_\alpha) \leq 2\varepsilon$.

We assume this is not the case and establish a contradiction. Thus there exist $\varepsilon > 0$ and $\alpha \in \mathcal{A}_+$ such that $\operatorname{diam}(C_{\alpha'}) > 2\varepsilon$ for all $\alpha' \in \mathcal{A}_+(\alpha)$. In particular, for every $\boldsymbol{x} \in X$ and $\alpha' \in \mathcal{A}_+(\alpha)$, there is a subset $\beta \in \mathcal{A}_+(\alpha')$ such that $\|\boldsymbol{x} - \boldsymbol{x}_\beta\|_X > \varepsilon$.[22]

Now consider a fixed $\alpha' \in \mathcal{A}_+(\alpha)$ and let $\{\beta_\lambda\}_{\lambda\in\Lambda}$ be a maximal collection of disjoint measurable elements of $\mathcal{A}_+(\alpha')$ such that $\|\boldsymbol{x}_{\alpha'} - \boldsymbol{x}_{\beta_\lambda}\|_X > \varepsilon$, where Λ is some indexing set. Since the sets β_λ are disjoint and have positive measure, and since

$$\sum_{\lambda\in\Lambda} \mu(\beta_\lambda) \leq \mu(\alpha') < \infty,$$

the indexing set Λ is at most countable. By construction we must have

$$\mu\Big(\alpha' \setminus \bigcup_{\lambda\in\Lambda} \beta_\lambda\Big) = 0; \tag{4.4}$$

otherwise we could find a set $\beta_! \in \mathcal{A}_+(\alpha' \setminus \cup_\lambda \beta_\lambda) \subset \mathcal{A}_+(\alpha')$ such that $\|\boldsymbol{x}_{\alpha'} - \boldsymbol{x}_{\beta_!}\|_X > \varepsilon$, contradicting the maximality of the set $\{\beta_\lambda\}$. Since $\boldsymbol{\nu} \ll \mu$, this yields

$$\boldsymbol{\nu}\Big(\alpha' \setminus \bigcup_{\lambda\in\Lambda} \beta_\lambda\Big) = 0,$$

or equivalently (using countable additivity)

$$\boldsymbol{\nu}(\alpha') = \sum_{\lambda\in\Lambda} \boldsymbol{\nu}(\beta_\lambda).$$

This lets us write

$$\begin{aligned}\boldsymbol{x}_{\alpha'} &= \mu(\alpha')^{-1}\boldsymbol{\nu}(\alpha')\\ &= \sum_{\lambda\in\Lambda} \mu(\alpha')^{-1}\boldsymbol{\nu}(\beta_\lambda)\\ &= \sum_{\lambda\in\Lambda} \frac{\mu(\beta_\lambda)}{\mu(\alpha')}\boldsymbol{x}_{\beta_\lambda}.\end{aligned}$$

By (4.4) we have that the coefficients of this series sum to 1, and the vectors in the series satisfy

$$\boldsymbol{x}_{\beta_\lambda} \in C_{\alpha'} \qquad \text{and} \qquad \|\boldsymbol{x}_{\alpha'} - \boldsymbol{x}_{\beta_\lambda}\|_X > \varepsilon,$$

which tells us that

$$\boldsymbol{x}_{\alpha'} \in \overline{\operatorname{conv}}(C_{\alpha'} \setminus B_\varepsilon(\boldsymbol{x}_{\alpha'})).$$

Since this is true for all $\alpha' \in \mathcal{A}_+(\alpha)$, we find that C_α is not dentable. This is a contradiction, which implies that our claim above is true.

Now we return to the construction of a Radon–Nikodym derivative f of $\boldsymbol{\nu}$ with respect to μ. Fix $\varepsilon > 0$. Using the claim we just established, let $(\alpha_\lambda)_{\lambda\in\Lambda}$ be a

[22] Otherwise there would exist a vector $\tilde{\boldsymbol{x}} \in X$ with $\|\tilde{\boldsymbol{x}} - \boldsymbol{x}_\beta\|_X \leq \varepsilon$ for all $\beta \in \mathcal{A}_+(\alpha')$, which implies $\|\boldsymbol{x}_\beta - \boldsymbol{x}_{\beta'}\|_X \leq 2\varepsilon$ for all $\beta, \beta' \in \mathcal{A}_+(\alpha')$ and hence $\operatorname{diam}(C_{\alpha'}) \leq 2\varepsilon$. Contradiction.

maximal disjoint collection of sets in $\mathcal{A}_+$ such that $\operatorname{diam}(C_{\alpha_\lambda}) \leq 2\varepsilon$. Then Λ is at most countable (by the same argument used in the last paragraph) and

$$\mu(S \setminus \bigcup_{\lambda \in \Lambda} \alpha_\lambda) = 0;$$

if this were not the case then we could select $\alpha' \in \mathcal{A}_+(S \setminus \cup_{\lambda\in\Lambda}\alpha_\lambda) \subset \mathcal{A}_+$ with $\operatorname{diam}(C_{\alpha'}) \leq 2\varepsilon$ (using the claim) and contradict maximality. Define

$$g_\varepsilon := \sum_{\lambda \in \Lambda} \mathbb{1}_{\alpha_\lambda} \otimes \boldsymbol{x}_{\alpha_\lambda}.$$

Then $g_\varepsilon \in L^1(S, \mathcal{A}, \mu; X)$ (since it is bounded and the measure is finite). We will show that

$$\|\boldsymbol{\nu} - g_\varepsilon \mu\|_{\mathrm{var}} \leq 2\mu(S)\varepsilon; \tag{4.5}$$

since this holds for all $\varepsilon > 0$, we find that $\boldsymbol{\nu}$ is in the closure in $M(S, \mathcal{A}; X)$ of the set of measures of the form $g\mu$ with $g \in L^1(S, \mathcal{A}, \mu; X)$. But this set is closed in $M(S, \mathcal{A}; X)$, so there exists $g \in L^1(S, \mu; X)$ with $\boldsymbol{\nu} = g\mu$, as desired.

provide the discussion that makes this argument work

It remains to show (4.5). To see this first note that for all $\alpha \in \mathcal{A}_+$

$$\begin{aligned}
\boldsymbol{\nu}(\alpha) - g_\varepsilon \mu(\alpha) &= \sum_{\lambda \in \Lambda} \Big(\boldsymbol{\nu}(\alpha \cap \alpha_\lambda) - \int_{\alpha \cap \alpha_\lambda} g_\varepsilon \, \mathrm{d}\mu \Big) \\
&= \sum_{\lambda \in \Lambda} \Big(\boldsymbol{\nu}(\alpha \cap \alpha_\lambda) - \mu(\alpha \cap \alpha_\lambda) \boldsymbol{x}_{\alpha_\lambda} \Big) \\
&= \sum_{\lambda \in \Lambda} \mu(\alpha \cap \alpha_\lambda)(\boldsymbol{x}_{\alpha \cap \alpha_\lambda} - \boldsymbol{x}_{\alpha_\lambda}),
\end{aligned}$$

and so

$$\|\boldsymbol{\nu}(\alpha) - g_\varepsilon \mu(\alpha)\|_X \leq \sum_{\lambda \in \Lambda} \mu(\alpha \cap \alpha_\lambda) \|\boldsymbol{x}_{\alpha \cap \alpha_\lambda} - \boldsymbol{x}_{\alpha_\lambda}\|_X \leq 2\mu(\alpha)\varepsilon$$

using that $\operatorname{diam}(C_{\alpha_\lambda}) \leq 2\varepsilon$. Taking the supremum over $\alpha \in \mathcal{A}_+$ proves (4.5) and completes the proof. $\square$

Combining this with everything else we know, we have proven the following theorem.

Theorem 4.17. *The following properties of a Banach space X are equivalent:*

- *X has the Radon–Nikodym property;*
- *X has the p-martingale convergence property for all $p \in [1, \infty]$;*
- *X has the ∞-martingale convergence property;*
- *for all $\delta > 0$, X does not contain a bounded δ-separated tree;*
- *every bounded subset of X is dentable.*

Remark 4.18. It is possible to prove that the ∞-MCP implies the RNP directly, but I don't think this is as nice as going via dentability. See [6, Proof of Theorem 2.9].

This set of equivalences says quite a bit. First, it says that a.s. convergence of L^p-bounded martingales holds for some $p \in [1, \infty]$ if and only if it holds for all $p \in [1, \infty]$. This p-independence of martingale-based Banach space properties turns out to be fairly typical; martingales satisfy miraculous extrapolation properties of this kind. Second, note that the first four properties are 'extrinsic': the RNP makes reference to all σ-finite measure spaces, and the MCP properties and the nonexistence of bounded separated trees make reference to martingales valued in X. In contrast, the last property is an intrinsic geometric property of X. It is always satisfying to find an intrinsic geometric characterisation of what seems to be an extrinsic property. One more point: by carefully looking at the proofs of these implications, one can show that it suffices to have the RNP with respect to the unit interval $[0,1]$ in order to show that every bounded subset X is dentable. This argument shows that a Banach space has the RNP if and only if it has the RNP with respect to the unit interval (see Exercise 26).

Let's rattle off some consequences of this theorem.

Corollary 4.19. *Reflexive spaces and separable dual spaces have the RNP.*

Proof. By Theorem 3.42 and Corollary 3.44, these spaces have the ∞-MCP. Thus they also have the RNP. ☐

Corollary 4.20. *The RNP is separably determined, i.e. it holds for a Banach space X if and only if it holds for all separable subspaces $Y \subset X$.*

Proof. By Lemma 3.43, this is true for the p-MCP (for all $p \in [1,\infty]$, but these properties are now known to be equivalent anyway). ☐

Remark 4.21. Not satisfied? See [2, §VII.6] for 29 characterisations of the Radon–Nikodym property.

4.4. **Duality of Bochner spaces.** Recall Proposition 2.12: for every σ-finite measure space $(S,\mathcal{A},\mu)$ and every Banach space X, for all $p \in [1,\infty]$, the map $\Phi\colon L^{p'}(S;X^*) \to L^p(S;X)^*$ given by

$$\Phi g(f) = \int_S \langle f(s), g(s)\rangle \,\mathrm{d}\mu(s) \qquad \forall f \in L^p(S;X)$$

is an isometry onto a closed norming subspace of $L^p(S;X)^*$. We will now complete this result with the help of the Radon–Nikodym property.

Theorem 4.22. *A dual space X^* has the Radon–Nikodym property if and only if for any countably generated probability space $(\Omega,\mathcal{A},\mathbb{P})$ and any $p \in [1,\infty)$ the isometric embedding $\Phi\colon L^{p'}(\Omega;X^*) \to L^p(\Omega;X)^*$ is an isomorphism.*

Proof. First suppose X^* has the RNP and let $(\Omega,\mathcal{A},\mathbb{P})$ be a countably generated probability space. Then there exists a filtration $(\mathcal{A}_n)_{n\in\mathbb{N}}$ of finite sub-σ-algebras of $\mathcal{A}$ such that $\mathcal{A}_\infty = \mathcal{A}$. Let $\varphi \in L^p(\mathcal{A};X)^*$ be a bounded linear functional. For each $n \in \mathbb{N}$ consider the restriction maps

$$R_n\colon L^p(\mathcal{A};X)^* \to L^p(\mathcal{A}_n;X)^* \qquad \text{and} \qquad r_n\colon L^p(\mathcal{A}_{n+1};X)^* \to L^p(\mathcal{A}_n;X)^*,$$

and define $\varphi_n := R_n\varphi \in L^p(\mathcal{A}_n;X)^*$. Then

$$r_n\varphi_{n+1} = r_n R_{n+1}\varphi = R_n\varphi = \varphi_n. \tag{4.6}$$

Now for each n consider the isometric embedding

$$\Phi_n\colon L^{p'}(\mathcal{A}_n;X^*) \to L^p(\mathcal{A}_n;X)^*.$$

Since each $\mathcal{A}_n$ is finite, each Φ_n is an isomorphism (Exercise 4), so there exist functions $g_n \in L^{p'}(\mathcal{A}_n;X^*)$ such that $\Phi_n g_n = \varphi_n$. The equality (4.6) yields

$$\Phi_n^{-1} r_n \Phi_{n+1} g_{n+1} = g_n.$$

We are going to show that $\Phi_n^{-1} r_n \Phi_{n+1} g_{n+1} = \mathbb{E}^{\mathcal{A}_n} g_{n+1}$, thus proving that the sequence $(g_n)_{n\in\mathbb{N}}$ is a martingale. ☐ up to here, want a clean proof of this

Exercises.

Exercise 22. Let X be a Banach space with the Radon–Nikodym property. Show that every bounded linear operator $T\colon L^1([0,1]) \to X$ is of the form

$$Tf = \int_0^1 f(t)g(t)\,\mathrm{d}t$$

for some $g \in L^1([0,1];X)$.

Exercise 23. Let X be a Banach space with the Radon–Nikodym property. Use this property to show that for every measure space $(S,\mathcal{A},\mu)$ and every sub-σ-algebra $\mathcal{B} \subset \mathcal{A}$, every $f \in L^1(\mathcal{A};X)$ has a conditional expectation with respect to $\mathcal{B}$. (Of course this holds for *every* Banach space, but your job here is to derive it in a simpler way under the RNP assumption.)

Exercise 24. Let X be a Banach space and $(S,\mathcal{A})$ a measure space. Let ν be a finitely additive X-valued vector measure on $(S,\mathcal{A})$, i.e. a function $\nu\colon \mathcal{A} \to X$ such that

$$\nu\Big(\sum_{n=1}^N E_n\Big) = \sum_{n=1}^N \nu(E_n)$$

for all $N \in \mathbb{N}$ and $E_1,\dots,E_n \in \mathcal{A}$. Suppose that there is a *countably additive* (scalar) measure μ on $(S,\mathcal{A})$ such that

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

Show that ν is countably additive, $\|\nu\|_{\mathrm{var}} \le \|\mu\|_{\mathrm{var}}$, and $|\nu| \ll \mu$.

Exercise 25 (Rademacher's theorem and the RNP)**.** Show that a Banach space X has the Radon–Nikodym property if and only if every X-valued Lipschitz function on $[0,1]$ is differentiable almost everywhere.

Exercise 26. Suppose that X has the Radon–Nikodym property with respect to the unit interval $[0,1]$ with Borel σ-algebra and Lebesgue measure. Show that X has the Radon–Nikodym property with respect to all σ-finite measure spaces.

5. The class of UMD spaces

5.1. **Unconditional sequences in Banach spaces.**

5.2. **The UMD property.** defn; easy examples; independence of p

5.3. p**-independence of the UMD property; Gundy's decomposition.**

5.4. **Dyadic UMD (is just UMD).**

5.5. **UMD and reflexivity.** UMD implies reflexive, hence also RNP

5.6. **Examples and counterexamples.**

6. Type, cotype, and Fourier type

6.1. **(Rademacher) type and cotype.**

- definitions and examples
- K-convexity: implied by UMD; implies type/cotype duality
- STATE that K-convexity iff nontrivial type (Maurey–Pisier theorem is too hard for this course)
- gaussian sums, orthonormal invariance/covariance domination. STATE the equivalence with rademacher for finite cotype

- detecting Hilbert spaces through type and cotype 2 (have to do the Lindenstrauss reduction to f.d. subspaces, and the LIndenstrauss–Pelczynski Theorem 7.3)

6.2. **Fourier type.**

- Fourier type with respect to $\mathbb{R}$, $\mathbb{T}$, and $\mathbb{Z}$; examples (e.g. by interpolation or Fubini)
- UMD implies nontrivial Fourier type (can't find an elementary proof)
- FT with respect to arbitrary lca group, in particular $\mathbb{Z}/N\mathbb{Z}$ and Cantor/Walsh group
- Fourier type 2 implies type and cotype 2
- STATE relations with type and cotype (some too subtle to prove)

7. Fourier multipliers and Littlewood–Paley theory

7.1. **The Fourier transform on Banach-valued functions.** We define the Fourier transform $\hat{f} \in C(\mathbb{R})$ of an integrable function $f \in L^1(\mathbb{R};\mathbb{C})$ by

$$\hat{f}(\xi) := \int_{\mathbb{R}} f(t)e^{-2\pi i t\xi}\,\mathrm{d}t.$$

- plancherel, dancing hat
- Fourier transform on L^2
- Hausdorff-Young
- tensor extension problem; direct definition on $L^1(\mathbb{R};X)$

7.2. **Fourier multipliers and transference.**

7.3. **The Hilbert transform and the HT property.**

7.4. **HT implies UMD.** (one-two lectures of work)

7.5. **UMD implies HT.** (use the argument in HNVW1 - two lectures of work)

8. Schatten class operators

- definition and basic properties
- example operators?
- the UMD property
- Schur multipliers
- Operator Lipschitz functions, Potapov–Sukochev theorem

9. Appendices

9.1. **Results from functional analysis.** Throughout this section X is a Banach space over the scalar field $\mathbb{K}$ (which is either $\mathbb{R}$ or $\mathbb{C}$). X^* denotes the dual Banach space, i.e. the Banach space of all bounded $\mathbb{K}$-linear functionals $\boldsymbol{x}^*\colon X \to \mathbb{K}$, under the norm

$$\|\boldsymbol{x}^*\|_{X^*} := \sup_{\boldsymbol{x}\in X\setminus\{0\}} \frac{|\boldsymbol{x}^*(\boldsymbol{x})|}{\|\boldsymbol{x}\|_X} = \sup_{\substack{\boldsymbol{x}\in X\\ \|\boldsymbol{x}\|_X=1}} |\boldsymbol{x}^*(\boldsymbol{x})|.$$

Often we write $\langle \boldsymbol{x}, \boldsymbol{x}^*\rangle := \boldsymbol{x}^*(\boldsymbol{x})$.

The following results are all standard the suggested references are essentially picked out at random.

[7, Section III.3]

Theorem 9.1 (Hahn–Banach: real case)**.** *Let X be a real Banach space. Let $p\colon X \to \mathbb{R}$ be a real-valued function satisfying*

$$p(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) \leq \alpha p(\boldsymbol{x}_1) + (1-\alpha)p(\boldsymbol{x}_2)$$

for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in X$ and all $\alpha \in [0,1]$. Let λ be an $\mathbb{R}$-linear functional defined on a subspace $Y \subset X$, satisfying

$$\lambda(\boldsymbol{y}) \leq p(\boldsymbol{y}) \qquad \forall \boldsymbol{y} \in Y.$$

Then there exists a functional $\Lambda \in X^*$ such that $\Lambda(\boldsymbol{x}) \leq p(\boldsymbol{x})$ for all $\boldsymbol{x} \in X$ and $\Lambda(\boldsymbol{y}) = \lambda(\boldsymbol{y})$ for all $\boldsymbol{y} \in Y$.

Theorem 9.2 (Hahn–Banach: complex case)**.** *Let X be a complex Banach space. Let $p\colon X \to \mathbb{R}$ be a real-valued function satisfying*

$$p(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) \leq |\alpha| p(\boldsymbol{x}_1) + |\beta| p(\boldsymbol{x}_2)$$

for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in X$ and all $\alpha, \beta \in \mathbb{C}$ with $|\alpha| + |\beta| = 1$. Let λ be a $\mathbb{C}$-linear functional defined on a subspace $Y \subset X$, satisfying

$$|\lambda(\boldsymbol{y})| \leq p(\boldsymbol{y}) \qquad \forall \boldsymbol{y} \in Y.$$

Then there exists a functional $\Lambda \in X^*$ such that $|\Lambda(\boldsymbol{x})| \leq p(\boldsymbol{x})$ for all $\boldsymbol{x} \in X$ and $\Lambda(\boldsymbol{y}) = \lambda(\boldsymbol{y})$ for all $\boldsymbol{y} \in Y$.

The Hahn–Banach theorem implies duality relations between subspaces and quotient spaces.

Proposition 9.3. *Let Y be a closed subspace of X, and define the annihilator*

$$Y^\perp := \{\boldsymbol{x}^* \in X^* : \langle \boldsymbol{y}, \boldsymbol{x}^* \rangle \text{ for all } \boldsymbol{y} \in Y\}.$$

Then $Y^\perp$ is a closed subspace of X^*, and we have isometric isomorphisms

$$X^*/Y^\perp = M^* \qquad \text{and} \qquad (X/Y)^* = M^\perp :$$

the first is given by restriction to Y, and the second is given by precomposition with the quotient map $X \to X/Y$.

The *weak topology* on X the weakest topology on X such that every functional $\boldsymbol{x}^* \in X^*$ is continuous. The weak toplogy is weaker than the usual (norm) topology on X, and these topologies are equal if and only if X is finite dimensional.

The *double dual* of X is the space $X^{**} = (X^*)^*$. Each $\boldsymbol{x} \in X$ induces an element of the double dual in a notationally confusing way: for each $\boldsymbol{x}^* \in X^*$, $\boldsymbol{x}$ acts on $\boldsymbol{x}^*$ by

$$\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle := \boldsymbol{x}^*(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{x}^* \rangle.$$

This identification yields a canonical isometric embedding $X \to X^{**}$, and X is called *reflexive* if this map is surjective (i.e. each functional on X^* is given by an element of X, using the above identification). In this case X and X^{**} are canonically isomorphic.[23]

The *weak-** topology on a dual space X^* is the weakest topology such that all the functions

$$\{\boldsymbol{x}^* \mapsto \boldsymbol{x}^*(\boldsymbol{x}) : \boldsymbol{x} \in X\}$$

are continuous. The weak-* topology is great because it has the following fundamental compactness property.

[7, Theorem IV.21]

Theorem 9.4 (Banach–Alaoglu)**.** *The closed unit ball of X^* is compact in the weak-* topology.*

[23] It is possible that X and X^{**} are isometrically isomorphic without X being reflexive [5].

One can show that X is reflexive if and only if the weak topology and the weak-* topology inherited from X^{**} coincide. Using Banach–Alaoglu, this can be restated in the following way.

Corollary 9.5. *X is reflexive if and only if the closed unit ball B_X is weakly compact.*

There are three notions of compactness in topological spaces that coincide for metric spaces: *compactness* (every open cover has a finite subcover), *sequential compactness* (every sequence has a convergent subsequence), and *countable compactness* (every sequence has a cluster point, or equivalently, every countable open cover has a finite subcover). Although the weak topology on X may not be metrisable, it behaves as if it were:

Theorem 9.6 (Eberlein–Smulian)**.** *Let A be a subset of a Banach space X. The following are equivalent:*

- *A is weakly compact,*
- *A is weakly sequentially compact,*
- *A is weakly countably compact.*

And thus Corollary 9.5 can be restated as follows:

Corollary 9.7. *X is reflexive if and only if every bounded sequence has a weakly convergent subsequence.*

A Banach space is *separable* if it has a countable dense subset. One can show via the Hahn–Banach theorem that if X^* is separable, then so is X [7, Theorem III.7].

Theorem 9.8. *X is separable if and only if the closed unit ball $\overline{B_{X^*}}$ of X^* is weak-* metrisable. Thus by Banach–Alaoglu and the equivalence of compactness notions for metrisable spaces, if X is separable, then $\overline{B_{X^*}}$ is weak-* sequentially compact.*

need a citation for this theorem

9.2. Results from probability theory.

- independence of random variables and σ-algebras
- mutual independence of sequences of RVs

9.3. Complex interpolation?

- defn
- basic structural results
- (co)retraction theorem
- examples

10. Misc. sections to move

10.1. John–Nirenberg for adapted sequences and the Kahane–Khintchine inequalities. Fix a Banach space X and a σ-finite measure space $(S, \mathcal{A}, \mu)$. Consider a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ and a sequence $\varphi = (\varphi_n)_{n\in\mathbb{N}}$ of X-valued functions adapted to the filtration (i.e. φ_n is $\mathcal{F}_n$-measurable for all $n \in \mathbb{N}$). For all $q \in (0,\infty)$ we consider the following measure of the oscillation of φ:

put this all on a probability space

Strongly?

$$\|\varphi\|_{*,q} := \sup_{\substack{k,n\in\mathbb{N}\\ k\le n}} \sup_{\substack{F\in\mathcal{F}_k\\ 0<\mu(F)<\infty}} \Big(\fint_F \|(\varphi_n - \varphi_{k-1})(s)\|_X^q \,\mathrm{d}\mu(s)\Big)^{1/q}.$$

Theorem 10.1 (John–Nirenberg inequality for adapted sequences)**.** *With notation as above, for all $p, q \in (0,\infty)$ there exists a finite constant $c_{p,q}$ independent of φ such that*

$$\|\varphi\|_{*,p} \le c_{p,q}\|\varphi\|_{*,q}.$$

We will prove this as a consequence of a series of lemmas, but before proving this, we demonstrate an important application to Rademacher sums.

Theorem 10.2 (Kahane–Khintchine inequality)**.** *Let X be a Banach space and let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a Rademacher sequence on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then for all $p, q \in (0,\infty)$ there exists a finite constant $\kappa_{p,q}$ such that for all finite sequences $(\boldsymbol{x}_n)_{n=1}^N$ in X,*

$$\Big\| \sum_{n=1}^N \varepsilon_n \boldsymbol{x}_n \Big\|_{L^p(\Omega;X)} \leq \kappa_{p,q} \Big\| \sum_{n=1}^N \varepsilon_n \boldsymbol{x}_n \Big\|_{L^q(\Omega;X)}.$$

That is, for all $p \in (0,\infty)$, the L^p-norms of a Rademacher sum are pairwise equivalent.

Since Ω is a probability space, Hölder's inequality yields the case $p \leq q$ with constant $\kappa_{p,q} = 1$, so we only need to consider the case $q < p$.

make sure to discuss filtrations generated by functions, particularly Rademacher variables

Proof, assuming the John–Nirenberg inequality. Consider the filtration and adapted sequence

$$\mathcal{F}_n := \sigma(\{\varepsilon_j : 1 \leq j \leq n\}), \qquad \varphi_n := \sum_{j=1}^n \varepsilon_j \boldsymbol{x}_j.$$

We claim that for all $q \in [1,\infty)$ we have

$$\|\varphi\|_{*,q} = \Big\| \sum_{j=1}^N \varepsilon_j \boldsymbol{x}_j \Big\|_{L^q(\Omega;X)}. \tag{10.1}$$

Assuming this for the moment, the John–Nirenberg inequality yields a finite constant $c_{p,q}$ such that

$$\Big\| \sum_{j=1}^N \varepsilon_j \boldsymbol{x}_j \Big\|_{L^p(\Omega;X)} = \|\varphi\|_{*,p} \leq c_{p,q} \|\varphi\|_{*,q} = \Big\| \sum_{j=1}^N \varepsilon_j \boldsymbol{x}_j \Big\|_{L^q(\Omega;X)}$$

whenever $1 \leq q < p$. If $q < 1 \leq p$, we fix $\theta \in (0,1)$ so that $1/p = \theta/q + (1-\theta)/2p$, and log-convexity of L^p-norms gives

$$\|\varphi_N\|_{L^p(\Omega;X)} \leq \|\varphi_N\|_{L^q(\Omega;X)}^\theta \|\varphi_N\|_{L^{2p}(\Omega;X)}^{1-\theta} \leq \|\varphi_N\|_{L^q(\Omega;X)}^\theta (c_{2p,p}\|\varphi_N\|_{L^p(\Omega;X)})^{1-\theta},$$

which yields

$$\|\varphi_N\|_{L^p(\Omega;X)} \leq c_{2p,p}^{(1-\theta)/\theta} \|\varphi_N\|_{L^q(\Omega;X)}.$$

Finally, if $q < p < 1$, we simply estimate

$$\|\varphi_N\|_{L^p(\Omega;X)} \leq \|\varphi_N\|_{L^1(\Omega;X)} \leq c_{1,q} \|\varphi_N\|_{L^q(\Omega;X)}.$$

This covers all nontrivial cases, and it remains to prove the claimed equality (10.1).

First recall that

$$\|\varphi\|_{*,q} = \sup_{\substack{k,n\in\mathbb{N} \\ k \leq n}} \sup_{\substack{F \in \mathcal{F}_k \\ \mu(F)>0}} \Big(\fint_F \|(\varphi_n - \varphi_{k-1})(s)\|_X^q \, \mathrm{d}\mathbb{P}(s) \Big)^{1/q}.$$

Fix $k \leq n$ and a set $F \in \mathcal{F}_k$ of positive measure. We will compute

$$\fint_F \|\mathbb{1}_F(\varphi_n - \varphi_{k-1})(s)\|_X^q \, \mathrm{d}\mathbb{P}(s) = \mathbb{P}(F)^{-1} \mathbb{E}\Big(\mathbb{1}_F \Big\| \sum_{j=k}^n \varepsilon_j \boldsymbol{x}_j \Big\|_X^q \Big).$$

Observe that for all $\omega \in \Omega$

$$\Big\| \sum_{j=k}^n \varepsilon_j(\omega) \boldsymbol{x}_j \Big\|_X = \Big\| \varepsilon_k(\omega) \Big(\boldsymbol{x}_k + \sum_{j=k+1}^n \varepsilon_j'(\omega) \boldsymbol{x}_j \Big) \Big\|_X = \Big\| \boldsymbol{x}_k + \sum_{j=k+1}^n \varepsilon_j' \boldsymbol{x}_j \Big\|_X$$

where

$$\varepsilon'_j := \begin{cases} \varepsilon_j & j \le k \\ \varepsilon_k \varepsilon_j & k+1 \le j, \end{cases}$$

since ε_j is ± 1-valued. Now note that the σ-algebra $\mathcal{F}'_{k+1,n} := \sigma(\{\varepsilon'_j : k+1 \le j \le n\})$ is independent of $\mathcal{F}_k$, since for all $1 \le j \le k$ and $k+1 \le j' \le n$ we have by independence of the original Rademacher sequence

$$\mathbb{E}(\varepsilon_j \varepsilon'_{j'}) = \mathbb{E}(\varepsilon_j \varepsilon_k \varepsilon_{j'}) = \begin{cases} \mathbb{E}(\varepsilon_j)\mathbb{E}(\varepsilon_k)\mathbb{E}(\varepsilon_{j'}) & \text{if } j < k \\ \mathbb{E}(\varepsilon_{j'}) & \text{if } j = k \end{cases} = 0.$$

Thus, since $F \in \mathcal{F}_k$, we have by independence of $\mathcal{F}_k$ and $\mathcal{F}'_{k+1,n}$

$$\begin{aligned} \mathbb{P}(F)^{-1}\mathbb{E}\Big(\mathbb{1}_F \Big\| \sum_{j=k}^n \varepsilon_j \boldsymbol{x}_j \Big\|_X^q\Big) &= \mathbb{P}(F)^{-1}\mathbb{E}\Big(\mathbb{1}_F \Big\| \boldsymbol{x}_k + \sum_{j=k+1}^n \varepsilon'_j \boldsymbol{x}_j \Big\|_X^q\Big) \\ &= \mathbb{E}\Big\| \boldsymbol{x}_k + \sum_{j=k+1}^n \varepsilon'_j \boldsymbol{x}_j \Big\|_X^q = \mathbb{E}\Big\| \sum_{j=k}^n \varepsilon_j \boldsymbol{x}_j \Big\|_X^q. \end{aligned}$$

Now letting $\mathcal{F}_{k,n} := \sigma(\{\varepsilon_j : k \le j \le n\})$, we have by the L^q-contraction property of conditional expectations

$$\mathbb{E}\Big\| \sum_{j=k}^n \varepsilon_j \boldsymbol{x}_j \Big\|_X^q = \mathbb{E}\Big\| \mathbb{E}^{\mathcal{F}_{k,n}}\Big(\sum_{j=1}^N \varepsilon_j \boldsymbol{x}_j \Big) \Big\|_X^q \le \mathbb{E}\Big\| \sum_{j=1}^N \varepsilon_j \boldsymbol{x}_j \Big\|_X^q$$

with equality when $k = 1$ and $n = N$. Thus

$$\|\varphi\|_{*,q} = \sup_{\substack{k,n \in \mathbb{N} \\ k \le n}} \Big(\mathbb{E}\Big\| \sum_{j=k}^n \varepsilon_j \boldsymbol{x}_j \Big\|_X^q \Big)^{1/q} = \Big(\mathbb{E}\Big\| \sum_{j=1}^N \varepsilon_j \boldsymbol{x}_j \Big\|_X^q \Big)^{1/q}$$

which proves the claimed equality (10.1) and completes the proof. □

In the setting of Hilbert-valued functions (and in particular, scalar-valued functions), the Kahane–Khintchine inequality leads to the classical Khintchine inequalities.

Corollary 10.3 (Khintchine's inequalities)**.** *Let H be a Hilbert space, and let $(\varepsilon_n)_{n \in \mathbb{N}}$ be a Rademacher sequence on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then for all $p \in (0, \infty)$ there exist finite constants A_p and B_p such that for all finite sequences $(\boldsymbol{h}_n)_{n=1}^N$ in H,*

$$A_p \Big(\sum_{n=1}^N \|\boldsymbol{h}_n\|_H^2 \Big)^{1/2} \le \Big\| \sum_{n=1}^N \varepsilon_n \boldsymbol{h}_n \Big\|_{L^p(\Omega;H)} \le B_p \Big(\sum_{n=1}^N \|\boldsymbol{h}_n\|_H^2 \Big)^{1/2}.$$

Proof. By independence of the Rademacher variables we have

$$\begin{aligned} \Big\| \sum_{n=1}^N \varepsilon_n \boldsymbol{h}_n \Big\|_{L^2(\Omega;H)}^2 &= \Big\langle \sum_{n=1}^N \varepsilon_n \boldsymbol{h}_n, \sum_{m=1}^N \varepsilon_m \boldsymbol{h}_m \Big\rangle \\ &= \sum_{n,m=1}^N \mathbb{E}(\varepsilon_n \varepsilon_m) \langle \boldsymbol{h}_n, \boldsymbol{h}_m \rangle = \sum_{n=1}^N \|\boldsymbol{h}_n\|_H^2, \end{aligned} \tag{10.2}$$

so the result is true for $p = 2$ with $A_2 = B_2 = 1$. Now use Kahane–Khintchine to extend the result to general $p \in (0, \infty)$. □

10.1.1. *Proof of the John–Nirenberg inequality.* We return to our analysis of a sequence $\varphi = (\varphi_n)_{n\in\mathbb{N}}$ of X-valued functions adapted to a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ on a σ-finite measure space $(S, \mathcal{A}, \mu)$. We prove the John–Nirenberg inequality via a series of lemmas in which we obtain increasingly fine control on the oscillation of φ.

Lemma 10.4. *For all $k \leq n$, $F \in \mathcal{F}_k$, and $\alpha > 0$,*

$$\mu(F \cap \{\|\varphi_n - \varphi_{k-1}\|_X > \alpha\}) \leq \Big(\frac{\|\varphi\|_{*,q}}{\alpha}\Big)^q \mu(F). \tag{10.3}$$

Proof. We can assume that $0 < \mu(F) < \infty$, otherwise there is nothing to prove. The left hand side of (10.3) is bounded by

$$\int_F \Big(\frac{\|\varphi_n - \varphi_{k-1}\|_X}{\alpha}\Big)^q \,\mathrm{d}\mu \leq \mu(F)\Big(\frac{\|\varphi\|_{*,q}}{\alpha}\Big)^q$$

since $F \in \mathcal{F}_k$ and $k \leq n$, by the definition of $\|\varphi\|_{*,q}$. □

Next we show that oscillation control of the form above extends to more general stopping times.

Lemma 10.5. *Suppose that there exist $\alpha > 0$ and $\eta > 0$ such that*

$$\mu(F \cap \{\|\varphi_n - \varphi_{k-1}\| > \alpha\}) \leq \eta\mu(F) \qquad \forall k \leq n, F \in \mathcal{F}_k.$$

Then for all $k \in \mathbb{N}$, $F \in \mathcal{F}_k$, and all stopping times ν such that $\nu \geq k$ on F,

$$\mu(F \cap \{\nu < \infty\} \cap \{\|\varphi_\nu - \varphi_{k-1}\| > 2\alpha\}) \leq 2\eta\mu(F). \tag{10.4}$$

Proof. Sum over all possible values of the stopping time:

$$\mu(F \cap \{\nu < \infty\} \cap \{\|\varphi_\nu - \varphi_{k-1}\| > 2\alpha\}) = \lim_{N\to\infty} \sum_{n=k}^{N} \mu(F_n \cap \{\|\varphi_n - \varphi_{k-1}\| > 2\alpha\}),$$

where $F_n := F \cap \{\nu = n\}$. For fixed $N \geq n > k$, since $F_n \in \mathcal{F}_n \subset \mathcal{F}_{n+1}$, we have by assumption

$$\begin{aligned}
&\mu(F_n \cap \{\|\varphi_n - \varphi_{k-1}\| > 2\alpha\}) \\
&\leq \mu(F_n \cap \{\|\varphi_n - \varphi_N\| > \alpha\}) + \mu(F_n \cap \{\|\varphi_{k-1} - \varphi_N\| > \alpha\}) \\
&\leq \eta\mu(F_n) + \mu(F_n \cap \{\|\varphi_{k-1} - \varphi_N\| > \alpha\}).
\end{aligned}$$

Thus we have

$$\begin{aligned}
&\lim_{N\to\infty} \sum_{n=k}^{N} \mu(F_n \cap \{\|\varphi_n - \varphi_{k-1}\| > 2\alpha\}) \\
&\leq \lim_{N\to\infty} \Big(\eta \sum_{n=k}^{N} \mu(F_n) + \sum_{n=k}^{N} \mu(F_n \cap \{\|\varphi_{k-1} - \varphi_N\| > \alpha\})\Big) \\
&\leq \eta\mu(F) + \lim_{N\to\infty} \mu(F \cap \{\|\varphi_{k-1} - \varphi_N\| > \alpha\}) \leq 2\eta\mu(F)
\end{aligned}$$

using the assumption and $F \in \mathcal{F}_k$ in the last estimate. □

In the following lemma we make use of the *started sequence*

$$^{k-1}\varphi = (\varphi_n - \varphi_{k-1})_{n\geq k-1}$$

and its maximal function

$$(^{k-1}\varphi)^*(s) := \sup_{n\geq k-1} \|(^{k-1}\varphi)_n(s)\|_X = \sup_{n\geq k} \|(\varphi_n - \varphi_{k-1})(s)\|_X.$$

Lemma 10.6. *Suppose that φ satisfies (10.4) for all $k \in \mathbb{N}$, $F \in \mathcal{F}_k$, and all stopping times ν such that $\nu \geq k$ on F. Then for all $\lambda > 0$,*

$$\mu(F \cap \{({}^{k-1}\varphi)^* > \lambda + 2\alpha\}) \leq 2\eta\mu(F \cap \{({}^{k-1}\varphi)^* > \lambda\}) \qquad \forall k \in \mathbb{N}, F \in \mathcal{F}_k. \tag{10.5}$$

Proof. Fix $k \in \mathbb{N}$ and consider the stopping times

$$\rho := \inf\{n \geq k : \|\varphi_n - \varphi_{k-1}\| > \lambda\},$$
$$\nu := \inf\{n \geq k : \|\varphi_n - \varphi_{k-1}\| > \lambda + 2\alpha\}.$$

Then $k \leq \rho \leq \nu$, and (10.5) can be rewritten as

$$\mu(F \cap \{\nu < \infty\}) \leq 2\eta\mu(F \cap \{\rho < \infty\}).$$

Now fix $n \geq k$ and let $F_n := F \cap \{\rho = n\} \in \mathcal{F}_n$. On $\{F_n \cap \{\nu < \infty\}\}$ we have

$$\|\varphi_\nu - \varphi_{n-1}\| \geq \|\varphi_\nu - \varphi_{k-1}\| - \|\varphi_{n-1} - \varphi_{k-1}\| > (\lambda + 2\alpha) - \lambda = 2\alpha,$$

so

$$\mu(F_n \cap \{\nu < \infty\}) = \mu(F_n \cap \{\nu < \infty\} \cap \{\|\varphi_\nu - \varphi_{n-1}\| > 2\alpha\}) \leq 2\eta\mu(F_n).$$

Summing over $n \geq k$ completes the proof. □

Lemma 10.7. *Suppose that f is a non-negative function supported in $F \in \mathcal{A}$, satisfying*

$$\mu(f > \lambda + \alpha) \leq \eta\mu(f > \lambda) \qquad \forall \lambda > 0$$

for some $\eta \in (0,1)$ and $\alpha > 0$. Then for all $p \in [1,\infty)$,

$$\|f\|_p \leq \frac{1+\eta^{1/p}}{1-\eta^{1/p}}\alpha\mu(F)^{1/p}.$$

Proof. WRITE PROOF □

UP TO HERE. HAVE TO COMPLETE PROOF OF JOHN-NIRENBERG.

$$\|\varphi\|_{**,q} := \sup_{k\in\mathbb{N}} \sup_{\substack{F\in\mathcal{F}_k \\ 0<\mu(F)<\infty}} \left(\fint_F ({}^{k-1}\varphi)^*(s)^q \, \mathrm{d}\mu(s) \right)^{1/q},$$

test citation [4]

References

1. F. Albiac and N. J. Kalton, *Topics in Banach space theory*, Graduate Texts in Mathematics, vol. 233, Springer, 2006.
2. J. Diestel and J. J. Uhl Jr., *Vector measures*, Mathematical Surveys, vol. 15, American Mathematical Society, 1977.
3. R. M. Dudley, *Real analysis and probability*, Cambridge studies in advanced mathematics, vol. 74, Cambridge University Press, 2004.
4. T. Hytönen, J. van Neerven, M. Veraar, and L. Weis, *Analysis in Banach spaces. Vol. I: Martingales and Littlewood-Paley theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge., vol. 63, Springer, Cham, 2016.
5. R. C. James, *A non-reflexive Banach space isometric with its second conjugate space*, Proc. Natl. Acad. Sci. U.S.A. **37** (1951), no. 7, 174–177.
6. G. Pisier, *Martingales in Banach spaces*, Cambridge Studies in Advanced Mathematics, vol. 155, Cambridge University Press, 2016.
7. M. Reed and B. Simon, *Methods of modern mathematical physics i: Functional analysis*, Academic Press, 1980, Revised and enlarged edition.

Mathematisches Institut
Universität Bonn, Bonn, Germany
Email address: amenta@math.uni-bonn.de