Jan van Neerven

# Stochastic Evolution Equations

ISEM Lecture Notes 2007/08

# Contents


**1 Integration in Banach spaces** ..... 1
- 1.1 Banach spaces ..... 1
- 1.2 The Pettis measurability theorem ..... 3
  - 1.2.1 Strong measurability ..... 3
  - 1.2.2 Strong $\mu$-measurability ..... 6
- 1.3 The Bochner integral ..... 7
  - 1.3.1 The Bochner integral ..... 8
  - 1.3.2 The Lebesgue-Bochner spaces $L^p(A;E)$ ..... 12
- 1.4 Exercises ..... 13

**2 Random variables in Banach spaces** ..... 15
- 2.1 Random variables ..... 15
- 2.2 Fourier transforms ..... 17
- 2.3 Convergence in probability ..... 20
- 2.4 Independence ..... 21
- 2.5 The Itô-Nisio theorem ..... 22
- 2.6 Exercises ..... 25

**3 Sums of independent random variables** ..... 29
- 3.1 Gaussian sums ..... 29
- 3.2 The Kahane-Khintchine inequality ..... 35
- 3.3 Exercises ..... 39

**4 Gaussian random variables** ..... 43
- 4.1 Fernique's theorem ..... 43
- 4.2 The covariance operator ..... 46
- 4.3 Series representation ..... 49
- 4.4 Convergence ..... 51
- 4.5 Exercises ..... 53

**5 $\gamma$-Radonifying operators** ..... 57
5.1 $\gamma$-Summing operators ..... 57
5.2 $\gamma$-Radonifying operators ..... 61
5.3 Examples of $\gamma$-radonifying operators ..... 66
5.4 Exercises ..... 69

**6 Stochastic integration I: the Wiener integral** ..... 73
6.1 Brownian motion ..... 73
6.1.1 Brownian motion ..... 73
6.1.2 Cylindrical Brownian motion ..... 77
6.2 The stochastic Wiener integral ..... 78
6.3 Exercises ..... 84

**7 Semigroups of linear operators** ..... 89
7.1 $C_0$-semigroups ..... 89
7.2 Duality ..... 94
7.3 The abstract Cauchy problem ..... 96
7.4 Examples of $C_0$-semigroups ..... 98
7.5 Exercises ..... 100

**8 Linear equations with additive noise I** ..... 105
8.1 Stochastic preliminaries ..... 105
8.2 Semigroup preliminaries ..... 108
8.3 Existence and uniqueness: cylindrical Brownian motion ..... 110
8.4 Existence and uniqueness: Brownian motion ..... 113
8.5 Non-existence ..... 116
8.6 Exercises ..... 116

**9 $\gamma$-Boundedness** ..... 121
9.1 Randomised boundedness ..... 121
9.2 Examples ..... 127
9.3 A multiplier result ..... 128
9.4 Exercises ..... 131

**10 Linear equations with additive noise II** ..... 135
10.1 Analytic semigroups ..... 135
10.2 Fractional powers ..... 139
10.3 Hölder regularity ..... 143
10.4 Exercises ..... 146

**11 Conditional expectations and martingales** ..... 151
11.1 Conditional expectations ..... 151
11.2 Vector-valued conditional expectations ..... 155
11.3 Martingales ..... 157
11.4 $L^p$-martingales ..... 160
11.5 Exercises ..... 162

**12 UMD-spaces** . . . . 165
12.1 $\mathrm{UMD}_p$-spaces . . . . 165
12.2 $p$-Independence of the $\mathrm{UMD}_p$-property . . . . 168
12.2.1 Reduction to Haar martingales . . . . 168
12.2.2 $p$-Independence for Haar martingales . . . . 172
12.3 The vector-valued Stein inequality . . . . 175
12.4 Exercises . . . . 177

**13 Stochastic integration II: the Itô integral** . . . . 181
13.1 Decoupling . . . . 181
13.2 Stochastic integration . . . . 183
13.3 Stochastic integrability of $L^p$-martingales . . . . 189
13.4 Exercises . . . . 191

**14 Linear equations with multiplicative noise** . . . . 195
14.1 $\gamma$-Lipschitz functions . . . . 196
14.2 Pisier's property . . . . 197
14.3 Stochastic convolutions . . . . 200
14.4 Existence and uniqueness . . . . 202
14.5 Space-time regularity . . . . 205
14.6 Exercises . . . . 207

**15 Applications to stochastic PDE** . . . . 211
15.1 Space-time white noise . . . . 211
15.2 The stochastic wave equation . . . . 212
15.2.1 Rank one multiplicative noise . . . . 212
15.2.2 Additive space-time white noise . . . . 214
15.3 The stochastic heat equation . . . . 216
15.3.1 Rank one multiplicative noise . . . . 216
15.3.2 Additive space-time white noise . . . . 218
15.4 Exercises . . . . 221

**References** . . . . 223

# 1

# Integration in Banach spaces

When integrating a continuous function $f : [a,b] \to E$, where $E$ is a Banach space, it usually suffices to use the Riemann integral. We shall be concerned frequently with $E$-valued functions defined on some abstract measure space (typically, a probability space), and in this context the notions of continuity and Riemann integral make no sense. For this reason we start this first lecture with generalising the Lebesgue integral to the $E$-valued setting.

## 1.1 Banach spaces

Throughout this lecture, $E$ is a Banach space over the scalar field $\mathbb{K}$, which may be either $\mathbb{R}$ or $\mathbb{C}$ unless otherwise stated. The norm of an element $x \in E$ is denoted by $\|x\|_E$, or, if no confusion can arise, by $\|x\|$. We write

$$B_E = \{x \in E : \ \|x\| \leqslant 1\}$$

for the closed unit ball of $E$.

The *Banach space dual* of $E$ is the vector space $E^*$ of all continuous linear mappings from $E$ to $\mathbb{K}$. This space is a Banach space with respect to the norm

$$\|x^*\|_{E^*} := \sup_{\|x\| \leqslant 1} |\langle x, x^* \rangle|.$$

Here, $\langle x, x^* \rangle := x^*(x)$ denotes the duality pairing of the elements $x \in E$ and $x^* \in E^*$. We shall simply write $\|x^*\|$ instead of $\|x^*\|_{E^*}$ if no confusion can arise. The elements of $E^*$ are often called (linear) *functionals* on $E$. The *Hahn-Banach separation theorem* guarantees an ample supply of functionals on $E$: for every convex closed set $C \subseteq E$ and convex compact set $K \subseteq E$ such that $C \cap K = \varnothing$ there exist $x^* \in E^*$ and real numbers $a < b$ such that

$$\mathrm{Re}\langle x, x^* \rangle \leqslant a < b \leqslant \mathrm{Re}\langle y, x^* \rangle$$

for all $x \in C$ and $y \in K$. As is well-known, from this one derives the *Hahn-Banach extension theorem*: if $F$ is a closed subspace of $E$, then for all $y^* \in F^*$ there exists an $x^* \in E^*$ such that $x^*|_F = y^*$ and $\|x^*\| = \|y^*\|$. This easily implies that for all $x \in E$ we have

$$\|x\| = \sup_{\|x^*\| \leqslant 1} |\langle x, x^* \rangle|.$$

A linear subspace $F$ of $E^*$ is called *norming* for a subset $S$ of $E$ if for all $x \in S$ we have

$$\|x\| = \sup_{\substack{x^* \in F \\ \|x^*\| \leqslant 1}} |\langle x, x^* \rangle|.$$

A subspace of $E^*$ which is norming for $E$ is simply called *norming*. The following lemma will be used frequently.

**Lemma 1.1.** *If $E_0$ is a separable subspace of $E$ and $F$ is a linear subspace of $E^*$ which is norming for $E_0$, then $F$ contains a sequence of unit vectors that is norming for $E_0$.*

*Proof.* Choose a dense sequence $(x_n)_{n=1}^\infty$ in $E_0$ and choose a sequence of unit vectors $(x_n^*)_{n=1}^\infty$ in $F$ such that $|\langle x_n, x_n^* \rangle| \geqslant (1-\varepsilon_n)\|x_n\|$ for all $n \geqslant 1$, where the numbers $0 < \varepsilon_n \leqslant 1$ satisfy $\lim_{n\to\infty} \varepsilon_n = 0$. The sequence $(x_n^*)_{n=1}^\infty$ is norming for $E_0$. To see this, fix an arbitrary $x \in E_0$ and let $\delta > 0$. Pick $n_0 \geqslant 1$ such that $0 < \varepsilon_{n_0} \leqslant \delta$ and $\|x - x_{n_0}\| \leqslant \delta$. Then,

$$\begin{aligned}(1-\delta)\|x\| \leqslant (1-\varepsilon_{n_0})\|x\| &\leqslant (1-\varepsilon_{n_0})\|x_{n_0}\| + (1-\varepsilon_{n_0})\delta \\ &\leqslant |\langle x_{n_0}, x_{n_0}^* \rangle| + \delta \leqslant |\langle x, x_{n_0}^* \rangle| + 2\delta.\end{aligned}$$

Since $\delta > 0$ was arbitrary it follows that $\|x\| \leqslant \sup_{n \geqslant 1} |\langle x, x_n^* \rangle|$. □

A linear subspace $F$ of $E^*$ is said to *separate the points* of a subset $S$ of $E$ if for every pair $x, y \in S$ with $x \neq y$ there exists an $x^* \in F$ with $\langle x, x^* \rangle \neq \langle y, x^* \rangle$. Clearly, norming subspaces separate points, but the converse need not be true.

**Lemma 1.2.** *If $E_0$ is a separable subspace of $E$ and $F$ is a linear subspace of $E^*$ which separates the points of $E_0$, then $F$ contains a sequence that separates the points of $E_0$.*

*Proof.* By the Hahn-Banach theorem, for each $x \in E_0 \setminus \{0\}$ there exists a vector $x^*(x) \in F$ such that $\langle x, x^*(x) \rangle \neq 0$. Defining

$$V_x := \{y \in E_0 \setminus \{0\} : \ \langle y, x^*(x) \rangle \neq 0\}$$

we obtain an open cover $\{V_x\}_{x \in E_0 \setminus \{0\}}$ of $E_0 \setminus \{0\}$. Since every open cover of a separable metric space admits a countable subcover it follows that there exists a sequence $(x_n)_{n=1}^\infty$ in $E_0 \setminus \{0\}$ such that $\{V_{x_n}\}_{n=1}^\infty$ covers $E_0 \setminus \{0\}$. Then the sequence $\{x^*(x_n)\}_{n=1}^\infty$ separates the points of $E_0$: indeed, every $x \in E_0 \setminus \{0\}$ belongs to some $V_{x_n}$, which means that $\langle x, x^*(x_n) \rangle \neq 0$. □

## 1.2 The Pettis measurability theorem

We begin with a discussion of weak and strong measurability of $E$-valued functions. The main result in this direction is the Pettis measurability theorem which states, roughly speaking, that an $E$-valued function is strongly measurable if and only if it is weakly measurable and takes its values in a separable subspace of $E$.

### 1.2.1 Strong measurability

Throughout this section $(A, \mathscr{A})$ denotes a *measurable space*, that is, $A$ is a set and $\mathscr{A}$ is a *$\sigma$-algebra* in $A$, that is, a collection of subsets of $A$ with the following properties:

1. $A \in \mathscr{A}$;
2. $B \in \mathscr{A}$ implies $\complement B \in \mathscr{A}$;
3. $B_1 \in \mathscr{A}, \ B_2 \in \mathscr{A}, \ \dots$ imply $\bigcup_{n=1}^{\infty} B_n \in \mathscr{A}$.

The first property guarantees that $\mathscr{A}$ is non-empty, the second expresses that $\mathscr{A}$ is closed under taking complements, and the third that $\mathscr{A}$ is closed under taking countable unions.

The *Borel $\sigma$-algebra* of a topological space $T$, notation $\mathscr{B}(T)$, is the smallest $\sigma$-algebra containing all open subsets of $T$. The sets in $\mathscr{B}(T)$ are the *Borel sets* of $T$.

**Definition 1.3.** *A function $f : A \to T$ is called* $\mathscr{A}$-measurable *if $f^{-1}(B) \in \mathscr{A}$ for all $B \in \mathscr{B}(T)$.*

The collection of all $B \in \mathscr{B}(T)$ satisfying $f^{-1}(B) \in \mathscr{A}$ is easily seen to be a $\sigma$-algebra. As a consequence, $f$ is $\mathscr{A}$-measurable if and only if $f^{-1}(U) \in \mathscr{A}$ for all open sets $U$ in $T$.

When $T_1$ and $T_2$ are topological spaces, a function $g : T_1 \to T_2$ is *Borel measurable* if $g^{-1}(B) \in \mathscr{B}(T_1)$ for all $B \in \mathscr{B}(T_2)$, that is, if $g$ is $\mathscr{B}(T_1)$-measurable. Note that if $f : A \to T_1$ is $\mathscr{A}$-measurable and $g : T_1 \to T_2$ is Borel measurable, then the composition $g \circ f : A \to T_2$ is $\mathscr{A}$-measurable. By the above observation, every continuous function $g : T_1 \to T_2$ is Borel measurable.

It is a matter of experience that the notion of $\mathscr{A}$-measurability does not lead to a satisfactory theory from the point of view of vector-valued analysis. Indeed, the problem is that this definition does not provide the means for approximation arguments. It is for this reason that we shall introduce next another notion of measurability. We shall restrict ourselves to Banach space-valued functions, although some of the results proved below can be generalised to functions with values in metric spaces.

Let $E$ be a Banach space and $(A, \mathscr{A})$ a measurable space. A function $f : A \to E$ is called *$\mathscr{A}$-simple* if it is of the form $f = \sum_{n=1}^{N} 1_{A_n} x_n$ with $A_n \in \mathscr{A}$ and $x_n \in E$ for all $1 \leqslant n \leqslant N$. Here $1_A$ denotes the *indicator function* of the set $A$, that is, $1_A(\xi) = 1$ if $\xi \in A$ and $1_A(\xi) = 0$ if $\xi \notin A$.

**Definition 1.4.** *A function $f : A \to E$ is* strongly $\mathscr{A}$-measurable *if there exists a sequence of $\mathscr{A}$-simple functions $f_n : A \to E$ such that* $\lim_{n\to\infty} f_n = f$ *pointwise on $A$.*

In order to be able to characterise strong $\mathscr{A}$-measurability of $E$-valued functions we introduce some terminology. A function $f : A \to E$ is called *separably valued* if there exists a separable closed subspace $E_0 \subseteq E$ such that $f(\xi) \in E_0$ for all $\xi \in A$, and *weakly $\mathscr{A}$-measurable* if the functions $\langle f, x^* \rangle : A \to \mathbb{K}$, $\langle f, x^* \rangle(\xi) := \langle f(\xi), x^* \rangle$, are $\mathscr{A}$-measurable for all $x^* \in E^*$.

**Theorem 1.5 (Pettis measurability theorem, first version).** *Let $(A, \mathscr{A})$ be a measurable space and let $F$ be a norming subspace of $E^*$. For a function $f : A \to E$ the following assertions are equivalent:*

(1) *$f$ is strongly $\mathscr{A}$-measurable;*
(2) *$f$ is separably valued and $\langle f, x^* \rangle$ is $\mathscr{A}$-measurable for all $x^* \in E^*$;*
(3) *$f$ is separably valued and $\langle f, x^* \rangle$ is $\mathscr{A}$-measurable for all $x^* \in F$.*

*Proof.* (1)⇒(2): Let $(f_n)_{n=1}^\infty$ be a sequence of $\mathscr{A}$-simple functions converging to $f$ pointwise and let $E_0$ be the closed subspace spanned by the countably many values taken by these functions. Then $E_0$ is separable and $f$ takes its values in $E_0$. Furthermore, each $\langle f, x^* \rangle$ is $\mathscr{A}$-measurable, being the pointwise limit of the $\mathscr{A}$-measurable functions $\langle f_n, x^* \rangle$.

(2)⇒(3): This implication is trivial.

(3)⇒(1): Using Lemma 1.1, choose a sequence $(x_n^*)_{n=1}^\infty$ of unit vectors in $F$ that is norming for a separable closed subspace $E_0$ of $E$ where $f$ takes its values. By the $\mathscr{A}$-measurability of the functions $\langle f, x_n^* \rangle$, for each $x \in E_0$ the real-valued function

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

is $\mathscr{A}$-measurable. Let $(x_n)_{n=1}^\infty$ be a dense sequence in $E_0$.

Define the functions $s_n : E_0 \to \{x_1, \dots, x_n\}$ as follows. For each $y \in E_0$ let $k(n, y)$ be the least integer $1 \leqslant k \leqslant n$ with the property that

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

and put $s_n(y) := x_{k(n,y)}$. Notice that

$$\lim_{n\to\infty} \|s_n(y) - y\| = 0 \quad \forall y \in E_0$$

since $(x_n)_{n=1}^\infty$ is dense in $E_0$. Now define $f_n : A \to E$ by

$$f_n(\xi) := s_n(f(\xi)), \qquad \xi \in A.$$

For all $1 \leqslant k \leqslant n$ we have

$$\begin{aligned}
&\{\xi \in A : \ f_n(\xi) = x_k\} \\
&\quad = \bigl\{\xi \in A : \ \|f(\xi) - x_k\| = \min_{1 \leqslant j \leqslant n} \|f(\xi) - x_j\|\bigr\} \\
&\qquad \cap \bigl\{\xi \in A : \ \|f(\xi) - x_l\| > \min_{1 \leqslant j \leqslant n} \|f(\xi) - x_j\| \text{ for } l = 1, \dots, k-1\bigr\}.
\end{aligned}$$

Note that the sets on the right hand side are in $\mathscr{A}$. Hence each $f_n$ is $\mathscr{A}$-simple, and for all $\xi \in A$ we have

$$\lim_{n\to\infty} \|f_n(\xi) - f(\xi)\| = \lim_{n\to\infty} \|s_n(f(\xi)) - f(\xi)\| = 0. \qquad \square$$

**Corollary 1.6.** *The pointwise limit of a sequence of strongly $\mathscr{A}$-measurable functions is strongly $\mathscr{A}$-measurable.*

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

**Corollary 1.7.** *If an $E$-valued function $f$ is strongly $\mathscr{A}$-measurable and $\phi : E \to F$ is continuous, where $F$ is another Banach space, then $\phi \circ f$ is strongly $\mathscr{A}$-measurable.*

*Proof.* Choose simple functions $f_n$ converging to $f$ pointwise. Then $\phi \circ f_n \to \phi \circ f$ pointwise and the result follows from the previous corollary. $\square$

**Proposition 1.8.** *For a function $f : A \to E$, the following assertions are equivalent:*

(1) *$f$ is strongly $\mathscr{A}$-measurable;*
(2) *$f$ is separably valued and for all $B \in \mathscr{B}(E)$ we have $f^{-1}(B) \in \mathscr{A}$.*

*Proof.* (1)$\Rightarrow$(2): Let $f$ be strongly $\mathscr{A}$-measurable. Then $f$ is separably-valued. To prove that $f^{-1}(B) \in \mathscr{A}$ for all $B \in \mathscr{B}(E)$ it suffices to show that $f^{-1}(U) \in \mathscr{A}$ for all open sets $U$.

Let $U$ be open and choose a sequence of $\mathscr{A}$-simple functions $f_n$ converging pointwise to $f$. For $r > 0$ let $U_r = \{x \in U : \ d(x, \complement U) > r\}$, where $\complement U$ denotes the complement of $U$. Then $f_n^{-1}(U_r) \in \mathscr{A}$ for all $n \geqslant 1$, by the definition of an $\mathscr{A}$-simple function. Since

$$f^{-1}(U) = \bigcup_{m \geqslant 1} \bigcup_{n \geqslant 1} \bigcap_{k \geqslant n} f_k^{-1}(U_{\frac{1}{m}})$$

(the inclusion '$\subseteq$' being a consequence of the fact that $U$ is open) it follows that also $f^{-1}(U) \in \mathscr{A}$.

(2)$\Rightarrow$(1): By assumption, $f$ is $\mathscr{A}$-measurable, and therefore $\langle f, x^*\rangle$ is $\mathscr{A}$-measurable for all $x^* \in E^*$. The result now follows from the Pettis measurability theorem. $\square$

Thus if $E$ is separable, then an $E$-valued function $f$ is strongly $\mathscr{A}$-measurable if and only if it is $\mathscr{A}$-measurable.

### 1.2.2 Strong $\mu$-measurability

So far, we have considered measurability properties of $E$-valued functions defined on a measurable space $(A, \mathscr{A})$. Next we consider functions defined on a *σ-finite measure space* $(A, \mathscr{A}, \mu)$, that is, $\mu$ is a non-negative measure on a measurable space $(A, \mathscr{A})$ and there exist sets $A^{(1)} \subseteq A^{(2)} \subseteq \dots$ in $\mathscr{A}$ with $\mu(A^{(n)}) < \infty$ for all $n \geqslant 1$ and $A = \bigcup_{n=1}^{\infty} A^{(n)}$.

A *μ-simple* function with values in $E$ is a function of the form

$$f = \sum_{n=1}^{N} 1_{A_n} x_n,$$

where $x_n \in E$ and the sets $A_n \in \mathscr{A}$ satisfy $\mu(A_n) < \infty$.

We say that a property holds *μ-almost everywhere* if there exists a $\mu$-null set $N \in \mathscr{A}$ such that the property holds on the complement $\complement N$ of $N$.

**Definition 1.9.** *A function $f : A \to E$ is* strongly $\mu$-measurable *if there exists a sequence $(f_n)_{n\geqslant 1}$ of μ-simple functions converging to $f$ μ-almost everywhere.*

Using the $\sigma$-finiteness of $\mu$ it is easy to see that every strongly $\mathscr{A}$-measurable function is strongly $\mu$-measurable. Indeed, if $f$ is strongly $\mathscr{A}$-measurable and $\lim_{n\to\infty} f_n = f$ pointwise with each $f_n$ an $\mathscr{A}$-simple functions, then also $\lim_{n\to\infty} 1_{A^{(n)}} f_n = f$ pointwise, where $A = \bigcup_{n=1}^{\infty} A^{(n)}$ as before, and each $1_{A^{(n)}} f_n$ is $\mu$-simple. The next proposition shows that in the converse direction, every strongly $\mu$-measurable function is equal $\mu$-almost everywhere to a strongly $\mathscr{A}$-measurable function.

Let us call two functions which agree $\mu$-almost everywhere *μ-versions* of each other.

**Proposition 1.10.** *For a function $f : A \to E$ the following assertions are equivalent:*

(1) *$f$ is strongly μ-measurable;*
(2) *$f$ has a μ-version which is strongly $\mathscr{A}$-measurable.*

*Proof.* (1)⇒(2): Suppose that $f_n \to f$ outside the null set $N \in \mathscr{A}$, with each $f_n$ $\mu$-simple. Then we have $\lim_{n\to\infty} 1_{\complement N} f_n = 1_{\complement N} f$ pointwise on $A$, and since the functions $1_{\complement N} f_n$ are $\mathscr{A}$-simple, $1_{\complement N} f$ is strongly $\mathscr{A}$-measurable. It follows that $1_{\complement N} f$ is a strongly $\mathscr{A}$-measurable $\mu$-version of $f$.

(2)⇒(1): Let $\widetilde{f}$ be a strongly $\mathscr{A}$-measurable $\mu$-version of $f$ and let $N \in \mathscr{A}$ be a null set such that $f = \widetilde{f}$ on $\complement N$. If $(\widetilde{f}_n)_{n=1}^{\infty}$ is a sequence of $\mathscr{A}$-simple functions converging pointwise to $\widetilde{f}$, then $\lim_{n\to\infty} \widetilde{f}_n = f$ on $\complement N$, which means that $\lim_{n\to\infty} \widetilde{f}_n = f$ $\mu$-almost everywhere.

Write $A = \bigcup_{n=1}^{\infty} A^{(n)}$ with $A^{(1)} \subseteq A^{(2)} \subseteq \dots \in \mathscr{A}$ and $\mu(A^{(n)}) < \infty$ for all $n \geqslant 1$. The functions $f_n := 1_{A^{(n)}} \widetilde{f}_n$ are $\mu$-simple and we have $\lim_{n\to\infty} f_n = f$ $\mu$-almost everywhere. □

We say that $f$ is *$\mu$-separably valued* if there exists a closed separable subspace $E_0$ of $E$ such that $f(\xi) \in E_0$ for $\mu$-almost all $\xi \in A$, and *weakly $\mu$-measurable* if $\langle f, x^* \rangle$ is $\mu$-measurable for all $x^* \in E^*$.

**Theorem 1.11 (Pettis measurability theorem, second version).** *Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space and let $F$ be a norming subspace of $E^*$. For a function $f : A \to E$ the following assertions are equivalent:*

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

*Proof.* The implication (1)⇒(2) follows the corresponding implication in Theorem 1.5 combined with Proposition 1.10, and (2)⇒(3) is trivial. The implication (3)⇒(1) is proved in the same way as the corresponding implication in Theorem 1.5, observing that this time the functions $f_n$ have $\mu$-versions $\widetilde{f_n}$ that are $\mathscr{A}$-simple. If we write $A = \bigcup_{n=1}^{\infty} A^{(n)}$ as before with each $A^{(n)}$ of finite $\mu$-measure, the functions $1_{A^{(n)}} \widetilde{f_n}$ are $\mu$-simple and converge to $f$ $\mu$-almost everywhere. □

By combining Proposition 1.10 with Corollaries 1.6 and 1.7 we obtain:

**Corollary 1.12.** *The $\mu$-almost everywhere limit of a sequence of strongly $\mu$-measurable $E$-valued functions is strongly $\mu$-measurable.*

**Corollary 1.13.** *If an $E$-valued function $f$ is strongly $\mu$-measurable and $\phi : E \to F$ is continuous, where $F$ is another Banach space, then $\phi \circ f$ is strongly $\mu$-measurable.*

The following result will be applied frequently.

**Corollary 1.14.** *If $f$ and $g$ are strongly $\mu$-measurable $E$-valued functions which satisfy $\langle f, x^* \rangle = \langle g, x^* \rangle$ $\mu$-almost everywhere for every $x^* \in F$, where $F$ is subspace of $E^*$ separating the points of $E$. Then $f = g$ $\mu$-almost everywhere.*

*Proof.* Both $f$ and $g$ take values in a separable closed subspace $E_0$ $\mu$-almost everywhere, say outside the $\mu$-null set $N$. Since $E_0$ is separable, by Lemma 1.2 some countable family of elements $(x_n^*)_{n=1}^{\infty}$ in $F$ separates the points of $E_0$. Since $\langle f, x_n^* \rangle = \langle g, x_n^* \rangle$ outside a $\mu$-null set $N_n$, we conclude that $f$ and $g$ agree outside the $\mu$-null set $N \cup \bigcup_{n=1}^{\infty} N_n$. □

## 1.3 The Bochner integral

The Bochner integral is the natural generalisation of the familiar Lebesgue integral to the vector-valued setting.

Throughout this section, $(A, \mathscr{A}, \mu)$ is a $\sigma$-finite measure space.

### 1.3.1 The Bochner integral

**Definition 1.15.** *A function $f : A \to E$ is* $\mu$-Bochner integrable *if there exists a sequence of $\mu$-simple functions $f_n : A \to E$ such that the following two conditions are met:*

(1) $\lim_{n\to\infty} f_n = f$ *$\mu$-almost everywhere;*
(2) $\displaystyle\lim_{n\to\infty} \int_A \|f_n - f\|\,d\mu = 0.$

Note that $f$ is strongly $\mu$-measurable. The functions $\|f_n - f\|$ are $\mu$-measurable by Corollary 1.13.

It follows trivially from the definitions that every $\mu$-simple function is $\mu$-Bochner integrable. For $f = \sum_{n=1}^{N} 1_{A_n} x_n$ we put

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

It is routine to check that this definition is independent of the representation of $f$. If $f$ is $\mu$-Bochner integrable, the limit

$$\int_A f\,d\mu := \lim_{n\to\infty} \int_A f_n\,d\mu$$

exists in $E$ and is called the *Bochner integral* of $f$ with respect to $\mu$. It is routine to check that this definition is independent of the approximating sequence $(f_n)_{n=1}^{\infty}$.

If $f$ is $\mu$-Bochner integrable and $g$ is a $\mu$-version of $f$, then $g$ is $\mu$-Bochner integrable and the Bochner integrals of $f$ and $g$ agree. In particular, in the definition of the Bochner integral the function $f$ need not be everywhere defined; it suffices that $f$ be $\mu$-almost everywhere defined.

If $f$ is $\mu$-Bochner integrable, then for all $x^* \in E^*$ we have the identity

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

For $\mu$-simple functions this is trivial, and the general case follows by approximating $f$ with $\mu$-simple functions.

**Proposition 1.16.** *A strongly $\mu$-measurable function $f : A \to E$ is $\mu$-Bochner integrable if and only if*

$$\int_A \|f\|\,d\mu < \infty,$$

*and in this case we have*

$$\Big\| \int_A f\,d\mu \Big\| \leqslant \int_A \|f\|\,d\mu.$$

*Proof.* First assume that $f$ is $\mu$-Bochner integrable. If the $\mu$-simple functions $f_n$ satisfy the two assumptions of Definition 1.15, then for large enough $n$ we obtain

$$\int_A \|f\|\,d\mu \leqslant \int_A \|f - f_n\|\,d\mu + \int_A \|f_n\|\,d\mu < \infty.$$

Conversely let $f$ be a strongly $\mu$-measurable function satisfying $\int_A \|f\|\,d\mu < \infty$. Let $g_n$ be $\mu$-simple functions such that $\lim_{n\to\infty} g_n = f$ $\mu$-almost everywhere and define

$$f_n := 1_{\{\|g_n\| \leqslant 2\|f\|\}} g_n.$$

Then $f_n$ is $\mu$-simple, and clearly we have $\lim_{n\to\infty} f_n = f$ $\mu$-almost everywhere. Since we have $\|f_n\| \leqslant 2\|f\|$ pointwise, the dominated convergence theorem can be applied and we obtain

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

The final inequality is trivial for $\mu$-simple functions, and the general case follows by approximation. □

As a simple application, note that if $f : A \to E$ is $\mu$-Bochner integrable, then for all $B \in \mathscr{A}$ the truncated function $1_B f : A \to E$ is $\mu$-Bochner integrable, the restricted function $f|_B : B \to E$ is $\mu|_B$-Bochner integrable, and we have

$$\int_A 1_B f\,d\mu = \int_B f|_B\,d\mu|_B.$$

Henceforth, both integrals will be denoted by $\int_B f\,d\mu$.

In the following result, $\text{conv}(V)$ denotes the convex hull of a subset $V \subseteq E$, i.e., the set of all finite sums $\sum_{j=1}^k \lambda_j x_j$ with $\lambda_j \geqslant 0$ satisfying $\sum_{j=1}^k \lambda_j = 1$ and $x_j \in V$ for $j = 1, \dots, k$. The closure of this set is denoted by $\overline{\text{conv}}(V)$.

**Proposition 1.17.** *Let $f : A \to E$ be a $\mu$-Bochner integrable function. If $\mu(A) = 1$, then*

$$\int_A f\,d\mu \in \overline{\text{conv}}\{f(\xi) :\ \xi \in A\}.$$

*Proof.* Let us say that an element $x \in E$ is *strictly separated* from a set $V \subseteq E$ by a functional $x^* \in E^*$ if there exists a number $\delta > 0$ such that

$$|\text{Re}\langle x, x^*\rangle - \text{Re}\langle v, x^*\rangle| \geqslant \delta \quad \forall v \in V.$$

The Hahn-Banach separation theorem asserts that if $V$ is convex and $x \notin \overline{V}$, then there exists a functional $x^* \in E^*$ which strictly separates $x$ from $V$.

For $x^* \in E^*$, let

$$m(x^*) := \inf\{\text{Re}\langle f(\xi), x^*\rangle :\ \xi \in A\},$$
$$M(x^*) := \sup\{\text{Re}\langle f(\xi), x^*\rangle :\ \xi \in A\},$$

allowing these values to be $-\infty$ and $\infty$, respectively. Then, since $\mu(A) = 1$,

$$\mathrm{Re}\Big\langle \int_A f\,d\mu, x^* \Big\rangle = \int_A \mathrm{Re}\langle f, x^*\rangle\,d\mu \in [m(x^*), M(x^*)].$$

This shows that $\int_A f\,d\mu$ cannot be strictly separated from the convex set $\mathrm{conv}\{f(\xi) : \ \xi \in A\}$ by functionals in $E^*$. Therefore the conclusion follows by an application of the Hahn-Banach separation theorem. $\square$

As a rule of thumb, results from the theory of Lebesgue integration carry over to the Bochner integral as long as there are no non-negativity assumptions involved. For example, there are no analogues of the Fatou lemma and the monotone convergence theorem, but we do have the following analogue of the dominated convergence theorem:

**Proposition 1.18 (Dominated convergence theorem).** *Let $f_n : A \to E$ be a sequence of functions, each of which is $\mu$-Bochner integrable. Assume that there exist a function $f : A \to E$ and a $\mu$-Bochner integrable function $g : A \to \mathbb{K}$ such that:*

(1) $\lim_{n\to\infty} f_n = f$ *$\mu$-almost everywhere;*
(2) $\|f_n\| \leqslant |g|$ *$\mu$-almost everywhere.*

*Then $f$ is $\mu$-Bochner integrable and we have*

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

*In particular we have*

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

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

It is immediate from the definition of the Bochner integral that if $f : A \to E$ is $\mu$-Bochner integrable and $T$ is a bounded linear operator from $E$ into another Banach space $F$, then $Tf : A \to F$ is $\mu$-Bochner integrable and

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

This identity has a useful extension to a suitable class of unbounded operators. A linear operator $T$, defined on a linear subspace $\mathscr{D}(T)$ of $E$ and taking values in another Banach space $F$, is said to be *closed* if its *graph*

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

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

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

and $T$ is a bounded operator from $\mathscr{D}(T)$ to $E$.

The *closed graph theorem* asserts that if $T : E \to F$ is a closed operator with domain $\mathscr{D}(T) = E$, then $T$ is bounded.

**Theorem 1.19 (Hille).** *Let $f : A \to E$ be $\mu$-Bochner integrable and let $T$ be a closed linear operator with domain $\mathscr{D}(T)$ in $E$ taking values in a Banach space $F$. Assume that $f$ takes its values in $\mathscr{D}(T)$ $\mu$-almost everywhere and the $\mu$-almost everywhere defined function $Tf : A \to F$ is $\mu$-Bochner integrable. Then $\int_A f\,d\mu \in \mathscr{D}(T)$ and*

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

*Proof.* We begin with a simple observation which is a consequence of Proposition 1.16 and the fact that the coordinate mappings commute with Bochner integrals: if $E_1$ and $E_2$ are Banach spaces and $f_1 : A \to E_1$ and $f_2 : A \to E_2$ are $\mu$-Bochner integrable, then $f = (f_1, f_2) : A \to E_1 \times E_2$ is $\mu$-Bochner integrable and

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

Turning to the proof of the proposition, by the preceding observation the function $g : A \to E \times F$, $g(\xi) := (f(\xi), Tf(\xi))$, is $\mu$-Bochner integrable. Moreover, since $g$ takes its values in the graph $\mathscr{G}(T)$, we have $\int_A g(\xi)\,d\mu(\xi) \in \mathscr{G}(T)$. On the other hand,

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

The result follows by combining these facts. $\square$

We finish this section with a result on integration of $E$-valued functions which may fail to be Bochner integrable.

**Theorem 1.20 (Pettis).** *Let $(A, \mathscr{A}, \mu)$ be a finite measure space and let $1 < p < \infty$ be fixed. If $f : A \to E$ is strongly $\mu$-measurable and satisfies $\langle f, x^*\rangle \in L^p(A)$ for all $x^* \in E^*$, then there exists a unique $x_f \in E$ satisfying*

$$\langle x_f, x^*\rangle = \int_A \langle f, x^*\rangle\,d\mu.$$

*Proof.* We may assume that $f$ is strongly $\mathscr{A}$-measurable.

It is easy to see that the linear mapping $S : E^* \to L^p(A)$, $Sx^* := \langle f, x^*\rangle$ is closed. Hence $S$ is bounded by the closed graph theorem.

Put $A_n := \{\|f\| \leqslant n\}$. Then $A_n \in \mathscr{A}$ and by Proposition 1.16 the integral $\int_{A_n} f\,d\mu$ exists as a Bochner integral in $E$. For all $x^* \in E^*$ and $n \geqslant m$, by Hölder's inequality we have

$$\Big|\Big\langle \int_{A_n\setminus A_m} f\,d\mu(x), x^*\Big\rangle\Big| \leqslant \big(\mu(A_n\setminus A_m)\big)^{\frac{1}{q}}\Big(\int_A |\langle f, x^*\rangle|^p\,d\mu(x)\Big)^{\frac{1}{p}}$$

$$\leqslant \big(\mu(A_n\setminus A_m)\big)^{\frac{1}{q}}\|S\|\,\|x^*\|.$$

Taking the supremum over all $x^* \in E^*$ with $\|x^*\| \leqslant 1$ we see that

$$\limsup_{m,n\to\infty}\Big\|\int_{A_n\setminus A_m} f\,d\mu\Big\| \leqslant \lim_{m,n\to\infty}\big(\mu(A_n\setminus A_m)\big)^{\frac{1}{q}}\|S\| = 0.$$

Hence the limit $x_f := \lim_{n\to\infty}\int_{A_n} f\,d\mu$ exists in $E$. Clearly,

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

for all $x^* \in E^*$. Uniqueness is obvious by the Hahn-Banach theorem. □

The element $x_f$ is called the *Pettis integral* of $f$ with respect to $\mu$.

### 1.3.2 The Lebesgue-Bochner spaces $L^p(A;E)$

Let $(A,\mathscr{A},\mu)$ be a $\sigma$-finite measure space. For $1 \leqslant p < \infty$ we define $L^p(A;E)$ as the linear space of all (equivalence classes of) strongly $\mu$-measurable functions $f : A \to E$ for which

$$\int_A \|f\|^p\,d\mu < \infty,$$

identifying functions which are equal $\mu$-almost everywhere. Endowed with the norm

$$\|f\|_{L^p(A;E)} := \Big(\int_A \|f\|^p\,d\mu\Big)^{\frac{1}{p}},$$

the space $L^p(A;E)$ is a Banach space; the proof for the scalar case carries over verbatim. Repeating the second part of the proof of Proposition 1.16 we see that the $\mu$-simple functions are dense in $L^p(A;E)$.

Note that the elements of $L^1(A;E)$ are precisely the equivalence classes of $\mu$-Bochner integrable functions.

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

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

the space $L^\infty(A;E)$ is a Banach space.

*Example 1.21.* For each $1 \leqslant p \leqslant \infty$, the Fubini theorem establishes a canonical isometric isomorphism

$$L^p(A_1; L^p(A_2;E)) \simeq L^p(A_1\times A_2;E),$$

which is uniquely defined by the mapping $1_{A_1}\otimes(1_{A_2}\otimes x) \mapsto 1_{A_1\times A_2}\otimes x$ and linearity. Here $1_A\otimes y \in L^p(A;F)$ is defined by $(1_A\otimes y)(\xi) := 1_A(\xi)y$.

## 1.4 Exercises

1. (!)[1] Let $E$ be a separable Banach space and let $C$ be a closed convex subset of $E$. Prove that there exists a sequence $(x_n^*)_{n=1}^\infty$ of norm one elements in $E^*$ and a sequence $(F_n)_{n=1}^\infty$ of closed sets in $\mathbb{K}$ such that
$$C = \bigcap_{n=1}^{\infty} \big\{x \in E : \langle x, x_n^* \rangle \in F_n\big\}.$$
*Hint:* Separate $C$ from the elements of a dense sequence in its complement $\complement C$ using the Hahn-Banach separation theorem.

2. Prove that the function $f : (0,1) \to L^\infty(0,1)$ defined by $f(t) = 1_{(0,t)}$ is weakly measurable, but not strongly measurable.
*Hint:* In the real case, elements in the dual of $L^\infty(0,1)$ can be decomposed into a positive and negative part. The complex case, consider real and imaginary parts separately.

3. Let $E$ be a Banach space and $f : [0,1] \to E$ a continuous function. Show that $f$ is Bochner integrable, and that its Bochner integral coincides with its Riemann integral.

4. A familiar theorem of calculus asserts that
$$\frac{d}{dx}\int_0^1 f(x,y)\,dy = \int_0^1 \frac{\partial f}{\partial x}(x,y)\,dy$$
for suitable functions $f : [0,1]\times[0,1] \to \mathbb{K}$. Show that this is a special case of Hille's theorem and deduce a set of rigorous conditions for this result.

5. Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space and let $1 \leqslant p, q \leqslant \infty$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Let $E$ be a Banach space and let $F$ be a norming subspace of $E^*$. Prove that $L^q(A;F)$ is a norming subspace of $(L^p(A;E))^*$ with respect to the duality pairing
$$\langle f, g\rangle = \int_A \langle f(\xi), g(\xi)\rangle\, d\mu(\xi), \quad f \in L^p(A;E),\ g \in L^q(A;E^*).$$
*Hint:* First find simple functions in $L^q(A;F)$ which norm simple functions in $L^p(A;E)$.

**Notes.** The material in this lecture is standard and can be found in many textbooks. More complete discussions of measurability in Banach spaces can be found in the monographs by Bogachev [8] and Vakhania, Tarieladze, Chobanyan [105]. Systematic expositions of the Bochner integral are presented in Arendt, Batty, Hieber, Neubrander [3], Diestel and Uhl [36], Dunford and Schwartz [37] and Lang [66].

[1] Results proved in the exercises marked with (!) are needed in the main text.

The Pettis measurability theorems 1.5 and 1.11 as well as Theorem 1.20 are due to Pettis [90]. Both versions of the Pettis measurability theorem remain correct if we only assume $f$ to be weakly measurable with respect to the functionals from a subspace $F$ of $E^*$ which separates the points of $E$, but the proof is more involved. For more details we refer to [36] and [105].

# 2

# Random variables in Banach spaces

In this lecture we take up the study of random variables with values in a Banach space $E$. The main result is the Itô-Nisio theorem (Theorem 2.17), which asserts that various modes of convergence of sums of independent symmetric $E$-valued random variables are equivalent. This result gives us a powerful tool to check the almost sure convergence of sums of independent symmetric random variables and will play an important role in the forthcoming lectures. The proof of the Itô-Nisio theorem is based on a uniqueness property of Fourier transforms (Theorem 2.8).

From this lecture onwards, we shall always assume that all spaces are real. This assumption is convenient when dealing with Fourier transforms and, in later lectures, when using the Riesz representation theorem to identify Hilbert spaces and their duals. However, much of the theory also works for complex scalars and can in fact be deduced from the real case. For some results it suffices to note that every complex vector space is a real space (by restricting the scalar multiplication to the reals); in others one proceeds by considering real and imaginary parts separately. We leave it to the interested reader to verify this in particular instances.

## 2.1 Random variables

A *probability space* is a triple $(\Omega, \mathscr{F}, \mathbb{P})$, where $\mathbb{P}$ is a *probability measure* on a measurable space $(\Omega, \mathscr{F})$, that is, $\mathbb{P}$ is a non-negative measure on $(\Omega, \mathscr{F})$ satisfying $\mathbb{P}(\Omega) = 1$.

**Definition 2.1.** *An $E$-valued* random variable *is an $E$-valued strongly $\mathbb{P}$-measurable function $X$ defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$.*

We think of $X$ as a 'random' element $x$ of $E$, which explains the choice of the letter '$X$'.

The underlying probability space $(\Omega, \mathscr{F}, \mathbb{P})$ will always be considered as fixed, and the prefix '$\mathbb{P}$-' will be omitted from our terminology unless confusion may arise. For instance, 'strongly measurable' means 'strongly $\mathbb{P}$-measurable' and 'almost surely' means '$\mathbb{P}$-almost surely', which is used synonymously with '$\mathbb{P}$-almost everywhere'. All integrals of $E$-valued random variables will be Bochner integrals unless stated otherwise, and the prefix 'Bochner' will usually be omitted.

The integral of an integrable random variable $X$ is called its *mean value* or *expectation* and is denoted by

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

If $X$ is an $E$-valued random variable, then by Proposition 1.10 $X$ has a strongly $\mathscr{F}$-measurable version $\widetilde{X}$ and by Proposition 1.8 the event

$$\{\widetilde{X} \in B\} := \{\omega \in \Omega : \ \widetilde{X}(\omega) \in B\}$$

belongs to $\mathscr{F}$ for all $B \in \mathscr{B}(E)$. The probability $\mathbb{P}\{\widetilde{X} \in B\}$ does not depend on the particular choice of the $\mathscr{F}$-measurable version $\widetilde{X}$, a fact which justifies the notation

$$\mathbb{P}\{X \in B\} := \mathbb{P}\{\widetilde{X} \in B\}$$

which will be used in the sequel without further notice.

**Definition 2.2.** *The* distribution *of an $E$-valued random variable $X$ is the Borel probability measure $\mu_X$ on $E$ defined by*

$$\mu_X(B) := \mathbb{P}\{X \in B\}, \qquad B \in \mathscr{B}(E).$$

*Random variables having the same distribution are said to be* identically distributed.

In the second part of this definition we allow the random variables to be defined on different probability spaces. If $X$ and $Y$ are identically distributed $E$-valued random variables and $f : E \to F$ is a Borel function, then $f(X)$ and $f(Y)$ are identically distributed. For example, for $1 \leqslant p < \infty$ it follows that

$$\mathbb{E}\|X\|^p = \mathbb{E}\|Y\|^p$$

if at least one (and then both) of these expectations are finite.

The next proposition shows that every $E$-valued random variable is *tight*:

**Proposition 2.3.** *If $X$ is a random variable in $E$, then for every $\varepsilon > 0$ there exists a compact set $K$ in $E$ such that $\mathbb{P}\{X \notin K\} < \varepsilon$.*

*Proof.* Since $X$ is separably valued outside some null set, we may assume that $E$ is separable. Let $(x_n)_{n=1}^{\infty}$ be a dense sequence in $E$ and fix $\varepsilon > 0$. For each integer $k \geqslant 1$ the closed balls $B(x_n, \frac{1}{k})$ cover $E$, and therefore there exists an index $N_k \geqslant 1$ such that

$$\mathbb{P}\Big\{X \in \bigcup_{n=1}^{N_k} B\big(x_n, \frac{1}{k}\big)\Big\} > 1 - \frac{\varepsilon}{2^k}.$$

The set $K := \bigcap_{k\geqslant 1} \bigcup_{n=1}^{N_k} B\big(x_n, \frac{1}{k}\big)$ is closed and totally bounded. Since $E$ is complete, $K$ is compact. Moreover,

$$\mathbb{P}\{X \notin K\} < \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon. \qquad \square$$

This result motivates the following definition.

**Definition 2.4.** *A family $\mathscr{X}$ of random variables in $E$ is* uniformly tight *if for every $\varepsilon > 0$ there exists a compact set $K$ in $E$ such that*

$$\mathbb{P}\{X \notin K\} < \varepsilon \qquad \forall X \in \mathscr{X}.$$

The following lemma will be useful in the proof of the Itô-Nisio theorem.

**Lemma 2.5.** *If $\mathscr{X}$ is uniformly tight, then $\mathscr{X} - \mathscr{X} = \{X_1 - X_2 : \ X_1, X_2 \in \mathscr{X}\}$ is uniformly tight.*

*Proof.* Let $\varepsilon > 0$ be arbitrary and fixed. Choose a compact set $K$ in $E$ such that $\mathbb{P}\{X \in K\} \geqslant 1 - \varepsilon$ for all $X \in \mathscr{X}$. The set $L = \{x - y : \ x, y \in K\}$ is compact, being the image of the compact set $K \times K$ under the continuous map $(x, y) \mapsto x - y$. Since $X_1(\omega), X_2(\omega) \in K$ implies $X_1(\omega) - X_2(\omega) \in L$,

$$\mathbb{P}\{X_1 - X_2 \notin L\} \leqslant \mathbb{P}\{X_1 \notin K\} + \mathbb{P}\{X_2 \notin K\} < 2\varepsilon. \qquad \square$$

## 2.2 Fourier transforms

We begin with a definition.

**Definition 2.6.** *The* Fourier transform *of a Borel probability measure $\mu$ on $E$ is the function $\widehat{\mu} : E^* \to \mathbb{C}$ defined by*

$$\widehat{\mu}(x^*) := \int_E \exp(-i\langle x, x^*\rangle)\, d\mu(x).$$

*The* Fourier transform *of a random variable $X : \Omega \to E$ is the Fourier transform of its distribution $\mu_X$.*

Note that the above integral converges absolutely, as $|\exp(-i\langle x, x^*\rangle)| = 1$ for all $x \in E$ since we are assuming that $E$ is a real Banach space. By a change of variable, the Fourier transform of a random variable $X$ on $E$ is given by

$$\widehat{X}(x^*) := \mathbb{E}\exp(-i\langle X, x^*\rangle) = \int_E \exp(-i\langle x, x^*\rangle)\, d\mu_X(x).$$

The proof of the next theorem is based upon a uniqueness result known as *Dynkin's lemma*. It states that two probability measures agree if they agree on a sufficiently rich family of sets.

**Lemma 2.7 (Dynkin).** *Let $\mu_1$ and $\mu_2$ be two probability measures defined on a measurable space $(\Omega, \mathscr{F})$. Let $\mathscr{A} \subseteq \mathscr{F}$ be a collection of sets with the following properties:*

(1) *$\mathscr{A}$ is closed under finite intersections;*
(2) *$\sigma(\mathscr{A})$, the $\sigma$-algebra generated by $\mathscr{A}$, equals $\mathscr{F}$.*

*If $\mu_1(A) = \mu_2(A)$ for all $A \in \mathscr{A}$, then $\mu_1 = \mu_2$.*

*Proof.* Let $\mathscr{D}$ denote the collection of all sets $D \in \mathscr{F}$ with $\mu_1(D) = \mu_2(D)$. Then $\mathscr{A} \subseteq \mathscr{D}$ and $\mathscr{D}$ is a *Dynkin system*, that is,

- $\Omega \in \mathscr{D}$;
- if $D_1 \subseteq D_2$ with $D_1, D_2 \in \mathscr{D}$, then also $D_2 \setminus D_1 \in \mathscr{D}$;
- if $D_1 \subseteq D_2 \subseteq \ldots$ with all $D_n \in \mathscr{D}$, then also $\bigcup_{n\geqslant 1} D_n \in \mathscr{D}$.

By assumption we have $\mathscr{D} \subseteq \mathscr{F} = \sigma(\mathscr{A})$; we will show that $\sigma(\mathscr{A}) \subseteq \mathscr{D}$. To this end let $\mathscr{D}_0$ denote the smallest Dynkin system in $\mathscr{F}$ containing $\mathscr{A}$. We will show that $\sigma(\mathscr{A}) \subseteq \mathscr{D}_0$. In view of $\mathscr{D}_0 \subseteq \mathscr{D}$, this will prove the lemma.

Let $\mathscr{C} = \{D_0 \in \mathscr{D}_0 :\ D_0 \cap A \in \mathscr{D}_0 \text{ for all } A \in \mathscr{A}\}$. Then $\mathscr{C}$ is a Dynkin system and $\mathscr{A} \subseteq \mathscr{C}$ since $\mathscr{A}$ is closed under taking finite intersections. It follows that $\mathscr{D}_0 \subseteq \mathscr{C}$, since $\mathscr{D}_0$ is the smallest Dynkin system containing $\mathscr{A}$. But obviously, $\mathscr{C} \subseteq \mathscr{D}_0$, and therefore $\mathscr{C} = \mathscr{D}_0$.

Now let $\mathscr{C}' = \{D_0 \in \mathscr{D}_0 :\ D_0 \cap D \in \mathscr{D}_0 \text{ for all } D \in \mathscr{D}_0\}$. Then $\mathscr{C}'$ is a Dynkin system and the fact that $\mathscr{C} = \mathscr{D}_0$ implies that $\mathscr{A} \subseteq \mathscr{C}'$. Hence $\mathscr{D}_0 \subseteq \mathscr{C}'$, since $\mathscr{D}_0$ is the smallest Dynkin system containing $\mathscr{A}$. But obviously, $\mathscr{C}' \subseteq \mathscr{D}_0$, and therefore $\mathscr{C}' = \mathscr{D}_0$.

It follows that $\mathscr{D}_0$ is closed under taking finite intersections. But a Dynkin system with this property is a $\sigma$-algebra. Thus, $\mathscr{D}_0$ is a $\sigma$-algebra, and now $\mathscr{A} \subseteq \mathscr{D}_0$ implies that also $\sigma(\mathscr{A}) \subseteq \mathscr{D}_0$. $\square$

**Theorem 2.8 (Uniqueness of the Fourier transform).** *Let $X_1$ and $X_2$ be $E$-valued random variables whose Fourier transforms are equal:*

$$\widehat{X_1}(x^*) = \widehat{X_2}(x^*) \qquad \forall x^* \in E^*.$$

*Then $X_1$ and $X_2$ are identically distributed.*

*Proof.* Since $X_1$ and $X_2$ are $\mu$-separably valued there is no loss of generality in assuming that $E$ is separable.

*Step 1* - First we prove: if $\lambda_1$ and $\lambda_2$ are Borel probability measures on $\mathbb{R}^d$ with the property that $\widehat{\lambda_1}(t) = \widehat{\lambda_2}(t)$ for all $t \in \mathbb{R}^d$, then $\lambda_1 = \lambda_2$. By Dynkin's lemma, for the latter it suffices to prove that $\lambda_1(K) = \lambda_2(K)$ for all compact subsets $K$ of $\mathbb{R}^d$. By the dominated convergence theorem, for the latter suffices to prove that

$$\int_{-\infty}^{\infty} f(\xi)\,d\lambda_1(\xi) = \int_{-\infty}^{\infty} f(\xi)\,d\lambda_2(\xi) \qquad \forall f \in C_c(\mathbb{R}^d), \tag{2.1}$$

where $C_c(\mathbb{R}^d)$ denote the space of all compactly supported continuous functions on $\mathbb{R}^d$.

Let $\varepsilon > 0$ be arbitrary and fix an $f \in C_c(\mathbb{R}^d)$. We may assume that $\|f\|_\infty \leqslant 1$. Let $r > 0$ be so large that the support of $f$ is contained in $[-r,r]^d$ and such that $\lambda_j\big(\complement[-r,r]^d\big) \leqslant \varepsilon$ for $j = 1,2$. By the Stone-Weierstrass theorem there exists a trigonometric polynomial $p : \mathbb{R}^d \to \mathbb{C}$ of period $2r$ such that $\sup_{t\in[-r,r]^d} |f(t) - p(t)| \leqslant \varepsilon$. Then,

$$\begin{aligned}
&\Big| \int_{\mathbb{R}^d} f(\xi)\,d\lambda_1(\xi) - \int_{\mathbb{R}^d} f(\xi)\,d\lambda_2(\xi) \Big| \\
&\quad \leqslant 4\varepsilon + 2(1+\varepsilon)\varepsilon + \Big| \int_{\mathbb{R}^d} p(\xi)\,d\lambda_1(\xi) - \int_{\mathbb{R}^d} p(\xi)\,d\lambda_2(\xi) \Big| \\
&\quad = 4\varepsilon + 2(1+\varepsilon)\varepsilon,
\end{aligned}$$

where the terms $2(1+\varepsilon)\varepsilon$ come from the estimate $\|p\|_\infty \leqslant 1+\varepsilon$ and the last equality follows from the equality of the Fourier transforms of $\lambda_1$ and $\lambda_2$. Since $\varepsilon > 0$ was arbitrary, this proves (2.1).

*Step 2* - If $\mu$ is any Borel probability measure on $E$, then for all $d \geqslant 1$ and all $t = (t_1, \dots, t_d) \in \mathbb{R}^d$ and $x_1^*, \dots, x_d^* \in E^*$ we have

$$\widehat{\mu}\Big(\sum_{j=1}^d t_j x_j^*\Big) = \int_E e^{-i\sum_{j=1}^d \langle x, t_j x_j^*\rangle}\,d\mu(x) = \int_{\mathbb{R}^d} e^{-i\langle t,\xi\rangle}\,d(T\mu)(\xi) = \widehat{T\mu}(t),$$

where $T\mu$ denotes Borel probability measure on $\mathbb{R}^d$ obtained as the image measure of $\mu$ under the map $T : E \to \mathbb{R}^d$, $x \mapsto (\langle x, x_1^*\rangle, \dots, \langle x, x_d^*\rangle)$, that is,

$$T\mu(B) := \mu\big\{x \in E :\ (\langle x, x_1^*\rangle,\ \dots,\ \langle x, x_d^*\rangle) \in B\big\}.$$

*Step 3* - Applying Step 2 to the measures $\mu_{X_1}$ and $\mu_{X_2}$ it follows that $\widehat{T\mu_{X_1}}(t) = \widehat{T\mu_{X_2}}(t)$ for all $t \in \mathbb{R}^d$. By Step 1, $T\mu_{X_1} = T\mu_{X_2}$. Hence $\mu_{X_1}$ and $\mu_{X_2}$ agree on the collection $\mathscr{C}(E)$ consisting of all Borel sets in $E$ of the form

$$\big\{x \in E :\ (\langle x, x_1^*\rangle,\ \dots,\ \langle x, x_d^*\rangle) \in B\big\}$$

with $d \geqslant 1$, $x_1^*, \dots, x_d^* \in E^*$ and $B \in \mathscr{B}(\mathbb{R}^d)$. Since $E$ is separable, every closed ball $\{x \in E :\ \|x - x_0\| \leqslant r\}$ can be written as a countable intersection of sets in $\mathscr{C}(E)$ (see Exercise 1.1). Thus the family $\mathscr{C}(E)$ generates the Borel $\sigma$-algebra $\mathscr{B}(E)$ and $\mu_{X_1} = \mu_{X_2}$ by Dynkin's Lemma. $\square$

## 2.3 Convergence in probability

In the absence of integrability conditions the following definition for convergence of random variables is often very useful.

**Definition 2.9.** *A sequence $(X_n)_{n=1}^\infty$ of $E$-valued random variables* converges in probability *to an $E$-valued random variable $X$ if for all $r > 0$ we have*

$$\lim_{n\to\infty} \mathbb{P}\{\|X_n - X\| > r\} = 0.$$

If $\lim_{n\to\infty} X_n = X$ in $L^p(\Omega; E)$ for some $1 \leqslant p < \infty$, then $\lim_{n\to\infty} X_n = X$ in probability. This follows from *Chebyshev's inequality*, which states that if $\xi \in L^p(\Omega)$, then for all $r > 0$ we have

$$\mathbb{P}\{|\xi| \geqslant r\} \leqslant \frac{1}{r^p}\mathbb{E}|\xi|^p.$$

The proof is simple:

$$\mathbb{P}\{|\xi| \geqslant r\} = \frac{1}{r^p}\int_{\{|\xi|^p \geqslant r^p\}} r^p \, d\mathbb{P} \leqslant \frac{1}{r^p}\int_{\{|\xi|^p \geqslant r^p\}} |\xi|^p \, d\mathbb{P} \leqslant \frac{1}{r^p}\mathbb{E}|\xi|^p.$$

Our first aim is to show that if $(X_n)_{n=1}^\infty$ converges in probability, then some subsequence converges almost surely. For this we need a lemma which is known as the *Borel-Cantelli lemma*.

**Lemma 2.10 (Borel-Cantelli).** *If $(A, \mathscr{A}, \mu)$ is a measure space and $(A_n)_{n=1}^\infty$ is a sequence in $\mathscr{A}$ satisfying $\sum_{n=1}^\infty \mu(A_n) < \infty$, then*

$$\mu\Big(\bigcap_{k\geqslant 1}\bigcup_{n\geqslant k} A_n\Big) = 0.$$

*Proof.* Let $k_0 \geqslant 1$. Then,

$$\mu\Big(\bigcap_{k\geqslant 1}\bigcup_{n\geqslant k} A_n\Big) \leqslant \mu\Big(\bigcup_{n\geqslant k_0} A_n\Big) \leqslant \sum_{n=k_0}^\infty \mu(A_n),$$

and the right hand side tends to 0 as $k_0 \to \infty$. □

Note that $\omega \in \bigcap_{k\geqslant 1}\bigcup_{n\geqslant k} A_n$ if and only if $\omega \in A_n$ for infinitely many indices $n$.

**Proposition 2.11.** *If a sequence $(X_n)_{n=1}^\infty$ of $E$-valued random variables converges in probability, then it has an almost surely convergent subsequence $(X_{n_k})_{k=1}^\infty$.*

*Proof.* Let $\lim_{n\to\infty} X_n = X$ in probability. Choose an increasing sequence of indices $n_1 < n_2 < \dots$ satisfying

$$\mathbb{P}\Big\{\|X_{n_k} - X\| > \frac{1}{k}\Big\} < \frac{1}{2^k} \qquad \forall k \geqslant 1.$$

By the Borel-Cantelli lemma,

$$\mathbb{P}\Big\{\|X_{n_k} - X\| > \frac{1}{k} \text{ for infinitely many } k \geqslant 1\Big\} = 0.$$

Outside this null set we have $\lim_{k\to\infty} X_{n_k} = X$ pointwise. □

## 2.4 Independence

Next we recall the notion of independence. The reader who is already familiar with it may safely skip this section.

**Definition 2.12.** *A family of random variables $(X_i)_{i\in I}$, where $I$ is some index set and each $X_i$ takes values in a Banach space $E_i$, is* independent *if for all choices of distinct indices $i_1, \dots, i_N \in I$ and all Borel sets $B_1, \dots, B_N$ in $E_{i_1}, \dots, E_{i_N}$ we have*

$$\mathbb{P}\{X_{i_1} \in B_1, \ \dots, \ X_{i_N} \in B_N\} = \prod_{n=1}^{N} \mathbb{P}\{X_{i_n} \in B_n\}.$$

Note that $(X_i)_{i\in I}$ is independent if and only if every finite subfamily of $(X_i)_{i\in I}$ is independent. Thus, in order to check independence of a given family of random variables it suffices to consider its finite subfamilies.

We assume that the reader is familiar with the elementary properties of independent real-valued random variables such as covered in a standard course on probability. Here we content ourselves recalling that if $\eta$ and $\xi$ are real-valued random variables which are integrable and independent, then their product $\eta\xi$ is integrable and $\mathbb{E}(\eta\xi) = \mathbb{E}\eta\,\mathbb{E}\xi$.

In the next two propositions, $X_1, \dots, X_N$ are random variables with values in the Banach spaces $E_1, \dots, E_N$, respectively. If $\nu_1, \dots, \nu_n$ are probability measures, we denote by $\nu_1 \times \dots \times \nu_n$ their product measure. The distribution of the $E^N$-valued random variable $(X_1, \dots, X_N)$ is denoted by $\mu_{(X_1,\dots,X_N)}$.

**Proposition 2.13.** *The random variables $X_1, \dots, X_N$ are independent if and only if*

$$\mu_{(X_1,\dots,X_N)} = \mu_{X_1} \times \dots \times \mu_{X_N}.$$

*Proof.* By definition, the random variables $X_1, \dots, X_N$ are independent if and only if $\mu_{(X_1,\dots,X_N)}$ and $\mu_{X_1} \times \dots \times \mu_{X_N}$ agree on all Borel rectangles $B_1 \times \dots \times B_N$ in $E_1 \times \dots \times E_N$. By Dynkin's lemma this happens if and only if $\mu_{(X_1,\dots,X_N)} = \mu_{X_1} \times \dots \times \mu_{X_N}$. □

We record two corollaries.

**Proposition 2.14.** *If* $\lim_{n\to\infty} X_n = X$ *and* $\lim_{n\to\infty} Y_n = Y$ *in probability and each* $X_n$ *is independent of* $Y_n$*, then* $X$ *and* $Y$ *are independent.*

*Proof.* By passing to a subsequence we may assume that $\lim_{n\to\infty} X_n = X$ and $\lim_{n\to\infty} Y_n = Y$ almost surely. We consider the $E \times E$-valued random variables $Z_n = (X_n, Y_n)$ and $Z = (X, Y)$. Identifying the dual of $E \times E$ with $E^* \times E^*$, by dominated convergence we obtain

$$\begin{aligned}
\widehat{\mu}_Z(x^*, y^*) &= \mathbb{E}\exp(-i(\langle X, x^*\rangle + \langle Y, y^*\rangle)) \\
&= \lim_{n\to\infty} \mathbb{E}\exp(-i(\langle X_n, x^*\rangle + \langle Y_n, y^*\rangle)) \\
&= \lim_{n\to\infty} \mathbb{E}\exp(-i\langle X_n, x^*\rangle)\mathbb{E}\exp(-i\langle Y_n, y^*\rangle) \\
&= \mathbb{E}\exp(-i\langle X, x^*\rangle)\mathbb{E}\exp(-i\langle Y, y^*\rangle) \\
&= \widehat{\mu_X}(x^*)\widehat{\mu_Y}(y^*) = \widehat{\mu_X \times \mu_Y}(x^*, y^*).
\end{aligned}$$

From Theorem 2.8 we conclude that $\mu_Z = \mu_X \times \mu_Y$. Now the result follows from Proposition 2.13. $\square$

**Definition 2.15.** *An* $E$*-valued random variable* $X$ *is called* symmetric *if* $X$ *and* $-X$ *are identically distributed.*

**Proposition 2.16.** *If* $X$ *is symmetric and independent of* $Y$*, then for all* $1 \leqslant p < \infty$ *we have*

$$\mathbb{E}\|X\|^p \leqslant \mathbb{E}\|X + Y\|^p.$$

*Proof.* The symmetry of $X$ and the independence of $X$ and $Y$ imply that $X + Y$ and $-X + Y$ are identically distributed, and therefore

$$\begin{aligned}
\big(\mathbb{E}\|X\|^p\big)^{\frac{1}{p}} &= \tfrac{1}{2}\big(\mathbb{E}\|(X+Y) + (X-Y)\|^p\big)^{\frac{1}{p}} \\
&\leqslant \tfrac{1}{2}\big(\mathbb{E}\|X+Y\|^p\big)^{\frac{1}{p}} + \tfrac{1}{2}\big(\mathbb{E}\|X-Y\|^p\big)^{\frac{1}{p}} = \big(\mathbb{E}\|X+Y\|^p\big)^{\frac{1}{p}}. \quad \square
\end{aligned}$$

## 2.5 The Itô-Nisio theorem

In this section we prove a celebrated result, due to Itô and Nisio, which states that a sum of symmetric and independent $E$-valued random variables converges (weakly) almost surely if and only if it converges in probability.

Here is the precise statement of the theorem:

**Theorem 2.17 (Itô-Nisio).** *Let* $X_n : \Omega \to E$*,* $n \geqslant 1$*, be independent symmetric random variables, put* $S_n := \sum_{j=1}^n X_j$*, and let* $S : \Omega \to E$ *be a random variable. The following assertions are equivalent:*

(1) *for all* $x^* \in E^*$ *we have* $\lim_{n\to\infty}\langle S_n, x^*\rangle = \langle S, x^*\rangle$ *almost surely;*
(2) *for all* $x^* \in E^*$ *we have* $\lim_{n\to\infty}\langle S_n, x^*\rangle = \langle S, x^*\rangle$ *in probability;*
(3) *we have* $\lim_{n\to\infty} S_n = S$ *almost surely;*
(4) *we have* $\lim_{n\to\infty} S_n = S$ *in probability.*

*If these equivalent conditions hold and* $\mathbb{E}\|S\|^p < \infty$ *for some* $1 \leqslant p < \infty$, *then*

$$\lim_{n\to\infty} \mathbb{E}\|S_n - S\|^p = 0.$$

We begin with a tail estimate known as *Lévy's inequality.*

**Lemma 2.18.** *Let* $X_1, \dots, X_n$ *be independent symmetric* $E$*-valued random variables, and put* $S_k := \sum_{j=1}^k X_j$ *for* $k = 1, \dots, n$. *Then for all* $r \geqslant 0$ *we have*

$$\mathbb{P}\Big\{\max_{1\leqslant k\leqslant n} \|S_k\| > r\Big\} \leqslant 2\mathbb{P}\{\|S_n\| > r\}.$$

*Proof.* Put

$$\begin{aligned} A &:= \Big\{\max_{1\leqslant k\leqslant n} \|S_k\| > r\Big\}, \\ A_k &:= \{\|S_1\| \leqslant r,\, \dots,\, \|S_{k-1}\| \leqslant r,\, \|S_k\| > r\}; \quad k = 1, \dots, n. \end{aligned}$$

The sets $A_1, \dots, A_n$ are disjoint and $\bigcup_{k=1}^n A_k = \Big\{\max_{1\leqslant k\leqslant n} \|S_k\| > r\Big\}$.

The identity $S_k = \frac{1}{2}(S_n + (2S_k - S_n))$ implies that

$$\{\|S_k\| > r\} \subseteq \{\|S_n\| > r\} \cup \{\|2S_k - S_n\| > r\}.$$

We also note $(X_1, \dots, X_n)$ and $(X_1, \dots, X_k, -X_{k+1}, \dots, -X_n)$ are identically distributed (see Exercise 2), which, in view of the identities

$$S_n = S_k + X_{k+1} + \dots + X_n, \qquad 2S_k - S_n = S_k - X_{k+1} - \dots - X_n,$$

implies that $(X_1, \dots, X_k, S_n)$ and $(X_1, \dots, X_k, 2S_k - S_n)$ are identically distributed. Hence,

$$\begin{aligned} \mathbb{P}(A_k) &\leqslant \mathbb{P}(A_k \cap \{\|S_n\| > r\}) + \mathbb{P}(A_k \cap \{\|2S_k - S_n\| > r\}) \\ &= 2\mathbb{P}(A_k \cap \{\|S_n\| > r\}). \end{aligned}$$

Summing over $k$ we obtain

$$\mathbb{P}(A) = \sum_{k=1}^n \mathbb{P}(A_k) \leqslant 2\sum_{k=1}^n \mathbb{P}(A_k \cap \{\|S_n\| > r\}) = 2\mathbb{P}\{\|S_n\| > r\}. \qquad \square$$

*Proof (Proof of Theorem 2.17).* We prove the implications (2)⇒(4)⇒(3), the implications (3)⇒(1)⇒(2) being clear.

(2)⇒(4): We split this proof into two steps.

*Step 1* – In this step we prove that the sequence $(S_n)_{n\geqslant 1}$ is uniformly tight.

For all $m \geqslant n$ and $x^* \in E^*$ the random variables $\langle S_m - S_n, x^*\rangle$ and $\pm\langle S_n, x^*\rangle$ are independent. Hence by Proposition 2.14, $\langle S - S_n, x^*\rangle$ and $\pm\langle S_n, x^*\rangle$ are independent. Next we claim that $S$ and $S - 2S_n$ are identically distributed. Indeed, denote their distributions by $\mu$ and $\lambda_n$, respectively. By the independence of $\langle S - S_n, x^*\rangle$ and $\pm\langle S_n, x^*\rangle$ and the symmetry of $S_n$, for all $x^* \in E^*$ we have

$$\begin{aligned}\widehat{\mu}(x^*) = \mathbb{E}\left(e^{-i\langle S,x^*\rangle}\right) &= \mathbb{E}\left(e^{-i\langle S-S_n,x^*\rangle}\right)\cdot \mathbb{E}\left(e^{-i\langle S_n,x^*\rangle}\right)\\ &= \mathbb{E}\left(e^{-i\langle S-S_n,x^*\rangle}\right)\cdot \mathbb{E}\left(e^{-i\langle -S_n,x^*\rangle}\right)\\ &= \mathbb{E}\left(e^{-i\langle S-2S_n,x^*\rangle}\right) = \widehat{\lambda_n}(x^*).\end{aligned}$$

By Theorem 2.8, this shows that $\mu = \lambda_n$ and the claim is proved.

Given $\varepsilon > 0$ we can find a compact set $K \subseteq E$ with $\mu(K) = \mathbb{P}\{S \in K\} > 1 - \varepsilon$. The set $L := \frac{1}{2}(K - K)$ is compact as well, and arguing as in the proof of Lemma 2.5 we have

$$\mathbb{P}\{S_n \notin L\} \leqslant \mathbb{P}\{S \notin K\} + \mathbb{P}\{S - 2S_n \notin K\} = 2\mathbb{P}\{S \notin K\} < 2\varepsilon.$$

It follows that $\mathbb{P}\{S_n \in L\} > 1 - 2\varepsilon$ for all $n \geqslant 1$, and therefore the sequence $(S_n)_{n=1}^{\infty}$ is uniformly tight.

*Step 2* – By Lemma 2.5, the sequence $(S_n - S)_{n\geqslant 1}$ is uniformly tight. Let $\nu_n$ denote the distribution of $S_n - S$. We need to prove that for all $\varepsilon > 0$ and $r > 0$ there exists an index $N \geqslant 1$ such that

$$\mathbb{P}\{\|S_n - S\| \geqslant r\} = \nu_n(\complement B(0,r)) < \varepsilon \qquad \forall n \geqslant N.$$

Suppose, for a contradiction, that such an $N$ does not exist for some $\varepsilon > 0$ and $r > 0$. Then there exists a subsequence $(S_{n_k})_{k\geqslant 1}$ such that

$$\nu_{n_k}(\complement B(0,r)) \geqslant \varepsilon, \qquad k \geqslant 1.$$

On the other hand, by uniform tightness we find a compact set $K$ such that $\nu_{n_k}(K) \geqslant 1 - \frac{1}{2}\varepsilon$ for all $k \geqslant 1$. It follows that

$$\nu_{n_k}(K \cap \complement B(0,r)) \geqslant \tfrac{1}{2}\varepsilon, \qquad k \geqslant 1.$$

By covering the compact set $K \cap \complement B(0,r)$ with open balls of radius $\frac{1}{2}r$ and passing to a subsequence, we find a ball $B$ not containing 0 and a number $\delta > 0$ such that

$$\nu_{n_{k_j}}(K \cap B) = \mathbb{P}\{S_{n_{k_j}} - S \in K \cap B\} \geqslant \delta, \qquad j \geqslant 1.$$

By the Hahn-Banach separation theorem, there is a functional $x^* \in E^*$ such that $\langle x, x^* \rangle \geqslant 1$ for all $x \in B$. For all $\omega \in \{S_{n_k} - S \in K \cap B\}$ it follows that $\langle S_{n_{k_j}}(\omega) - S(\omega), x^* \rangle \geqslant 1$. Thus, $\langle S_{n_k}, x^* \rangle$ fails to converge to $\langle S, x^* \rangle$ in probability. This contradiction concludes the proof.

(4)⇒(3): Assume that $\lim_{n\to\infty} S_n = S$ in probability for some random variable $S$. By Proposition 2.11 there is a subsequence $(S_{n_k})_{k=1}^\infty$ converging almost surely to $S$. Fix $k$ and let $m > n_k$. Then by Lévy's inequality,

$$\begin{aligned}\mathbb{P}\Big\{ \sup_{n_k \leqslant j \leqslant m} \|S_j - S_{n_k}\| \geqslant r \Big\} &\leqslant 2\mathbb{P}(\|S_m - S_{n_k}\| \geqslant r) \\ &\leqslant 2\mathbb{P}\Big\{ \|S_m - S\| \geqslant \frac{r}{2} \Big\} + 2\mathbb{P}\Big\{ \|S - S_{n_k}\| \geqslant \frac{r}{2} \Big\}.\end{aligned}$$

Letting $m \to \infty$ we find

$$\mathbb{P}\Big\{ \sup_{j \geqslant n_k} \|S_j - S_{n_k}\| \geqslant r \Big\} \leqslant 2\mathbb{P}\Big\{ \|S - S_{n_k}\| \geqslant \frac{r}{2} \Big\},$$

and hence, upon letting $k \to \infty$,

$$\lim_{k\to\infty} \mathbb{P}\Big\{ \sup_{j \geqslant n_k} \|S_j - S_{n_k}\| \geqslant r \Big\} = 0.$$

Since $S_{n_k} \to S$ pointwise a.e., it follows that

$$\begin{aligned}\mathbb{P}\Big\{ \lim_{k\to\infty} \sup_{j \geqslant n_k} \|S_j - S\| \geqslant 2r \Big\} &\leqslant \lim_{k\to\infty} \mathbb{P}\Big\{ \sup_{j \geqslant n_k} \|S_j - S\| \geqslant 2r \Big\} \\ &\leqslant \lim_{k\to\infty} \mathbb{P}\Big\{ \sup_{j \geqslant n_k} \|S_j - S_{n_k}\| \geqslant r \Big\} + \lim_{k\to\infty} \mathbb{P}\Big\{ \sup_{j \geqslant n_k} \|S_{n_k} - S\| \geqslant r \Big\} = 0.\end{aligned}$$

It remains to prove the assertion about $L^p$-convergence. First we note that $S = S_n + (S - S_n)$ with $S_n$ and $S - S_n$ independent (by the independence of $S_n$ and $S_m - S_n$ for $m \geqslant n$ and Proposition 2.14), and therefore $\mathbb{E}\|S_n\|^p \leqslant \mathbb{E}\|S\|^p$ by Proposition 2.16. Hence by an integration by parts (see Exercise 1) and Lévy inequality,

$$\begin{aligned}\mathbb{E} \sup_{1 \leqslant k \leqslant n} \|S_k\|^p &= \int_0^\infty p r^{p-1} \mathbb{P}\Big\{ \sup_{1 \leqslant k \leqslant n} \|S_k\| > r \Big\}\, dr \\ &\leqslant 2 \int_0^\infty p r^{p-1} \mathbb{P}\{\|S_n\| > r\}\, dr = 2\mathbb{E}\|S_n\|^p \leqslant 2\mathbb{E}\|S\|^p.\end{aligned}$$

Hence $\mathbb{E} \sup_{k \geqslant 1} \|S_k\|^p \leqslant 2\mathbb{E}\|S\|^p$ by the monotone convergence theorem. Now $\lim_{n\to\infty} \|S_n - S\|^p = 0$ follows from the dominated convergence theorem. □

## 2.6 Exercises

1. (!) Let $\xi$ be a non-negative random variable and let $1 \leqslant p < \infty$. Prove the *integration by parts formula*

$$\mathbb{E}\xi^p = \int_0^\infty p\lambda^{p-1}\mathbb{P}\{\xi > \lambda\}\, d\lambda.$$

*Hint:* Write $\mathbb{P}\{\xi > \lambda\} = \mathbb{E}1_{\{\xi>\lambda\}}$ and apply Fubini's theorem.

2. (!) Let $X_1, \dots, X_N$ be independent symmetric $E$-valued random variables. Show that for all choices of $\varepsilon_1, \dots, \varepsilon_N \in \{-1, +1\}$ the $E^N$-valued random variables $(X_1, \dots, X_N)$ and $(\varepsilon_1 X_1, \dots, \varepsilon_N X_N)$ are identically distributed.
3. (!) Define the *convolution* of two Borel measures $\mu$ and $\nu$ on $E$ by

$$\mu * \nu(B) := \int_E \int_E 1_B(x+y)\, d\mu(x)\, d\nu(y), \quad B \in \mathscr{B}(E).$$

Prove that for all $x^* \in E^*$ we have $\widehat{\mu * \nu}(x^*) = \widehat{\mu}(x^*)\widehat{\nu}(x^*)$.

4. A sequence of $E$-valued random variables $(X_n)_{n=1}^\infty$ is *Cauchy in probability* if for all $\varepsilon > 0$ and $r > 0$ there exists an index $N \geqslant 1$ such that

$$\mathbb{P}\{\|X_n - X_m\| > r\} < \varepsilon \qquad \forall m, n \geqslant N.$$

Show that $(X_n)_{n=1}^\infty$ is Cauchy in probability if and only if $(X_n)_{n=1}^\infty$ converges in probability.

*Hint:* For the 'if' part, first show that some subsequence of $(X_n)_{n=1}^\infty$ converges almost surely.

5. Let $(X_n)_{n=1}^\infty$ be a sequence of $E$-valued random variables. Prove that if $\lim_{n\to\infty} X_n = X$ in probability, then $(X_n)_{n=1}^\infty$ is uniformly tight.

**Notes.** There are many excellent introductory texts on Probability Theory, among them the classic by Chung [21]. The more analytically inclined reader might consult Stromberg [101]. A comprehensive treatment of modern Probability Theory is offered by Kallenberg [55].

Thorough discussions of Banach space-valued random variables can be found in the monographs by Kwapień and Woyczyński [65], Ledoux and Talagrand [69], and Vakhania, Tarieladze, and Chobanyan [105].

The Itô-Nisio theorem was proved by Itô and Nisio in their beautiful paper [52] which we recommend for further reading. The usual proofs of this theorem are based upon the following celebrated and non-trivial compactness theorem due to Prokhorov:

**Theorem 2.19 (Prokhorov).** *For a family $\mathscr{M}$ of Borel probability measures on a separable complete metric space $M$ the following assertions are equivalent:*

(1) *$\mathscr{M}$ is uniformly tight;*
(2) *Every sequence $(\mu_n)_{n=1}^\infty$ in $\mathscr{M}$ has a weakly convergent subsequence.*

Here, (1) means that for all $\varepsilon > 0$ there exists a compact set $K$ in $M$ such that $\mu(\complement K) < \varepsilon$ for all $\mu \in \mathscr{M}$, and (2) means that there exist a subsequence $(\mu_{n_k})_{k\geqslant 1}$ and a Borel probability measure $\mu$ such that

$$\lim_{k\to\infty} \int_M f\, d\mu_{n_k} = \int_M f\, d\mu$$

for all bounded continuous functions $f : M \to \mathbb{R}$. This theorem is the starting point of measure theory on metric spaces. Expositions of this subject can be found in the monographs by Billingsley [7] and Parthasarathy [88], as well as in the recent two-volume treatise on measure theory by Bogachev [9]. Readers familiar with it will have noticed that some of the results which we have stated for $E$-valued random variables, such as Proposition 2.3 and Theorem 2.8, could just as well be stated for probability measures on $E$.

# 3

# Sums of independent random variables

This lecture collects a number of estimates for sums of independent random variables with values in a Banach space $E$. We concentrate on sums of the form $\sum_{n=1}^{N} \gamma_n x_n$, where the $\gamma_n$ are real-valued Gaussian variables and the $x_n$ are vectors in $E$. As we shall see later on such sums are the building blocks of general $E$-valued Gaussian random variables and, perhaps more importantly, stochastic integrals of $E$-valued step functions are of this form. Furthermore, they are used in the definition of various geometric properties of Banach spaces, such as type and cotype.

The highlights of this lecture are the Kahane contraction principle (Theorem 3.1), a covariance domination principle (Theorem 3.9) and the Kahane-Khintchine inequalities (Theorems 3.11 and 3.12).

## 3.1 Gaussian sums

We begin with an important inequality for sums of independent symmetric random variables, due to Kahane.

**Theorem 3.1 (Kahane contraction principle).** *Let $(X_n)_{n=1}^{\infty}$ be a sequence of independent symmetric $E$-valued random variables. Then for all $a_1, \dots, a_N \in \mathbb{R}$ and $1 \leqslant p < \infty$,*

$$\mathbb{E}\Big\| \sum_{n=1}^{N} a_n X_n \Big\|^p \leqslant \big( \max_{1 \leqslant n \leqslant N} |a_n| \big)^p \ \mathbb{E}\Big\| \sum_{n=1}^{N} X_n \Big\|^p.$$

*Proof.* For all $(\varepsilon_1, \dots, \varepsilon_N) \in \{-1, +1\}^N$ the $E^N$-valued random variables $(X_1, \dots, X_N)$ and $(\varepsilon_1 X_1, \dots, \varepsilon_N X_N)$ are identically distributed and therefore

$$\mathbb{E}\Big\| \sum_{n=1}^{N} \varepsilon_n X_n \Big\|^p = \mathbb{E}\Big\| \sum_{n=1}^{N} X_n \Big\|^p.$$

For the general case we may assume that $|a_n| \leqslant 1$ for all $n = 1, \dots, N$. Then $a = (a_1, \dots, a_N)$ is a convex combination of the $2^N$ elements of $\{-1, +1\}^N$, say $a = \sum_{j=1}^{2^N} \lambda^{(j)} \varepsilon^{(j)}$. Hence,

$$\begin{aligned}
\mathbb{E}\Big\| \sum_{n=1}^{N} a_n X_n \Big\|^p &= \mathbb{E}\Big\| \sum_{j=1}^{2^N} \lambda^{(j)} \sum_{n=1}^{N} \varepsilon_n^{(j)} X_n \Big\|^p \\
&\leqslant \mathbb{E}\Big( \sum_{j=1}^{2^N} \lambda^{(j)} \Big\| \sum_{n=1}^{N} \varepsilon_n^{(j)} X_n \Big\| \Big)^p \\
&\leqslant \mathbb{E} \sum_{j=1}^{2^N} \lambda^{(j)} \Big\| \sum_{n=1}^{N} \varepsilon_n^{(j)} X_n \Big\|^p \\
&= \sum_{j=1}^{2^N} \lambda^{(j)} \mathbb{E}\Big\| \sum_{n=1}^{N} X_n \Big\|^p = \mathbb{E}\Big\| \sum_{n=1}^{N} X_n \Big\|^p,
\end{aligned}$$

where the third step follows from the convexity of the function $t \mapsto t^p$ (or an application of Jensen's inequality). $\square$

As an application of the Kahane contraction principle we shall prove an inequality which shows that Rademacher sums have the 'smallest' $L^p$-norms among all random sums. Rademacher sums are easier to handle than the Gaussian sums in which we are ultimately interested, and, as we shall see, there are various techniques to pass on results for Rademacher sums to Gaussian sums.

Let us begin with a definition. An $\{-1, +1\}$-valued random variable $r$ is called a *Rademacher variable* if

$$\mathbb{P}\{r = -1\} = \mathbb{P}\{r = +1\} = \frac{1}{2}.$$

Throughout these lectures, the notation $(r_n)_{n=1}^{\infty}$ will be used for a *Rademacher sequence*, that is, a sequence of independent Rademacher variables.

**Theorem 3.2 (Comparison).** *Let $(\varphi_n)_{n=1}^{\infty}$ be a sequence of independent symmetric integrable real-valued random variables satisfying $\mathbb{E}|\varphi_n| \geqslant 1$ for all $n \geqslant 1$. Then for all $x_1, \dots, x_N \in E$ and $1 \leqslant p < \infty$ we have*

$$\mathbb{E}\Big\| \sum_{n=1}^{N} r_n x_n \Big\|^p \leqslant \mathbb{E}\Big\| \sum_{n=1}^{N} \varphi_n x_n \Big\|^p.$$

The proof of this theorem relies on an auxiliary lemma, for which we need two definitions based on the following easy observation: if $X_1, \dots, X_N$ are random variables with values in $E_1, \dots, E_N$, then $(X_1, \dots, X_N)$ is a random variable with values in $E_1 \times \dots \times E_N$.

**Definition 3.3.** *Two families of random variables $(X_i)_{i\in I}$ and $(Y_i)_{i\in I}$, where $I$ is some index set and $X_i$ and $Y_i$ take values in a Banach space $E_i$, are* identically distributed *if for all choices of $i_1,\dots,i_N\in I$ the random variables $(X_{i_1},\dots,X_{i_N})$ and $(Y_{i_1},\dots,Y_{i_N})$ are identically distributed.*

Note that by Proposition 2.13, if $(X_i)_{i\in I}$ and $(Y_i)_{i\in I}$ are families of independent random variables such that $X_i$ and $Y_i$ are identically distributed for all $i\in I$, then $(X_i)_{i\in I}$ and $(Y_i)_{i\in I}$ are identically distributed.

**Definition 3.4.** *Two families of random variables $(X_i)_{i\in I}$ and $(Y_j)_{j\in J}$, where $I$ and $J$ are index sets, $X_i$ takes values in $E_i$ for all $i\in I$ and $Y_j$ takes values in $F_j$ for all $j\in J$, are* independent of each other *if for all choices $i_1,\dots,i_M\in I$ and $j_1,\dots,j_N\in I$ the random variables $(X_{i_1},\dots,X_{i_M})$ and $(Y_{j_1},\dots,Y_{i_N})$ are independent.*

**Lemma 3.5.** *Let $(\varphi_n)_{n=1}^\infty$ be a sequence of independent symmetric real-valued random variables and let $(r_n)_{n=1}^\infty$ be a Rademacher sequence independent of $(\varphi_n)_{n=1}^\infty$. The sequences $(\varphi_n)_{n=1}^\infty$ and $(r_n|\varphi_n|)_{n=1}^\infty$ are identically distributed.*

*Proof.* By independence and symmetry we have

$$\begin{aligned}
&\mathbb{P}\{r_n|\varphi_n|\in B\}\\
&\quad=\mathbb{P}\{r_n=1,\ \varphi_n\geqslant 0,\ \varphi_n\in B\}+\mathbb{P}\{r_n=1,\ \varphi_n<0,\ \varphi_n\in -B\}\\
&\qquad+\mathbb{P}\{r_n=-1,\ \varphi_n\geqslant 0,\ \varphi_n\in -B\}+\mathbb{P}\{r_n=-1,\ \varphi_n<0,\ \varphi_n\in B\}\\
&\quad=\tfrac12\mathbb{P}\{\varphi_n\geqslant 0,\ \varphi_n\in B\}+\tfrac12\mathbb{P}\{\varphi_n<0,\ \varphi_n\in -B\}\\
&\qquad+\tfrac12\mathbb{P}\{\varphi_n\geqslant 0,\ \varphi_n\in -B\}+\tfrac12\mathbb{P}\{\varphi_n<0,\ \varphi_n\in B\}\\
&\quad=\tfrac12\mathbb{P}\{\varphi_n\geqslant 0,\ \varphi_n\in B\}+\tfrac12\mathbb{P}\{\varphi_n>0,\ \varphi_n\in B\}\\
&\qquad+\tfrac12\mathbb{P}\{\varphi_n\leqslant 0,\ \varphi_n\in B\}+\tfrac12\mathbb{P}\{\varphi_n<0,\ \varphi_n\in B\}\\
&\quad=\mathbb{P}\{\varphi_n\in B\}.
\end{aligned}$$

Since $(\varphi_n)_{n=1}^\infty$ and $(r_n|\varphi_n|)_{n=1}^\infty$ are sequences of independent random variables, the lemma now follows from the observation preceding Definition 3.4. □

*Proof (Proof of Theorem 3.2).* We may assume that the sequences $(\varphi_n)_{n=1}^\infty$ and $(r_n)_{n=1}^\infty$ are defined on distinct probability spaces $\Omega_\varphi$ and $\Omega_r$. By considering the $\varphi_n$ and $r_n$ as random variables on the probability space $\Omega_\varphi\times\Omega_r$, we may assume that $(\varphi_n)_{n=1}^\infty$ and $(r_n)_{n=1}^\infty$ are independent of each other.

Since $\mathbb{E}_\varphi|\varphi_n|\geqslant 1$, with the Kahane contraction principle and Jensen's inequality we obtain

$$\begin{aligned}
\mathbb{E}_r\Big\|\sum_{n=1}^N r_nx_n\Big\|^p &\leqslant \mathbb{E}_r\Big\|\mathbb{E}_\varphi\sum_{n=1}^N r_n|\varphi_n|x_n\Big\|^p\\
&\leqslant \mathbb{E}_r\mathbb{E}_\varphi\Big\|\sum_{n=1}^N r_n|\varphi_n|x_n\Big\|^p=\mathbb{E}_\varphi\Big\|\sum_{n=1}^N\varphi_nx_n\Big\|^p,
\end{aligned}$$

where the last identity follows from Lemma 3.5. $\square$

A real-valued random variable $\gamma$ is called *standard Gaussian* if its distribution has density

$$f_\gamma(t) = \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}t^2)$$

with respect to the Lebesgue measure on $\mathbb{R}$. For later reference we note that $\gamma$ is standard Gaussian if and only if its Fourier transform is given by

$$\mathbb{E}\exp(-i\xi\gamma) = \exp(-\tfrac{1}{2}\xi^2), \quad \xi \in \mathbb{R}. \tag{3.1}$$

The 'only if' statement follows from the identity

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-i\xi t - \tfrac{1}{2}t^2)\,dt = \exp(-\tfrac{1}{2}\xi^2)$$

which can be proved by completing the squares in the exponential and then shifting the path of integration from $i\xi+\mathbb{R}$ to $\mathbb{R}$ by using Cauchy's formula; the 'if' part then follows from the injectivity of the Fourier transform (Theorem 2.8).

For a standard Gaussian random variable $\gamma$ we have

$$\mathbb{E}|\gamma| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |t| \exp(-\tfrac{1}{2}t^2)\,dt = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} t \exp(-\tfrac{1}{2}t^2)\,dt = \sqrt{2/\pi}. \tag{3.2}$$

From this point on, $(\gamma_n)_{n=1}^\infty$ will always denote a *Gaussian sequence*, that is, a sequence of independent standard Gaussian variables.

From (3.2) and Theorem 3.2 we obtain the following comparison result.

**Corollary 3.6.** *For all $x_1, \dots, x_N \in E$ and $1 \leqslant p < \infty$,*

$$\mathbb{E}\Big\| \sum_{n=1}^N r_n x_n \Big\|^p \leqslant (\pi/2)^{\frac{p}{2}}\ \mathbb{E}\Big\| \sum_{n=1}^N \gamma_n x_n \Big\|^p. \tag{3.3}$$

The geometric notions of *type* and *cotype* will be introduced in the exercises. Without proof we state the following important converse to Corollary 3.6 for Banach spaces with finite cotype. Examples of spaces with finite cotype are Hilbert spaces, $L^p$-spaces for $1 \leqslant p < \infty$, and the UMD spaces which will be introduced in later lectures.

**Theorem 3.7.** *If $E$ has finite cotype, there exists a constant $C \geqslant 0$ such that for all $x_1, \dots, x_N \in E$,*

$$\mathbb{E}\Big\| \sum_{n=1}^N \gamma_n x_n \Big\|^2 \leqslant C^2\, \mathbb{E}\Big\| \sum_{n=1}^N r_n x_n \Big\|^2.$$

The Kahane-Khintichine inequalities (Theorems 3.11 and 3.12 below) can be used to extend this inequality to arbitrary exponents $1 \leqslant p < \infty$.

The proof of Theorem 3.7 is beyond the scope of these lectures; we refer to the Notes at the end of the lecture for references to the literature. When taken together, Corollary 3.6 and Theorem 3.7 show that in spaces with finite cotype, Gaussian sequences and Rademacher sums can be used interchangeably.

Without any assumptions on $E$, Theorem 3.7 fails. This is shown by the next example.

*Example 3.8.* Let $E = c_0$ and let $(u_n)_{n=1}^\infty$ be the standard unit basis of $c_0$. Then

$$\mathbb{E}\Big\| \sum_{n=1}^N r_n u_n \Big\|_{c_0} = \mathbb{E}\Big( \max_{1\leqslant n\leqslant N} |r_n| \Big) = 1.$$

Next we estimate $\mathbb{E}\big\| \sum_{n=1}^N \gamma_n u_n \big\|_{c_0}$ from below. First, if $\gamma$ is standard Gaussian, the inequality $1 - x \leqslant e^{-x}$ implies

$$\mathbb{P}\Big\{ \max_{1\leqslant n\leqslant N} |\gamma_n| \leqslant r \Big\} = \big[1 - \mathbb{P}\{|\gamma| > r\}\big]^N \leqslant \exp(-N\mathbb{P}\{|\gamma| > r\}).$$

For $r = \frac12\sqrt{\log N}$ we estimate

$$\begin{aligned}\mathbb{P}\{|\gamma| > \tfrac12\sqrt{\log N}\} &\geqslant \frac{2}{\sqrt{2\pi}} \int_{\frac12\sqrt{\log N}}^{\sqrt{\log N}} e^{-\frac12 x^2}\,dx \\ &\geqslant \frac{2}{\sqrt{2\pi}} \cdot \tfrac12\sqrt{\log N} \cdot e^{-\frac12 \log N} = \sqrt{\frac{\log N}{2\pi N}}.\end{aligned}$$

Hence, using the integration by parts formula of Exercise 2.1,

$$\begin{aligned}\mathbb{E}\Big\| \sum_{n=1}^N \gamma_n u_n \Big\|_{c_0} &= \mathbb{E}\Big( \max_{1\leqslant n\leqslant N} |\gamma_n| \Big) \\ &= \int_0^\infty \mathbb{P}\Big\{ \max_{1\leqslant n\leqslant N} |\gamma_n| > r \Big\}\,dr \\ &\geqslant \int_0^{\frac12\sqrt{\log N}} \big[1 - \exp(-N\mathbb{P}\{|\gamma| > r\})\big]\,dr \\ &\geqslant \tfrac12\sqrt{\log N} \cdot \Big[1 - \exp\Big(-\sqrt{\frac{N\log N}{2\pi}}\Big)\Big] \\ &\eqsim \tfrac12\sqrt{\log N} \quad \text{as} \quad N \to \infty.\end{aligned}$$

Similar estimates show that the bound $\mathscr{O}(\sqrt{\log N})$ for $N \to \infty$ is of the correct order.

We conclude this section with an important comparison result for Gaussian sums.

**Theorem 3.9 (Covariance domination).** *Let* $(\gamma_m)_{m=1}^\infty$ *and* $(\gamma'_n)_{n=1}^\infty$ *be Gaussian sequences on probability spaces* $\Omega$ *and* $\Omega'$*, respectively, and let* $x_1, \dots, x_M$ *and* $y_1, \dots, y_N$ *be elements of* $E$ *satisfying*

$$\sum_{m=1}^{M} \langle x_m, x^* \rangle^2 \leqslant \sum_{n=1}^{N} \langle y_n, x^* \rangle^2 \qquad \forall x^* \in E^*.$$

*Then, for all* $1 \leqslant p < \infty$,

$$\mathbb{E} \Big\| \sum_{m=1}^{M} \gamma_m x_m \Big\|^p \leqslant \mathbb{E}' \Big\| \sum_{n=1}^{N} \gamma'_n y_n \Big\|^p.$$

*Proof.* Denote by $F$ the linear span of $\{x_1, \dots, x_M, y_1, \dots, y_N\}$ in $E$. Define $Q \in \mathscr{L}(F^*, F)$ by

$$Qz^* := \sum_{n=1}^{N} \langle y_n, z^* \rangle y_n - \sum_{m=1}^{M} \langle x_m, z^* \rangle x_m, \qquad z^* \in F^*.$$

The assumption of the theorem implies that $\langle Qz^*, z^* \rangle \geqslant 0$ for all $z^* \in F^*$, and it is clear that $\langle Qz_1^*, z_2^* \rangle = \langle Qz_2^*, z_1^* \rangle$ for all $z_1^*, z_2^* \in F^*$. Since $F$ is finite-dimensional, by linear algebra we can find a sequence $(x_j)_{j=M+1}^{M+k}$ in $F$ such that $Q$ is represented as

$$Qz^* = \sum_{j=M+1}^{M+k} \langle x_j, z^* \rangle x_j, \quad z^* \in F^*.$$

We leave the verification of this statement as an exercise for the moment and shall return to this issue from a more general point of view in the next lecture.

Now,

$$\sum_{m=1}^{M+k} \langle x_m, z^* \rangle^2 = \sum_{n=1}^{N} \langle y_n, z^* \rangle^2, \qquad z^* \in F^*. \tag{3.4}$$

It follows from (3.1) that the random variables $X := \sum_{m=1}^{M+k} \gamma_m x_m$ and $Y := \sum_{n=1}^{N} \gamma'_n y_n$ have Fourier transforms

$$\begin{aligned} \mathbb{E} \exp(-i \langle X, x^* \rangle) &= \prod_{m=1}^{M+k} \mathbb{E} \exp(-i \gamma_m \langle x_m, x^* \rangle) \\ &= \prod_{m=1}^{M+k} \exp\big(-\frac{1}{2} \langle x_m, x^* \rangle^2\big) = \exp\Big(-\frac{1}{2} \sum_{m=1}^{M+k} \langle x_m, x^* \rangle^2\Big) \end{aligned}$$

and similarly $\mathbb{E}' \exp(-i \langle Y, x^* \rangle) = \exp(-\frac{1}{2} \sum_{n=1}^{N} \langle y_n, x^* \rangle^2)$. Hence by (3.4) and Theorem 2.8, $X$ and $Y$ are identically distributed. Thus, for all $1 \leqslant p < \infty$,

$$\mathbb{E}\Big\|\sum_{m=1}^{M+k}\gamma_m x_m\Big\|^p = \mathbb{E}'\Big\|\sum_{n=1}^{N}\gamma_n' y_n\Big\|^p.$$

By Proposition 2.16,

$$\mathbb{E}\Big\|\sum_{m=1}^{M}\gamma_m x_m\Big\|^p \leqslant \mathbb{E}\Big\|\sum_{m=1}^{M+k}\gamma_m x_m\Big\|^p,$$

and the proof is complete. □

## 3.2 The Kahane-Khintchine inequality

The main result of this section states that all $L^p$-norms of an $E$-valued Gaussian sum are comparable, with universal constants depending only on $p$. First we prove the analogous result for Rademacher sums; then we use the central limit theorem to pass it on to Gaussian sums.

The starting point is the following inequality, which is a consequence of Lévy's inequality.

**Lemma 3.10.** *For all $x_1, \dots, x_N \in E$ and $r > 0$ we have*

$$\mathbb{P}\Big\{\Big\|\sum_{n=1}^{N} r_n x_n\Big\| > 2r\Big\} \leqslant 4\Big[\mathbb{P}\Big\{\Big\|\sum_{n=1}^{N} r_n x_n\Big\| > r\Big\}\Big]^2.$$

*Proof.* Let us write $S_n := \sum_{j=1}^{n} r_j x_j$. As in the proof of Lemma 2.18 we put

$$A_n := \{\|S_1\| \leqslant r,\ \dots,\ \|S_{n-1}\| \leqslant r,\ \|S_n\| > r\}.$$

If for an $\omega \in A_n$ we have $\|S_N(\omega)\| > 2r$, then $\|S_N(\omega) - S_{n-1}(\omega)\| > r$. Now the crucial observation is that $(r_1, \dots, r_N)$ and $(r_1, \dots, r_n, r_n r_{n+1}, \dots, r_n r_N)$ are identically distributed; we leave the easy proof as an exercise. From this and the fact that $|r_n| = 1$ almost surely we obtain

$$\begin{aligned}
\mathbb{P}(A_n \cap \{\|S_N - S_{n-1}\| > r\}) &= \mathbb{P}\Big(A_n \cap \Big\{\Big\|\sum_{j=n}^{N} r_j x_j\Big\| > r\Big\}\Big) \\
&= \mathbb{P}\Big(A_n \cap \Big\{\Big\|r_n \sum_{j=n}^{N} r_j x_j\Big\| > r\Big\}\Big) \\
&= \mathbb{P}\Big(A_n \cap \Big\{\Big\|x_n + \sum_{j=n+1}^{N} r_n r_j x_j\Big\| > r\Big\}\Big) \\
&= \mathbb{P}\Big(A_n \cap \Big\{\Big\|x_n + \sum_{j=n+1}^{N} r_j x_j\Big\| > r\Big\}\Big) \\
&= \mathbb{P}(A_n \cap \{\|x_n + (S_N - S_n)\| > r\}),
\end{aligned}$$

and similarly $\mathbb{P}\{\|S_N - S_{n-1}\| > r\} = \mathbb{P}\{\|x_n + (S_N - S_n)\| > r\}$. Hence, by the independence of $A_n$ and $S_N - S_n$,

$$\begin{aligned}\mathbb{P}(A_n \cap \{\|S_N\| > 2r\}) &\leqslant \mathbb{P}(A_n \cap \{\|S_N - S_{n-1}\| > r\}) \\ &= \mathbb{P}(A_n)\mathbb{P}\{\|x_n + (S_N - S_n)\| > r\} \\ &= \mathbb{P}(A_n)\mathbb{P}\{\|S_N - S_{n-1}\| > r\} \leqslant 2\mathbb{P}(A_n)\mathbb{P}\{\|S_N\| > r\},\end{aligned}$$

where the last step follows from Lévy's inequality after changing the order of summation. Summing over $n = 1, \dots, N$ and using Lévy's inequality once more we obtain

$$\begin{aligned}\mathbb{P}\{\|S_N\| > 2r\} &= \sum_{n=1}^N \mathbb{P}(A_n \cap \{\|S_N\| > 2r\}) \leqslant 2\sum_{n=1}^N \mathbb{P}(A_n)\mathbb{P}\{\|S_N\| > r\} \\ &= 2\mathbb{P}\Big\{\max_{1\leqslant n\leqslant N} \|S_n\| > r\Big\}\mathbb{P}\{\|S_N\| > r\} \leqslant 4[\mathbb{P}\{\|S_N\| > r\}]^2. \quad \square\end{aligned}$$

We are now ready to prove the following result, which is the Banach space generalisation due to KAHANE of a classical result for scalar random variables of KHINTCHINE.

**Theorem 3.11 (Kahane-Khintchine inequality - Rademacher sums).** *For all $1 \leqslant p, q < \infty$ there exists a constant $K_{p,q}$, depending only on $p$ and $q$, such that for all finite sequences $x_1, \dots, x_N \in E$ we have*

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^N r_n x_n\Big\|^p\Big)^{\frac{1}{p}} \leqslant K_{p,q}\Big(\mathbb{E}\Big\|\sum_{n=1}^N r_n x_n\Big\|^q\Big)^{\frac{1}{q}}.$$

*Proof.* By Hölder's inequality it suffices to consider the case $p > 1$ and $q = 1$.

Fix vectors $x_1, \dots, x_N \in E$. Writing $X_n = r_n x_n$ and $S_N = \sum_{n=1}^N X_n$, we may assume that $\mathbb{E}\|S_N\| = 1$.

Let $j \geqslant 1$ be the unique integer such that $2^{j-1} < p \leqslant 2^j$. By successive applications of Lemma 3.10 for $r > 0$ we have

$$\mathbb{P}\{\|S_N\| > 2^j r\} \leqslant 4^{2^j-1}(\mathbb{P}\{\|S_N\| > r\})^{2^j}.$$

Chebyshev's inequality gives $r\mathbb{P}\{\|S_N\| > r\} \leqslant \mathbb{E}\|S_N\| = 1$. Hence,

$$\begin{aligned}\mathbb{E}\|S_N\|^p &= \int_0^\infty pt^{p-1}\mathbb{P}\{\|S_N\| > t\}\,dt \\ &= 2^{jp}\int_0^\infty pr^{p-1}\mathbb{P}\{\|S_N\| > 2^j r\}\,dr \\ &\leqslant 2^{jp}4^{2^j-1}\int_0^\infty pr^{p-1}(\mathbb{P}\{\|S_N\| > r\})^{2^j}\,dr \\ &\leqslant (2p)^p 4^{2p-1}\int_0^\infty pr^{p-1}(\mathbb{P}\{\|S_N\| > r\})^p\,dr \\ &\leqslant (2p)^p 4^{2p-1}\int_0^\infty p\mathbb{P}\{\|S_N\| > r\}\,dr \\ &\leqslant (2p)^p 4^{2p-1}p.\end{aligned}$$

□

The best possible constants $K_{p,q}$ in this inequality are called the *Kahane-Khintchine constants*. Note that $K_{p,q} = 1$ if $p \leqslant q$ by Hölder's inequality. The bound on $K_{p,1}$ produced in the above proof is not the best possible: for instance it is known that $K_{p,1} = 2^{1-\frac{1}{p}}$; see the Notes at the end of the lecture.

By an application of the central limit theorem, the Kahane-Khintchine inequality extends to Gaussian sums:

**Theorem 3.12 (Kahane-Khintchine inequality - Gaussian sums).** *For all $1 \leqslant p, q < \infty$ and all finite sequences $x_1, \dots, x_N \in E$ we have*

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^N \gamma_n x_n\Big\|^p\Big)^{\frac{1}{p}} \leqslant K_{p,q}\Big(\mathbb{E}\Big\|\sum_{n=1}^N \gamma_n x_n\Big\|^q\Big)^{\frac{1}{q}},$$

*where $K_{p,q}$ is the Kahane-Khintchine constant.*

*Proof.* Fix $k = 1, 2, \dots$ and define $\varphi_n^{(k)} := \frac{1}{\sqrt{k}}\sum_{j=1}^k r_{nk+j}$. For each $k$ we have

$$\begin{aligned}\Big(\mathbb{E}\Big\|\sum_{n=1}^N \varphi_n^{(k)} x_n\Big\|^p\Big)^{\frac{1}{p}} &= \Big(\mathbb{E}\Big\|\sum_{n=1}^N\sum_{j=1}^k r_{nk+j}\frac{x_n}{\sqrt{k}}\Big\|^p\Big)^{\frac{1}{p}} \\ &\leqslant K_{p,q}\Big(\mathbb{E}\Big\|\sum_{n=1}^N\sum_{j=1}^k r_{nk+j}\frac{x_n}{\sqrt{k}}\Big\|^q\Big)^{\frac{1}{q}} = K_{p,q}\Big(\mathbb{E}\Big\|\sum_{n=1}^N \varphi_n^{(k)} x_n\Big\|^q\Big)^{\frac{1}{q}}.\end{aligned}$$

The proof is completed by passing to the limit $k \to \infty$ and using the central limit theorem. □

The attentive reader has noticed that we are cheating a bit in the above proof, as the usual formulation of the central limit theorem only asserts that $\lim_{k\to\infty}(\varphi_1^{(k)}, \dots, \varphi_N^{(k)}) = (\gamma_1, \dots, \gamma_N)$ in distribution, that is,

$$\lim_{k\to\infty} \mathbb{E}f(\varphi_1^{(k)}, \dots, \varphi_N^{(k)}) = \mathbb{E}f(\gamma_1, \dots, \gamma_N)$$

for all bounded continuous functions $f : \mathbb{R}^N \to \mathbb{R}$. We will show next how, in the present situation, the convergence of the $L^r$-norms (with $r = p, q$) of the sums can be deduced from this. The main idea is contained in the next lemma.

**Lemma 3.13.** *Suppose $\varphi_0, \varphi_1, \dots$ and $\varphi$ are $\mathbb{R}^N$-valued random variables such that for all bounded continuous functions $f : \mathbb{R}^N \to \mathbb{R}$ we have*

$$\lim_{k\to\infty} \mathbb{E}f(\varphi_k) = \mathbb{E}f(\varphi).$$

*Let* $\Phi : \mathbb{R}^N \to \mathbb{R}$ *be a Borel function such that* $\sup_{k\geqslant 1} \mathbb{E}\,|\Phi(\varphi_k)| < \infty$ *and* $\mathbb{E}\,|\Phi(\varphi)| < \infty$*. If* $g : \mathbb{R}^N \to \mathbb{R}$ *is a continuous function satisfying*

$$|g(t)| \leqslant |c(t)||\Phi(t)|, \quad t \in \mathbb{R}^N,$$

*where* $c : \mathbb{R}^N \to \mathbb{R}$ *is a bounded function satisfying* $\lim_{|t|\to\infty} |c(t)| = 0$*, then*

$$\lim_{k\to\infty} \mathbb{E}g(\varphi_k) = \mathbb{E}g(\varphi).$$

*Proof.* Let $g_R := g \cdot 1_{\{|g|<R\}} + R \cdot 1_{\{g \geqslant R\}} - R \cdot 1_{\{g \leqslant -R\}}$ denote the truncation of $g$ at the levels $\pm R$. By assumption we have

$$\lim_{k\to\infty} \mathbb{E}g_R(\varphi_k) = \mathbb{E}g_R(\varphi). \tag{3.5}$$

Furthermore, by dominated convergence,

$$\lim_{R\to\infty} \mathbb{E}g_R(\varphi) = \mathbb{E}g(\varphi). \tag{3.6}$$

Fix $\varepsilon > 0$ and choose $R_0 > 0$ so large that $\sup_{|t|>R_0} |c(t)| < \varepsilon$. Choose $R_1 > 0$ so large that $|g(t)| > R_1$ implies $|t| > R_0$. Then, for all $R \geqslant R_1$,

$$\begin{aligned}\sup_{k\geqslant 0} \mathbb{E}|g(\varphi_k) - g_R(\varphi_k)| &\leqslant \sup_{k\geqslant 0} \mathbb{E}(1_{\{|g|>R\}}(\varphi_k)|g(\varphi_k)|) \\ &\leqslant \sup_{k\geqslant 0} \mathbb{E}(1_{\{|g|>R\}}(\varphi_k)|c(t)||\Phi(\varphi_k)|) \\ &\leqslant \varepsilon \sup_{k\geqslant 0} \mathbb{E}|\Phi(\varphi_k)|,\end{aligned} \tag{3.7}$$

Combined with (3.6) and (3.5), this gives the desired result. □

Now we can finish the proof of Theorem 3.12:

**Lemma 3.14.** *With the notations of Theorem 3.12, for all* $1 \leqslant r < \infty$ *and* $x_1, \ldots, x_N \in E$ *we have*

$$\lim_{k\to\infty} \mathbb{E}\Big\| \sum_{n=1}^{N} \varphi_n^{(k)} x_n \Big\|^r = \mathbb{E}\Big\| \sum_{n=1}^{N} \gamma_n x_n \Big\|^r,$$

*where* $\gamma_1, \ldots, \gamma_N$ *are independent standard Gaussian variables.*

*Proof.* Without loss of generality we may assume that $\max_{1\leqslant n\leqslant N} \|x_n\| \leqslant 1$. We fix $1 \leqslant r < \infty$ and check the condition of Lemma 3.13 for the functions $\Phi : \mathbb{R}^N \to \mathbb{R}$ and $g : \mathbb{R}^N \to \mathbb{R}$ defined by

$$\Phi(t) := \exp\Big( \sum_{n=1}^{N} |t_n| \|x_n\| \Big), \qquad g(t) := \Big\| \sum_{n=1}^{N} t_n x_n \Big\|^r,$$

where $\varphi_k := (\varphi_1^{(k)}, \ldots, \varphi_N^{(k)})$ and $\varphi := (\gamma_1, \ldots, \gamma_N)$.

If $\varphi$ is a symmetric real-valued random variable, then

$$\begin{aligned}\mathbb{E}\exp(|\varphi|) &= \mathbb{E}\exp(-1_{\{\varphi<0\}}\varphi) + \mathbb{E}\exp(1_{\{\varphi\geqslant 0\}}\varphi)\\ &= \mathbb{E}\exp(1_{\{-\varphi<0\}}\varphi) + \mathbb{E}\exp(1_{\{\varphi\geqslant 0\}}\varphi) \leqslant 2\mathbb{E}\exp(\varphi).\end{aligned}$$

Hence, since $\max_{1\leqslant n\leqslant N}\|x_n\| \leqslant 1$,

$$\begin{aligned}\mathbb{E}\,\Phi(\varphi_k) &\leqslant \prod_{n=1}^{N}\mathbb{E}\exp(|\varphi_n^{(k)}|) \leqslant 2^N\prod_{n=1}^{N}\mathbb{E}\exp(\varphi_n^{(k)})\\ &= 2^N\prod_{n=1}^{N}\prod_{j=1}^{k}\mathbb{E}\exp\big(\frac{r_{nk+j}}{\sqrt{k}}\big) = 2^N\Big(\frac12\exp\big(\frac{1}{\sqrt{k}}\big)+\frac12\exp\big(\frac{-1}{\sqrt{k}}\big)\Big)^{kN}\\ &= 2^N\mathscr{O}\Big(1+\frac{1}{2k}\Big)^{kN} = 2^N\exp(N/2)\cdot\mathscr{O}(1) \quad \text{as } k\to\infty.\end{aligned}$$

□

## 3.3 Exercises

1. Let $(X_n)_{n=1}^N$ be a sequence of independent symmetric $E$-valued random variables, and let $(r_n)_{n=1}^N$ be a Rademacher sequence which is independent of $(X_n)_{n=1}^N$. Prove that the sequences $(X_n)_{n=1}^N$ and $(r_nX_n)_{n=1}^N$ are identically distributed.

   *Hint:* As in the proof of Theorem 3.2 it may be assumed that $(X_n)_{n=1}^N$ and $(r_n)_{n=1}^N$ are defined on distinct probability spaces. Use Fubini's theorem together with the result of Exercise 2.2.

   *Remark:* This technique for introducing Rademacher variables is known as *randomisation*. It enables one to apply inequalities for Rademacher sums in $E$ to sums of independent symmetric random variables in $E$.

2. (!) Let $(r'_n)_{n=1}^\infty$ and $(r''_n)_{n=1}^\infty$ be independent Rademacher sequences on probability spaces $(\Omega',\mathscr{F}',\mathbb{P}')$ and $(\Omega'',\mathscr{F}'',\mathbb{P}'')$. Prove that on the product $(\Omega,\mathscr{F},\mathbb{P}) = (\Omega'\times\Omega'',\mathscr{F}'\otimes\mathscr{F}'',\mathbb{P}'\otimes\mathbb{P}'')$, the sequence $(r'_mr''_n)_{m,n=1}^\infty$ consists of Rademacher variables, but as a (doubly indexed) sequence it fails to be a Rademacher sequence (that is, the random variables $r'_mr''_n$ fail to be independent).

3. (!) We continue with the notations of the previous exercise. Prove that for $1\leqslant p<\infty$ the following version of the contraction principle holds for double Rademacher sums in the spaces $L^p(A)$, where $(A,\mathscr{A},\mu)$ is a $\sigma$-finite measure space: there exists a constant $C_p\geqslant 0$ such that for all finite sequences $(f_{mn})_{m,n=1}^N$ in $L^p(A)$ and all scalars $(a_{mn})_{m,n=1}^N$ we have

$$\mathbb{E}\Big\|\sum_{m,n=1}^{N}a_{mn}r'_mr''_nf_{mn}\Big\|^p \leqslant C_p^p\big(\max_{1\leqslant m,n\leqslant N}|a_{mn}|^p\big)\mathbb{E}\Big\|\sum_{m,n=1}^{N}r'_mr''_nf_{mn}\Big\|^p.$$

*Hint:* Proceed in three steps: (i) the result holds for $E = \mathbb{R}$ with exponent 2; (ii) the result holds for $E = \mathbb{R}$ with exponent $p$; (iii) the result holds for $E = L^p(A)$ with exponent $p$.

4. Let $1 \leqslant p \leqslant 2$. A Banach space $E$ is said to have *type* $p$ if there exists a constant $C_p \geqslant 0$ such that for all finite sequences $x_1, \dots, x_N$ in $E$ we have

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^N r_n x_n\Big\|^2\Big)^{\frac12} \leqslant C_p\Big(\sum_{n=1}^N \|x_n\|^p\Big)^{\frac1p}.$$

Let $2 \leqslant q \leqslant \infty$. The space $E$ is said to have *cotype* $q$ if there exists a constant $C_q \geqslant 0$ such that for all finite sequences $x_1, \dots, x_N$ in $E$ we have

$$\Big(\sum_{n=1}^N \|x_n\|^q\Big)^{\frac1q} \leqslant C_q\Big(\mathbb{E}\Big\|\sum_{n=1}^N r_n x_n\Big\|^2\Big)^{\frac12}.$$

For $q = \infty$ we make the obvious adjustment in the second definition. Prove the following assertions:

a) Every Banach space has type 1 and cotype $\infty$ (accordingly, a Banach space is said to have *non-trivial type* if it has type $p \in (1, 2]$ and *finite cotype* if it has cotype $q \in [2, \infty)$).
b) Every Hilbert space has type 2 and cotype 2.
c) If a Banach space has type $p$ for some $p \in [1, 2]$, then it has type $p'$ for all $p' \in [1, p]$; if a Banach space has cotype $q$ for some $q \in [2, \infty]$, then it has cotype $q'$ for all $q' \in [q, \infty]$.
d) Let $p \in [1, 2]$. Prove that if $E$ has type $p$, then the dual space $E^*$ has cotype $p'$, $\frac1p + \frac1{p'} = 1$.

*Hint:* For each $x_n^* \in E^*$ choose $x_n \in E$ of norm one such that $\|x_n^*\| \geqslant \frac12 |\langle x_n, x_n^*\rangle|$. Then use Hölder's inequality to the effect that for all scalar sequences $(b_n)_{n=1}^N$ one has

$$\Big(\sum_{n=1}^N |b_n|^{p'}\Big)^{\frac1{p'}} = \sup\Big\{\sum_{n=1}^N a_n b_n : \ \Big(\sum_{n=1}^N |a_n|^p\Big)^{\frac1p} \leqslant 1\Big\}.$$

*Remark:* The analogous result for spaces with cotype fails. Indeed, the reader is invited to check that $l^1$ has cotype 2 while its dual $l^\infty$ fails to have non-trivial type.

5. Let $p \in [1, 2]$. Prove that a Banach space $E$ has type $p$ if and only if it has *Gaussian type* $p$, that is, if and only if there exists a constant $C \geqslant 0$ such that for all finite sequences $x_1, \dots, x_N$ in $E$ we have

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^N \gamma_n x_n\Big\|^2\Big)^{\frac12} \leqslant C\Big(\sum_{n=1}^N \|x_n\|^p\Big)^{\frac1p}.$$

*Hint*: One direction follows from Corollary 3.6. For the other direction use a randomisation argument.

*Remark:* The corresponding assertion for cotype is also true but much harder to prove; see the Notes.

**Notes.** The results of this lecture are classical and can be found in many textbooks. Our presentation borrows from ALBIAC and KALTON [1] and DIESTEL, JARCHOW, TONGE [35]. Both are excellent starting points for further reading.

The Kahane contraction principle is due to KAHANE [54], who also extended the classical scalar Khintchine inequality to arbitrary Banach spaces. It is an open problem to determine the best constants $K_{p,q}$ in the Kahane-Khintchine inequality; a recent result of LATALA and OLESZKIEWICZ [67] asserts that the constant $K_{p,1} = 2^{1-\frac{1}{p}}$ is optimal for $1 \leqslant p \leqslant 2$.

For a proof of Theorem 3.7 see, e.g., [35]. The proofs of Theorems 3.9 and 3.11 are taken from ALBIAC and KALTON [1]. The central limit argument in Lemma 3.14 is adapted from TOMCZAK-JAEGERMANN [102].

The contraction principle for double Rademacher sums of Exercise 3 has been introduced by PISIER [92]. This property, nowadays known under the rather unsuggestive name 'property $(\alpha)$' plays an important role in many advanced results in Banach space-valued harmonic analysis. It can be shown that the Rademachers can be replaced by Gaussians without changing the class of spaces under consideration. Not every Banach space has property $(\alpha)$; a counterexample is the space $c_0$.

The notions of type and cotype were developed in the 1970s by MAUREY and PISIER. As we have seen in Exercise 4, Hilbert spaces have type 2 and cotype 2. A celebrated theorem of KWAPIEŃ [64] asserts that Hilbert spaces are the only spaces with this property: a Banach space $E$ is isomorphic to a Hilbert space if and only if $E$ has type 2 and cotype 2. Another class of spaces of which the type and cotype can be computed are the $L^p$-spaces. For the interested reader we include a proof that the spaces $L^p(A)$, with $1 \leqslant p < \infty$ and $(A, \mathscr{A}, \mu)$ $\sigma$-finite, have type $\min\{p, 2\}$. A similar argument can be used to prove that they have cotype $\max\{p, 2\}$.

Let $f_1, \ldots, f_N \in L^p(A)$ and put $r := \min\{p, 2\}$. Using the Fubini theorem, the scalar Kahane-Khintchine inequality, the type $p$ inequality, Hölder's inequality, and the triangle inequality in $L^{\frac{p}{r}}(A)$, we obtain

$$
\begin{aligned}
\Big(\mathbb{E}\Big\|\sum_{n=1}^N r_n f_n\Big\|_{L^p(A)}^p\Big)^{\frac1p} &= \Big(\int_A \mathbb{E}\Big|\sum_{n=1}^N r_n f_n(\xi)\Big|^p\,d\mu(\xi)\Big)^{\frac1p}\\
&\leqslant K_{p,2}\Big(\int_A \Big(\mathbb{E}\Big|\sum_{n=1}^N r_n f_n(\xi)\Big|^2\Big)^{\frac p2}\,d\mu(\xi)\Big)^{\frac1p}\\
&= K_{p,2}\Big(\int_A \Big(\sum_{n=1}^N |f_n(\xi)|^2\Big)^{\frac p2}\,d\mu(\xi)\Big)^{\frac1p}\\
&\leqslant K_{p,2}\Big(\int_A \Big(\sum_{n=1}^N |f_n(\xi)|^r\Big)^{\frac pr}\,d\mu(\xi)\Big)^{\frac1p}\\
&= K_{p,2}\Big\|\sum_{n=1}^N |f_n|^r\Big\|_{L^{\frac pr}(A)}^{\frac1r}\\
&\leqslant K_{p,2}\Big(\sum_{n=1}^N \Big\||f_n|^r\Big\|_{L^{\frac pr}(A)}\Big)^{\frac1r}\\
&= K_{p,2}\Big(\sum_{n=1}^N \|f_n\|_{L^p(A)}^r\Big)^{\frac1r}.
\end{aligned}
$$

An application of the Kahane-Khintchine inequality for $L^p(A)$ to replace the $L^p$-moment in the left hand side by the $L^2$-moment finishes the proof.

It was noted in Exercise 4 that if $E$ has type $p$, then $E^*$ has cotype $p'$ (where $\frac1p + \frac1{p'} = 1$) and that the analogous duality result for cotype fails. It is a deep result of Pisier [93] that if $E$ has cotype $q \in [2,\infty)$ *and non-trivial type*, then $E^*$ has type $q'$, $\frac1q + \frac1{q'} = 1$.

The fact that a Banach space has cotype $q$ if and only if it has Gaussian cotype $q$ can be deduced from a deep result of Maurey and Pisier (see [1, Chapter 11]) which gives a purely geometric characterisation of type and cotype. For the details we refer to [35].

# 4

# Gaussian random variables

Having studied $E$-valued Gaussian sums of the form $\sum_{n=1}^{N} \gamma_n x_n$ in the previous lecture, we now turn to general theory of Gaussian random variables with values in a Banach space $E$. The results of this lecture will be important for the construction of an $E$-valued stochastic integral with respect to Brownian motion.

We start with a proof of the Fernique theorem on integrability of Gaussian random variables. This theorem makes it possible to investigate $L^p$-convergence of sequences of Gaussian random variables. As it turns out, every $E$-valued Gaussian random variable can be represented in a canonical way as an $L^p$-convergent (finite or infinite) sum $\sum_{n\geqslant 1} \gamma_n x_n$. This representation theorem permits us to extend the covariance domination principle and the Kahane-Khintchine inequality to arbitrary $E$-valued Gaussians.

## 4.1 Fernique's theorem

A real-valued random variable $\gamma$ is called *Gaussian* if there exists a number $q \geqslant 0$ such that its Fourier transform is given by

$$\mathbb{E}\big(\exp(-i\xi\gamma)\big) = \exp(-\tfrac{1}{2}q\xi^2), \qquad \xi \in \mathbb{R}.$$

By uniqueness of Fourier transforms one deduces that $\gamma = 0$ almost surely if $q = 0$, and that $\gamma$ has a distribution with density

$$f_\gamma(t) = \frac{1}{\sqrt{2\pi q}} \exp(\frac{-t^2}{2q})$$

if $q > 0$. It follows that $\mathbb{E}\gamma = 0$ and $\mathbb{E}\gamma^2 = q$, which means that $\gamma$ is centred and has variance $q$. We call $\gamma$ *standard Gaussian* if $q = 1$; this definition is consistent with the one given in Lecture 3.

Let $E$ be a real Banach space.

**Definition 4.1.** *An $E$-valued random variable $X$ is* Gaussian *if the real-valued random variable $\langle X, x^*\rangle$ is Gaussian for all $x^* \in E^*$.*

Much of the theory of Banach space-valued Gaussian random variables depends on a fundamental integrability result due to Fernique. For its proof we need a lemma.

**Lemma 4.2.** *Let $X$ and $Y$ be independent and identically distributed $E$-valued Gaussian random variables. Then $U := (X+Y)/\sqrt{2}$ and $V := (X-Y)/\sqrt{2}$ are independent and have the same distribution as $X$ and $Y$.*

*Proof.* Let $\mu$ be the common distribution of $X$ and $Y$. Then $\widehat{\mu}(x^*) = \exp(-\frac{1}{2}q(x^*))$, where $q(x^*) = \mathbb{E}\langle X, x^*\rangle^2 = \mathbb{E}\langle Y, x^*\rangle^2$. Using the independence of $X$ and $Y$ we have

$$\begin{aligned}\mathbb{E}\exp(-i\langle U, x^*\rangle) &= \mathbb{E}\exp(-i\tfrac{1}{2}\sqrt{2}\langle X, x^*\rangle)\mathbb{E}\exp(-i\tfrac{1}{2}\sqrt{2}\langle Y, x^*\rangle)\\ &= \exp(-\tfrac{1}{4}q(x^*))\exp(-\tfrac{1}{4}q(x^*)) = \exp(-\tfrac{1}{2}q(x^*)).\end{aligned}$$

By the uniqueness theorem for the Fourier transform, this shows that $U$ has the same distribution as $X$ and $Y$. A similar computation shows that $V$ has the same distribution as $X$ and $Y$.

We will prove that $U$ and $V$ are independent by checking that $\mu_{(U,V)} = \mu \times \mu$, where $\mu_{(U,V)}$ is the distribution of the $E \times E$-valued random variable $(U, V)$. Identifying $(E \times E)^*$ with $E^* \times E^*$ with pairing $\langle (x, y), (x^*, y^*)\rangle = \langle x, x^*\rangle + \langle y, y^*\rangle$, by the uniqueness theorem for the Fourier transform it is enough to prove that $\widehat{\mu}_{(U,V)}(x^*, y^*) = \widehat{\mu}(x^*)\widehat{\mu}(y^*)$ for all $x^*, y^* \in E^*$. But this follows from

$$\begin{aligned}\widehat{\mu}_{(U,V)}(x^*, y^*) &= \mathbb{E}\,\exp\Big(-i\big(\langle U, x^*\rangle + \langle V, y^*\rangle\big)\Big)\\ &= \mathbb{E}\,\exp\Big(-\tfrac{1}{2}i\sqrt{2}\big(\langle X, x^* + y^*\rangle + \langle Y, x^* - y^*\rangle\big)\Big)\\ &= \mathbb{E}\,\exp\big(-\tfrac{1}{2}i\sqrt{2}\langle X, x^* + y^*\rangle\big)\mathbb{E}\,\exp\big(-\tfrac{1}{2}i\sqrt{2}\langle Y, x^* - y^*\rangle\big)\\ &= \exp\big(-\tfrac{1}{4}q(x^* + y^*)\big)\exp\big(-\tfrac{1}{4}q(x^* - y^*)\big)\\ &= \exp\big(-\tfrac{1}{2}(q(x^*) + q(y^*))\big)\\ &= \widehat{\mu}(x^*)\widehat{\mu}(y^*).\end{aligned}$$

□

**Theorem 4.3 (Fernique).** *Let $X$ be an $E$-valued Gaussian variable. There exists a constant $\beta > 0$ such that*

$$\mathbb{E}\exp(\beta\|X\|^2) < \infty. \tag{4.1}$$

*Proof.* On a possibly larger probability space, let $X'$ be independent copy of $X$. For instance, identify $X$ with the random variable $X(\omega_1, \omega_2) := X(\omega_1)$ on $\Omega \times \Omega$ and define $X'$ on $\Omega \times \Omega$ by $X'(\omega_1, \omega_2) := X(\omega_2)$.

Fix $t \geqslant s > 0$. By the lemma,

$$
\begin{aligned}
&\mathbb{P}\{\|X\| \leqslant s\} \cdot \mathbb{P}\{\|X'\| > t\} \\
&\quad = \mathbb{P}\Big\{\Big\|\frac{X - X'}{\sqrt{2}}\Big\| \leqslant s\Big\} \cdot \mathbb{P}\Big\{\Big\|\frac{X + X'}{\sqrt{2}}\Big\| > t\Big\} \\
&\quad \leqslant \mathbb{P}\Big\{\Big|\frac{\|X\| - \|X'\|}{\sqrt{2}}\Big| \leqslant s,\ \frac{\|X\| + \|X'\|}{\sqrt{2}} > t\Big\} \\
&\quad \overset{(*)}{\leqslant} \mathbb{P}\Big\{\|X\| > \frac{t-s}{\sqrt{2}},\ \|X'\| > \frac{t-s}{\sqrt{2}}\Big\} \\
&\quad = \mathbb{P}\Big\{\|X\| > \frac{t-s}{\sqrt{2}}\Big\} \cdot \mathbb{P}\Big\{\|X'\| > \frac{t-s}{\sqrt{2}}\Big\},
\end{aligned}
$$

where in $(*)$ we used that the set

$$\big\{(\xi, \eta) \in \mathbb{R}_+^2 :\ |\xi - \eta| \leqslant s\sqrt{2} \ \text{ and } \ \xi + \eta > t\sqrt{2}\big\}$$

is contained in the set

$$\Big\{(\xi, \eta) \in \mathbb{R}_+^2 :\ \xi > \frac{t-s}{\sqrt{2}} \ \text{ and } \ \eta > \frac{t-s}{\sqrt{2}}\Big\}.$$

Hence, since $X$ and $X'$ have the same distribution,

$$\mathbb{P}\{\|X\| \leqslant s\}\mathbb{P}\{\|X\| > t\} \leqslant \Big(\mathbb{P}\Big\{\|X\| > \frac{t-s}{\sqrt{2}}\Big\}\Big)^2. \tag{4.2}$$

Choose $r > 0$ such that $\mathbb{P}\{\|X\| \leqslant r\} \geqslant \frac{2}{3}$. Define $t_0 := r$ and $t_n := r + \sqrt{2}t_{n-1}$ for $n \geqslant 1$. By induction it is checked that $t_n = r((\sqrt{2})^{n+1} - 1)/(\sqrt{2} - 1)$, so $t_n \leqslant r(\sqrt{2})^{n+4}$. Put

$$\alpha_n := \frac{\mathbb{P}\{\|X\| > t_n\}}{\mathbb{P}\{\|X\| \leqslant r\}}.$$

Note that $\alpha_0 \leqslant (1 - \frac{2}{3})/\frac{2}{3} = \frac{1}{2}$. From (4.2) with $s = r$, $t = t_{n+1}$ we obtain

$$\alpha_{n+1} \leqslant \left(\frac{\mathbb{P}\{\|X\| > t_n\}}{\mathbb{P}\{\|X\| \leqslant r\}}\right)^2 = \alpha_n^2.$$

Therefore $\alpha_n \leqslant \alpha_0^{2^n} \leqslant 2^{-2^n}$ and it follows that

$$\mathbb{P}\{\|X\| > t_n\} = \alpha_n\mathbb{P}\{\|X\| \leqslant r\} \leqslant 2^{-2^n}.$$

Using these estimates we obtain

$$\begin{aligned}\mathbb{E}(\exp(\beta\|X\|^2)) &\leqslant \mathbb{P}\{\|X\| \leqslant t_0\} \cdot \exp(\beta t_0^2) \\ &\qquad + \sum_{n=0}^{\infty} \mathbb{P}\{t_n < \|X\| \leqslant t_{n+1}\} \cdot \exp(\beta t_{n+1}^2) \\ &\leqslant \exp(\beta r^2) + \sum_{n=0}^{\infty} 2^{-2^n} \exp(\beta r^2 2^{n+5}) \\ &= \exp(\beta r^2) + \sum_{n=0}^{\infty} \exp(2^n[-\log 2 + 32\beta r^2]),\end{aligned}$$

and this sum converges if $\beta > 0$ is taken small enough. □

In what follows we need much less: it will suffice to know that $\mathbb{E}\|X\|^p < \infty$ for all $1 \leqslant p < \infty$.

As a simple corollary to Fernique's theorem we note that the expectation of a Gaussian random variable is well-defined. In fact we have the following result:

**Corollary 4.4.** *If $X$ is $E$-valued Gaussian, then $\mathbb{E}X = 0$.*

*Proof.* For all $x^* \in E^*$ we have $\langle \mathbb{E}X, x^*\rangle = \mathbb{E}\langle X, x^*\rangle = 0$ and we may appeal to the Hahn-Banach theorem. □

## 4.2 The covariance operator

In order to characterise Gaussian variables in terms of their Fourier transforms we introduce the following terminology.

**Definition 4.5.** *A bounded operator $Q \in \mathscr{L}(E^*, E)$ is called*

- positive*, if $\langle Qx^*, x^*\rangle \geqslant 0$ for all $x^* \in E^*$;*
- symmetric*, if $\langle Qx^*, y^*\rangle = \langle Qy^*, x^*\rangle$ for all $x^*, y^* \in E^*$.*

**Proposition 4.6.** *For an $E$-valued random variable $X$ the following assertions are equivalent:*

(1) *$X$ is Gaussian;*
(2) *there exists a positive symmetric operator $Q \in \mathscr{L}(E^*, E)$ such that the Fourier transform of $X$ is given by*

$$\mathbb{E}\exp(-i\langle X, x^*\rangle) = \exp\bigl(-\tfrac{1}{2}\langle Qx^*, x^*\rangle\bigr), \qquad x^* \in E^*.$$

*The operator $Q$ is uniquely determined by* (2)*. Moreover,*

$$\mathbb{E}\langle X, x^*\rangle^2 = \langle Qx^*, x^*\rangle, \qquad x^* \in E^*.$$

*Proof.* (1)⇒(2): Since $X$ is square integrable by Theorem 4.3, the random variable $\langle X, x^*\rangle X$ is integrable and we may define

$$Qx^* := \mathbb{E}\langle X, x^*\rangle X, \quad x^* \in E^*.$$

From $\langle Qx^*, y^*\rangle = \mathbb{E}\langle X, x^*\rangle\langle X, y^*\rangle$ we see that $Q$ is positive and symmetric. Since $\langle X, x^*\rangle$ is Gaussian with variance $\mathbb{E}\langle X, x^*\rangle^2 = \langle Qx^*, x^*\rangle$, we have

$$\mathbb{E}\exp(-i\langle X, x^*\rangle) = \exp(-\tfrac{1}{2}\langle Qx^*, x^*\rangle).$$

(2)⇒(1): Replacing $x^*$ by $\xi x^*$ in the assumption, we see that the Fourier transform of $\langle X, x^*\rangle$ equals

$$\mathbb{E}\exp(-i\xi\langle X, x^*\rangle) = \exp\bigl(-\tfrac{1}{2}\xi^2\langle Qx^*, x^*\rangle\bigr).$$

Thus $\langle X, x^*\rangle$ is Gaussian with variance $\langle Qx^*, x^*\rangle$.

If $R$ is another positive symmetric operator satisfying condition (2), then $\langle Qx^*, x^*\rangle = \langle Rx^*, x^*\rangle$ for all $x^* \in E^*$. By polarisation this implies $\langle Qx^*, y^*\rangle = \langle Rx^*, y^*\rangle$ for all $x^*, y^* \in E^*$, and therefore $Q = R$. □

The operator $Q$ is called the *covariance operator* of $X$. The reader is warned that not every positive symmetric operator $Q \in \mathscr{L}(E^*, E)$ is the covariance of an $E$-valued random variable $X$. This may happen even if $E$ is a separable infinite-dimensional Hilbert space (see Exercise 2).

**Corollary 4.7.** *Every $E$-valued Gaussian random variable is symmetric.*

*Proof.* Just note that $X$ and $-X$ have the same Fourier transforms. □

We proceed with two simple constructions to produce new Gaussian variables from old ones. The first asserts that sums of independent Gaussian variables are Gaussian.

**Proposition 4.8.** *Let $X_1, \dots, X_N$ be independent $E$-valued Gaussian random variables with covariance operators $Q_1, \dots, Q_N$. Then the sum $X := \sum_{n=1}^N X_n$ is Gaussian with covariance operator $Q = \sum_{n=1}^N Q_n$.*

*Proof.* For all $x^* \in E^*$ we have, by independence,

$$\begin{aligned}\mathbb{E}\exp(-i\langle X, x^*\rangle) &= \mathbb{E}\prod_{n=1}^N \exp(-i\langle X_n, x^*\rangle) = \prod_{n=1}^N \mathbb{E}(\exp(-i\langle X_n, x^*\rangle))\\ &= \prod_{n=1}^N \exp(-\tfrac{1}{2}\langle Q_n x^*, x^*\rangle) = \exp(-\tfrac{1}{2}\langle Qx^*, x^*\rangle).\end{aligned}$$ □

Compositions of Gaussians with bounded operators are Gaussian again:

**Proposition 4.9.** *If $X$ is $E$-valued Gaussian with covariance operator $Q$, and if $T \in \mathscr{L}(E,F)$ is a bounded operator, then $TX$ is $F$-valued Gaussian with covariance operator $TQT^*$.*

*Proof.* This follows by computing the Fourier transform of $TX$:

$$\begin{aligned}\mathbb{E}(\exp(-i\langle TX, x^*\rangle)) &= \mathbb{E}(\exp(-i\langle X, T^*x^*\rangle)) \\ &= \exp(-\tfrac{1}{2}\langle QT^*x^*, T^*x^*\rangle)) = \exp(-\tfrac{1}{2}\langle TQT^*x^*, x^*\rangle)). \qquad \square\end{aligned}$$

As an application we prove next that if the $E$-valued random variables $X_1, \dots, X_N$ are *jointly Gaussian*, that is, if the $E^N$-valued random variable $(X_1, \dots, X_N)$ is Gaussian, then $X_1, \dots, X_N$ are independent if and only if they are *uncorrelated* in the sense that

$$\mathbb{E}\langle X_m, x^*\rangle\langle X_n, y^*\rangle = 0, \quad \forall m \neq n, \ x^*, y^* \in E^*.$$

**Proposition 4.10.** *Let $X_1, \dots, X_N$ be $E$-valued random variables such that the $E^N$-valued random variable $X = (X_1, \dots, X_N)$ is Gaussian. The following assertions are equivalent:*

(1) $X_1, \dots, X_N$ *are independent;*
(2) $X_1, \dots, X_N$ *are uncorrelated.*

*Proof.* We proceed in two steps.

*Step 1* – First we consider the scalar case. Let $\gamma_1, \dots, \gamma_N$ be real-valued random variables such that the $\mathbb{R}^N$-valued random variable $\gamma = (\gamma_1, \dots, \gamma_N)$ is Gaussian. Note that each $\gamma_n$ is Gaussian; this follows from Proposition 4.9 by applying coordinate projections. We shall prove that $\gamma_1, \dots, \gamma_N$ are independent if and only if $\gamma_1, \dots, \gamma_N$ are uncorrelated.

The 'only if' part follows from $\mathbb{E}\gamma_m\gamma_n = \mathbb{E}\gamma_m\mathbb{E}\gamma_n = 0$ for all $m \neq n$. For the 'if' part we note that if $\gamma_1, \dots, \gamma_N$ are uncorrelated, the covariance matrix of $\gamma$ is diagonal: $Q = \operatorname{diag}(q_1, \dots, q_N)$ with $q_n = \mathbb{E}\,\gamma_n^2$. Then the Fourier transform of $\gamma$ is given by

$$\begin{aligned}\mathbb{E}\big(\exp(-i\langle\gamma,\xi\rangle)\big) &= \exp(-\tfrac{1}{2}\langle Q\xi,\xi\rangle) = \exp\big(-\tfrac{1}{2}\sum_{n=1}^N q_n\xi_n^2\big) \\ &= \prod_{n=1}^N \exp(-\tfrac{1}{2}q_n\xi_n^2) = \prod_{n=1}^N \mathbb{E}\exp(-i\xi_n\gamma_n), \quad \xi \in \mathbb{R}^N.\end{aligned}$$

Let $\mu$ and $\mu_n$ denote the distributions of $\gamma$ and $\gamma_n$, respectively. The above identity implies that $\mu$ and the product measure $\mu_1 \times \dots \times \mu_N$ have the same Fourier transform. Hence from Theorem 2.8 we deduce that $\mu = \mu_1 \times \dots \times \mu_N$. This implies that $\gamma_1, \dots, \gamma_N$ are independent.

*Step 2* – Next we turn to the proof of the proposition. For all choices of $x_1^*, \dots, x_N^* \in E^*$ the $\mathbb{R}^N$-valued random variable $(\langle X_1, x_1^*\rangle, \dots, \langle X_N, x_N^*\rangle)$

is Gaussian by Proposition 4.9, since it is the image of $(X_1, \dots, X_N)$ under the linear transformation from $E^N$ to $\mathbb{R}^N$, $(x_1, \dots, x_n) \mapsto (\langle x_1, x_1^* \rangle, \dots, \langle x_N, x_N^* \rangle)$.

(1)⇒(2): This implication follows from the corresponding implication in Step 1 since the independence of $X_1, \dots, X_n$ implies the independence of $\langle X_1, x_1^* \rangle, \dots, \langle X_N, x_N^* \rangle$.

(2)⇒(1): By Step 1, for all $x_1^*, \dots, x_N^* \in E^*$ the random variables $\langle X_1, x_1^* \rangle, \dots, \langle X_N, x_N^* \rangle$ are independent and therefore

$$\widehat{\mu_{(X_1,\dots,X_N)}}(x_1^*, \dots, x_N^*) = \mathbb{E} \exp(-i \sum_{n=1}^{N} \langle X_n, x_n^* \rangle) = \prod_{n=1}^{N} \mathbb{E} \exp(-i \langle X_n, x_n^* \rangle)$$
$$= \prod_{n=1}^{N} \widehat{\mu_{X_n}}(x_n^*) = \mu_{X_1} \widehat{\times \cdots \times} \mu_{X_N} (x_1^*, \dots, x_N^*).$$

Hence $\mu_{(X_1,\dots,X_N)} = \mu_{X_1} \times \cdots \times \mu_{X_N}$ by Theorem 2.8. $\square$

The joint Gaussianity condition cannot be relaxed to Gaussianity of each of the $X_n$; see Exercise 1.

## 4.3 Series representation

The main result of this section states that every $E$-valued Gaussian random variable can be represented as a Gaussian sum of the form $\sum_{n\geqslant 1} \gamma_n x_n$, where $(\gamma_n)_{n\geqslant 1}$ is a Gaussian sequence and $(x_n)_{n\geqslant 1}$ is a (finite or infinite) sequence in $E$. This fact enables us to extend various results for Gaussian sums, such as the Kahane-Khintchine inequality, to arbitrary Gaussian random variables.

We start with a simple proposition stating that limits of Gaussian variables are Gaussian.

**Proposition 4.11.** *If $(X_n)_{n=1}^\infty$ is a sequence of $E$-valued Gaussian variables and $X$ is a random variable such that*

$$\lim_{n\to\infty} \langle X_n, x^* \rangle = \langle X, x^* \rangle \quad \textit{in probability for all } x^* \in E^*,$$

*then $X$ is Gaussian. Its covariance operator $Q \in \mathscr{L}(E^*, E)$ is given by $\langle Qx^*, y^* \rangle = \lim_{n\to\infty} \langle Q_n x^*, y^* \rangle$ for $x^*, y^* \in E^*$.*

*Proof.* Fixing $x^* \in E^*$, after passing to a subsequence we may assume that $\lim_{n\to\infty} \langle X_n, x^* \rangle = \langle X, x^* \rangle$ almost surely. Then, by the dominated convergence theorem,

$$\mathbb{E} \exp(-i\xi \langle X, x^* \rangle) = \lim_{n\to\infty} \mathbb{E} \exp(-i\xi \langle X_n, x^* \rangle) = \lim_{n\to\infty} \exp(-\tfrac{1}{2}\xi^2 \langle Q_n x^*, x^* \rangle).$$

Since each of the terms $\langle Q_n x^*, x^* \rangle$ is non-negative, this implies that the limit $q(x^*) := \lim_{n\to\infty} \langle Q_n x^*, x^* \rangle$ exists. From the resulting identity

$$\mathbb{E}\exp(-i\xi\langle X, x^*\rangle) = \exp(-\tfrac{1}{2}\xi^2 q(x^*))$$

we conclude that $\langle X, x^*\rangle$ is Gaussian for all $x^* \in E^*$. By definition this means that $X$ is Gaussian.

Denote by $Q$ the covariance operator of $X$. For all $\xi \in \mathbb{R}$,

$$\exp(-\tfrac{1}{2}\xi^2\langle Qx^*, x^*\rangle) = \mathbb{E}\exp(-i\xi\langle X, x^*\rangle) = \exp(-\tfrac{1}{2}\xi^2 q(x^*)).$$

From this we deduce that $\langle Qx^*, x^*\rangle = q(x^*) = \lim_{n\to\infty}\langle Q_n x^*, x^*\rangle$. Applying this to $x^* + y^*$ we find $\langle Qx^*, y^*\rangle = \lim_{n\to\infty}\langle Q_n x^*, y^*\rangle$ for all $x^*, y^* \in E^*$. $\square$

For a Gaussian random variable $X$ with covariance operator $Q$, we denote by $H_X$ the closed linear subspace in $L^2(\Omega)$ spanned by the random variables $\{\langle X, x^*\rangle : \ x^* \in E^*\}$. The operator $i_X : H_X \to E$,

$$i_X\langle X, x^*\rangle := \mathbb{E}\langle X, x^*\rangle X = Qx^*, \tag{4.3}$$

is well-defined and bounded by Hölder's inequality and Fernique's theorem. Its adjoint is given by $i_X^* x^* = \langle X, x^*\rangle$. This leads to the factorisation

$$Q = i_X i_X^*. \tag{4.4}$$

Here, and in similar situations later on, we identify $H_X$ and its dual $H_X^*$ by means of the Riesz representation theorem. Since we are working over the real scalar field this identification is linear and should never lead to any confusion. For a generalisation of the factorisation (4.4) to arbitrary positive symmetric operators $Q$ see Exercise 3.

**Theorem 4.12 (Karhunen-Loève expansion).** *Let $X$ be an $E$-valued Gaussian random variable.*

(1) *The space $H_X$ is separable.*
(2) *If $(\gamma_n)_{n\geqslant 1}$ is an orthonormal basis of $H_X$, then $(\gamma_n)_{n\geqslant 1}$ is a Gaussian sequence and*

$$\sum_{n\geqslant 1}\gamma_n i_X \gamma_n = X,$$

*where convergence holds almost surely and in $L^p(\Omega; E)$ for all $1 \leqslant p < \infty$.*

*Proof.* Define $\widetilde{H_X}$ as the closed linear subspace of $L^2(E, \mu_X)$ spanned by $E^*$; here we think of the functionals $x^* \in E^*$ as functions on $E$. In view of $\mathbb{E}\langle X, x^*\rangle^2 = \int_E \langle x, x^*\rangle^2\, d\mu_X(x)$, the mapping $x^* \mapsto \langle X, x^*\rangle$ extends uniquely to an isometry of Hilbert spaces $\widetilde{H_X} \simeq H_X$.

Let $E_0$ be a separable closed subspace of $E$ containing the essential range of $X$. Then $\mu_X(E_0) = 1$ and therefore the identity mapping gives an isometry $L^2(E_0, \mu_X) \simeq L^2(E, \mu_X)$. Since the Borel $\sigma$-algebra $\mathscr{B}(E_0)$ is generated by a countable family of open sets (take a dense sequence $(x_n)_{n=1}^\infty$ in $E_0$ and consider the open balls $B(x_n, q)$ with rational $q > 0$), the space $L^2(E_0, \mu_X)$ is separable. It follows that $L^2(E, \mu_X)$ is separable and hence so is $\widetilde{H_X}$, it being a closed subspace of $L^2(E, \mu_X)$. It follows that $H_X$ is separable.

Let $(\gamma_n)_{n\geqslant 1}$ be a (finite or countably infinite) orthonormal basis of $H_X$. Every random variable in $H_X$ is Gaussian by Proposition 4.11. In particular, linear combinations of the $\gamma_n$ are Gaussian, which means that all vectors $(\gamma_{n_1}, \dots, \gamma_{n_N})$ are Gaussian as $\mathbb{R}^N$-valued random variables. Therefore Proposition 4.10 implies that the $\gamma_n$ are independent.

For all $x^* \in E^*$ we have the identities

$$\sum_{n\geqslant 1} \gamma_n \langle i_X \gamma_n, x^* \rangle = \sum_{n\geqslant 1} \gamma_n \, \mathbb{E}\gamma_n \langle X, x^* \rangle = \langle X, x^* \rangle$$

in $H_X$, noting that the middle expression is the expansion of $\langle X, x^* \rangle$ with respect to the orthonormal basis $(\gamma_n)_{n\geqslant 1}$ of $H_X$. The result now follows from the Itô-Nisio theorem. $\square$

For the readers familiar with weak$^*$-topologies we sketch an alternative, more functional analytic proof of the separability of $H_X$. The dual $E_0^*$ is weak$^*$-separable, by the separability of $E_0$. Regarding $i_X$ as a bounded injective operator from $H_X$ to $E_0$, the adjoint $i_X^*$ is weak$^*$-continuous and maps $E_0^*$ onto a weak$^*$-separable and weak$^*$-dense subspace of $H_X$. But the weak$^*$-topology of the Hilbert space $H_X$ is the same as the weak topology. By the Hahn-Banach theorem, the weak closure of $i_X^* E_0^*$ equals its strong closure, and the separability of $H_X$ follows.

As an application of Theorem 4.12 we extend Theorem 3.12 to arbitrary $E$-valued Gaussian random variables.

**Corollary 4.13 (Kahane-Khintchine inequality).** *Let $X$ be an $E$-valued Gaussian variable. Then for all $1 \leqslant p, q < \infty$ we have*

$$(\mathbb{E}\|X\|^p)^{\frac{1}{p}} \leqslant K_{p,q} (\mathbb{E}\|X\|^q)^{\frac{1}{q}}.$$

*Proof.* For the special case where $X = \sum_{n=1}^N \gamma_n x_n$ this was proved in Theorem 3.12. The general case follows by combining this with the Karhunen-Loève expansion. $\square$

## 4.4 Convergence

As an application of the Kahane-Khintchine inequality, we show next that if a sequence of Gaussian random variables converges in probability, then it converges in $L^p$ for all $1 \leqslant p < \infty$.

We start with a classical inequality for non-negative random variables.

**Lemma 4.14 (Paley-Zygmund inequality).** *Let $\xi$ be a non-negative random variable. If $0 < \mathbb{E}\xi^2 \leqslant c(\mathbb{E}\xi)^2 < \infty$ for some $c > 0$, then for all $0 < r < 1$ we have*

$$\mathbb{P}\{\xi > r\mathbb{E}\xi\} \geqslant \frac{(1-r)^2}{c}.$$

*Proof.* Using the non-negativity of $\xi$ we have

$$(1-r)\mathbb{E}\xi = \mathbb{E}(\xi - r\mathbb{E}\xi) \leqslant \mathbb{E}(1_{\{\xi > r\mathbb{E}\xi\}}(\xi - r\mathbb{E}\xi)) \leqslant \mathbb{E}(1_{\{\xi > r\mathbb{E}\xi\}}\xi)$$

and therefore, by the Cauchy-Schwarz inequality,

$$(1-r)^2(\mathbb{E}\xi)^2 \leqslant \big(\mathbb{E}(1_{\{\xi > r\mathbb{E}\xi\}}\xi)\big)^2 \leqslant \mathbb{E}1_{\{\xi > r\mathbb{E}\xi\}}\mathbb{E}\xi^2.$$

The result follows upon dividing both sides by $\mathbb{E}\xi^2$. □

**Theorem 4.15.** *For a sequence $(X_n)_{n=1}^\infty$ of $E$-valued Gaussian random variables the following assertions are equivalent:*

(1) *the sequence $(X_n)_{n=1}^\infty$ converges in probability to a random variable $X$;*
(2) *for some $1 \leqslant p < \infty$ the sequence $(X_n)_{n=1}^\infty$ converges in $L^p(\Omega; E)$ to a random variable $X$;*
(3) *for all $1 \leqslant p < \infty$ the sequence $(X_n)_{n=1}^\infty$ converges in $L^p(\Omega; E)$ to a random variable $X$.*

*In this situation the limit random variable $X$ is Gaussian.*

*Proof.* Fix $1 \leqslant p < \infty$. It suffices to prove that $\lim_{n\to\infty} X_n = X$ in probability implies $\lim_{n\to\infty} X_n = X$ in $L^p(\Omega; E)$. Note that $X$ is Gaussian by Proposition 4.11.

*Step 1* - Fix $1 \leqslant q < \infty$. By Fernique's theorem we have $\mathbb{E}\|X_n\|^q < \infty$ for all $n \geqslant 1$. In this step we prove the uniform bound

$$\sup_{n\geqslant 1} \mathbb{E}\|X_n\|^q < \infty. \tag{4.5}$$

From the Paley-Zygmund inequality, for all $n \geqslant 1$ we obtain

$$\mathbb{P}\Big\{\|X_n\|^2 > \frac{1}{2}\mathbb{E}\|X_n\|^2\Big\} \geqslant \frac{1}{4K_{4,2}^4}, \tag{4.6}$$

where $K_{4,2}$ is the Kahane-Khintchine constant corresponding to $p = 4$ and $q = 2$. On the other hand, given $\varepsilon > 0$, for any $r > 0$ we find an index $N \geqslant 1$ such that for all $n \geqslant N$,

$$\begin{aligned}&\mathbb{P}\big\{\|X_n\|^2 > r\big\} \\ &\qquad \leqslant \mathbb{P}\Big\{\|X\| > \frac{1}{2}\sqrt{r}\Big\} + \mathbb{P}\Big\{\|X_n - X\| > \frac{1}{2}\sqrt{r}\Big\} \leqslant \mathbb{P}\Big\{\|X\| > \frac{1}{2}\sqrt{r}\Big\} + \varepsilon.\end{aligned}$$

Thus for large enough $r_0 > 0$ we find an index $N_0 \geqslant 1$ such that for $n \geqslant N_0$,

$$\mathbb{P}\big\{\|X_n\|^2 > r_0\big\} < 2\varepsilon.$$

If for some subsequence we had $\lim_{k\to\infty} \mathbb{E}\|X_{n_k}\|^2 = \infty$, then for all sufficiently large $k$ we would obtain

$$\mathbb{P}\big\{\|X_{n_k}\|^2 > \frac{1}{2}\mathbb{E}\|X_{n_k}\|^2\big\} \leqslant \mathbb{P}\big\{\|X_{n_k}\|^2 > r_0\big\} < 2\varepsilon,$$

contradicting (4.6). We conclude that $\sup_{n\geqslant 1}\mathbb{E}\|X_n\|^2 < \infty$. Now (4.5) follows from the Kahane-Khintchine inequality.

*Step 2* - Fix $1 \leqslant p < q < \infty$. By Step 1, the triangle inequality in $L^q(\Omega; E)$, and a scaling argument we may assume that $\sup_{k\geqslant 1}\mathbb{E}\|X_k - X\|^q \leqslant 1$. Using this together with Hölder's inequality (with $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$), for fixed $\varepsilon > 0$ we obtain

$$\begin{aligned}\mathbb{E}\|X_k - X\|^p &= \mathbb{E}(1_{\{\|X_k - X\|\leqslant\varepsilon\}}\|X_k - X\|^p) + \mathbb{E}(1_{\{\|X_k - X\|>\varepsilon\}}\|X_k - X\|^p)\\ &\leqslant \varepsilon^p + \mathbb{E}(1_{\{\|X_k - X\|>\varepsilon\}}\|X_k - X\|^p)\\ &\leqslant \varepsilon^p + (\mathbb{P}\{\|X_k - X\| > \varepsilon\})^{\frac{p}{r}}.\end{aligned}$$

Since $\lim_{k\to\infty} X_k = X$ in probability, it follows that

$$\limsup_{k\to\infty}\mathbb{E}\|X_k - X\|^p \leqslant \varepsilon^p.$$

This being true for all $\varepsilon > 0$ we arrive at $\limsup_{k\to\infty}\mathbb{E}\|X_k - X\|^p = 0$. □

## 4.5 Exercises

1. This exercise presents an example of two uncorrelated Gaussian random variables which are not independent. This shows that the joint Gaussianity condition in Proposition 4.10 cannot be omitted.
   Let $\gamma$ be a standard Gaussian random variable on a probability space $(\Omega_1, \mathscr{F}_1, \mathbb{P}_1)$ and let $r$ be a Rademacher variable on a probability space $(\Omega_2, \mathscr{F}_2, \mathbb{P}_2)$. Define the random variables $\varphi_1$ and $\varphi_2$ on the product space $(\Omega, \mathscr{F}, \mathbb{P}) = (\Omega_1 \times \Omega_2, \mathscr{F}_1 \times \mathscr{F}_2, \mathbb{P}_1 \times \mathbb{P}_2)$ by
   $$\varphi_1(\omega_1, \omega_2) = \gamma(\omega_1), \quad \varphi_2(\omega_1, \omega_2) = \gamma(\omega_1)r(\omega_2).$$
   a) Show that $\varphi_1$ and $\varphi_2$ are Gaussian.
   b) Show that $\varphi_1$ and $\varphi_2$ are uncorrelated.
   c) Show that $\varphi_1$ and $\varphi_2$ fail to be independent.
   *Hint:* Consider, for instance, the events $\{|\varphi_1| \leqslant 1\}$ and $\{|\varphi_2| \leqslant 1\}$.
2. In this exercise we prove Sazanov*'s theorem*: a bounded linear operator $Q$ on a separable Hilbert space $H$ with inner product $[\cdot,\cdot]$ is a Gaussian covariance operator if and only if $Q$ is positive, self-adjoint and the sum $\sum_{n=1}^{\infty}[Qh_n, h_n]$ converges for some (equivalently, for every) orthonormal basis $(h_n)_{n=1}^{\infty}$ of $H$.
   a) Suppose $Q$ satisfies the conditions of the Sazanov theorem, let $(h_n)_{n=1}^{\infty}$ be an orthonormal basis of $H$, and put $x_n := Q^{\frac{1}{2}}h_n$. Show that the Gaussian sum $\sum_{n=1}^{\infty}\gamma_n x_n$ converges in $L^2(\Omega; H)$ and defines a Gaussian $H$-valued random variable with covariance $Q$.

b) Suppose conversely that $X$ is an $H$-valued Gaussian random variable with covariance operator $Q$. Then $Q$ is positive and symmetric. Show that if $(h_n)_{n=1}^{\infty}$ is any orthonormal basis for $H$, then

$$\sum_{n=1}^{\infty} [Qh_n, h_n] = \mathbb{E}\|X\|^2.$$

c) Deduce that the identity operator on a separable infinite-dimensional Hilbert space fails to be a Gaussian covariance operator.

3. (!) The identity (4.4) shows that every Gaussian covariance operator can be written as $Q = TT^*$ for a suitable operator $T$ from a Hilbert space into $E$. In this exercise we generalise this observation to arbitrary positive symmetric operators.
Let $Q \in \mathscr{L}(E^*, E)$ be positive and symmetric.
a) Show that the formula

$$[Qx^*, Qy^*] := \langle Qx^*, y^* \rangle$$

defines an inner product on the range of $Q$.

The Hilbert space completion of the range of $Q$ with respect to this inner product is denoted by $H_Q$.
b) Show that the identity mapping $Qx^* \mapsto Qx^*$ extends uniquely to a bounded operator $i_Q$ from $H_Q$ into $E$.
c) Prove the identity

$$i_Q i_Q^* = Q.$$

d) Prove the statement concerning $Q$ in the proof of Theorem 3.9.
*Hint:* Consider an orthonormal basis of the (finite-dimensional) Hilbert space $H_Q$.

4. Suppose that $X$ is an $E$-valued Gaussian random variable with covariance operator $Q$. We compare the mappings $i_X : H_X \to E$ defined by (4.3) and $i_Q : H_Q \to E$ of the previous exercise.
a) Show that the mapping $\langle X, x^* \rangle \mapsto i_Q^* x^*$ extends uniquely to an isometry from $H_X$ onto $H_Q$.
b) Prove that $i_X(H_X) = i_Q(H_Q)$ and show that $X$ takes its values in $\overline{i_X(H_X)} = \overline{i_Q(H_Q)}$ almost surely.

5. Let $Q$ be a positive self-adjoint operator on a Hilbert space $H$ and let $\sqrt{Q}$ be its unique positive square root.
a) Show that the range of $\sqrt{Q}$ is a Hilbert space with respect to the norm

$$|||\sqrt{Q}h||| := \inf\{\|h'\| : \ h' \in H, \ \sqrt{Q}h' = \sqrt{Q}h\}.$$

b) Show that the identity mapping $Qh \mapsto Qh$ extends uniquely to an isometry

$$H_Q \simeq \mathrm{range}(\sqrt{Q}),$$

where $H_Q$ is defined as in the previous two exercises.

**Notes.** A comprehensive treatment of the theory of Gaussian variables is given in Bogachev [8]. See also the monographs of Janson [53], Vakhania, Tarieladze, Chobanyan [105], and the older lecture notes of Kuo [62].

Theorem 4.3 is a celebrated result due to Fernique [39]. By a (non-trivial) modification of the proof one obtains the following stronger result: if $\mathscr{X}$ is a uniformly tight family of $E$-valued Gaussian random variables, then there exist constants $\beta > 0$ and $C > 0$ such that

$$\mathbb{E}(\exp(\beta\|X\|^2) \leqslant C \qquad \forall X \in \mathscr{X}.$$

Using powerful concentration of measure inequalities it can be shown that the supremum of all admissible constants $\beta$ for which the conclusion of Fernique's theorem holds is equal to $1/2\sigma^2(X)$, where

$$\sigma^2(X) = \sup_{\|x^*\|\leqslant 1} \mathbb{E}\langle X, x^*\rangle^2.$$

We refer to Kwapień and Woyczyński [65], Ledoux [68], and Ledoux and Talagrand [69] for expositions of this result and further reading.

The proof of Theorem 4.15 is taken from Rosiński and Suchanecki [96].

For more on the Karhunen-Loève expansion of Gaussian variables we recommend [65]. The convergence of the series can be alternatively deduced from the martingale convergence theorem for Banach space-valued martingales, but we have chosen not to do so here in order to keep the presentation self-contained.

A Borel measure $\mu$ on a Banach space $E$ is called *Gaussian* if it is the distribution of an $E$-valued Gaussian random variable $X$, or equivalently, if the image measure $\langle \mu, x^*\rangle$ are Gaussian on $\mathbb{R}$ for all $x^* \in E^*$ (to see that the latter implies the former consider the random variable $X(x) := x$ on the probability space $(E, \mu)$). The *covariance operator* of $\mu$ is then defined as the covariance operator $Q$ of $X$. In view of the identities $\langle Qx^*, x^*\rangle = \mathbb{E}\langle X, x^*\rangle^2 = \int_E \langle x, x^*\rangle^2 \, d\mu(x)$ this is well-defined. For the sake of unity of presentation we have stated all results in terms of random variables. Some results, such as Theorem 4.3 and Propositions 4.6 and 4.9, can equally well be formulated in terms of Gaussian measures.

Exercise 2 c) tells us that on an infinite-dimensional Hilbert space $H$ there is no *standard* Gaussian measure, that is, a Gaussian measure whose covariance operator is the identity operator. More can be said, however. Let us call a subset $C$ of $H$ *cylindrical* if it is of the form

$$C = \{h \in H : \ ([h, h_1], \dots, [h, h_n]) \in B\}$$

for certain $h_1, \dots, h_n \in H$ and a Borel set $B$ in $\mathbb{R}^n$. More generally, cylindrical sets in Banach spaces can be defined by replacing the role of the $h_j$ by functionals $x_j^*$. We have already used cylindrical sets in the proof of the uniqueness theorem for the Fourier transform (Theorem 2.8). The cylindrical sets form an algebra of sets in $H$. It can be shown that there exists a unique finitely additive measure $\gamma_H$ on this algebra with the property that the restrictions of $\gamma_H$ to finite-dimensional subspaces of $H$ are standard Gaussian measures.

The pair $(i_Q, H_Q)$ constructed in Exercise 3 is called the *reproducing kernel* associated with $Q$. The operator $i_Q : H_Q \to E$ is in fact injective, and the factorisation $Q = i_Q i_Q^*$ is minimal in the following sense: if $H$ is a Hilbert space and $T : H \to E$ is a bounded operator such that $Q = TT^*$, then there exists a bounded surjection $P : H \to H_Q$ such that $T = i_Q P$. For more information on reproducing kernel Hilbert spaces as well as an explanation of the terminology we refer the interested reader to SCHWARTZ [98] and the book by VAKHANIA, TARIELADZE, CHOBANYAN [105].

# 5

# $\gamma$-Radonifying operators

Experience has taught that many results in analysis involving $L^2$-techniques, such as the Plancherel theorem in harmonic analysis and the Itô isometry in stochastic analysis, carry over without difficulty to the Hilbert space-valued setting. Often this fact characterises Hilbert spaces among all Banach spaces.

It has only been recently realised that many results do generalise beyond the Hilbert space case if one does three things:

- Replace 'functions' by '$\gamma$-radonifying (integral) operators';
- Replace 'uniform boundedness' by '$\gamma$-boundedness';
- Replace 'orthogonality' by 'unconditionality'.

This paradigm has had enormous impact in the areas of (parabolic) evolution equations and harmonic analysis, more recently, in the theory of stochastic (parabolic) evolution equations.

In this lecture we address the first item in the list and investigate properties of $\gamma$-radonifying operators. These operators will be used in the next lecture to give necessary and sufficient conditions for stochastic integrability, the main idea being that the $L^2$-norms occurring in the Itô isometry are replaced by the $\gamma$-radonifying norms of associated integral operators.

## 5.1 $\gamma$-Summing operators

We begin with a discussion of the class of $\gamma$-summing operators. In the next section, $\gamma$-radonifying operators are defined as the $\gamma$-summing operators which can be approximated in the $\gamma$-summing norm by finite rank operators.

Continuing the notational conventions of the previous lectures, $(\gamma_n)_{n=1}^\infty$ always denotes a Gaussian sequence, $H$ is a Hilbert space (with inner product $[\cdot,\cdot]$), and $E$ is a Banach space. Although we have made the standing assumption that all spaces are real, most results of this lecture extend with only minor changes to complex scalars.

**Definition 5.1.** *A linear operator $S : H \to E$ is called* $\gamma$-summing *if for some (equivalently, for all) $1 \leqslant p < \infty$,*

$$\|S\|_{\gamma_p^\infty(H,E)} := \sup \Big(\mathbb{E}\Big\|\sum_{j=1}^k \gamma_j\, Sh_j\Big\|^p\Big)^{\frac{1}{p}} < \infty,$$

*the supremum being taken over all finite orthonormal systems $\{h_1, \dots, h_k\}$.*

By considering singletons $\{h\}$ we see that every $\gamma$-summing operator is bounded and satisfies $\|S\| \leqslant \|S\|_{\gamma_p^\infty(H,E)}$.

With respect to any one of the norms $S \mapsto \|S\|_{\gamma_p^\infty(H,E)}$, which are mutually equivalent by the Kahane-Khintchine inequalities, the linear space $\gamma^\infty(H,E)$ of all $\gamma$-summing operators from $H$ to $E$ is a normed space. Unless otherwise stated we shall write

$$\|S\|_{\gamma^\infty(H,E)} := \|S\|_{\gamma_2^\infty(H,E)}.$$

**Proposition 5.2.** *The space $\gamma^\infty(H,E)$ is a Banach space.*

*Proof.* If $(S_n)_{n=1}^\infty$ is Cauchy in $\gamma^\infty(H,E)$, then $\sup_{n\geqslant 1} \|S_n\|_{\gamma^\infty(H,E)} < \infty$. Let us denote this supremum by $C$. Since $(S_n)_{n=1}^\infty$ is a Cauchy sequence in $\mathscr{L}(H,E)$ it tends to an operator $S$ in $\mathscr{L}(H,E)$. We will prove that $S \in \gamma^\infty(H,E)$ and that $\lim_{n\to\infty} S_n = S$ in the norm of $\gamma^\infty(H,E)$.

If $\{h_1, \dots, h_k\}$ is an orthonormal system in $H$, then by Fatou's lemma,

$$\mathbb{E}\Big\|\sum_{j=1}^k \gamma_j\, Sh_j\Big\|^2 \leqslant \liminf_{n\to\infty} \mathbb{E}\Big\|\sum_{j=1}^k \gamma_j\, S_n h_j\Big\|^2 \leqslant C.$$

It follows that $S \in \gamma^\infty(H,E)$ and $\|S\|_{\gamma^\infty(H,E)} \leqslant C$.

Next we check that $\lim_{n\to\infty} S_n = S$ in the norm of $\gamma^\infty(H,E)$. Given $\varepsilon > 0$, we choose $N \geqslant 1$ such that $\|S_n - S_m\|_{\gamma^\infty(H,E)} < \varepsilon$ for all $m, n \geqslant N$. Let $\{h_1, \dots, h_k\}$ be an orthonormal system in $H$. By another application of the Fatou lemma,

$$\mathbb{E}\Big\|\sum_{j=1}^k \gamma_j\, (S_n - S)h_j\Big\|^2 \leqslant \liminf_{m\to\infty} \mathbb{E}\Big\|\sum_{j=1}^k \gamma_j\, (S_n - S_m)h_j\Big\|^2 < \varepsilon^2.$$

Therefore, $\|S_n - S\|_{\gamma^\infty(H,E)} \leqslant \varepsilon$ for all $n \geqslant N$. ☐

**Proposition 5.3 ($\gamma$-Fatou lemma).** *Let $(S_n)_{n=1}^\infty$ be a bounded sequence in $\gamma_p^\infty(H,E)$. If $S \in \mathscr{L}(H,E)$ is an operator such that*

$$\lim_{n\to\infty} \langle S_n h, x^*\rangle = \langle Sh, x^*\rangle \quad \forall h \in H,\ x^* \in E^*,$$

*then $S \in \gamma^\infty(H,E)$ and for all $1 \leqslant p < \infty$ we have*

$$\|S\|_{\gamma_p^\infty(H,E)} \leqslant \liminf_{n\to\infty} \|S_n\|_{\gamma_p^\infty(H,E)}.$$

*Proof.* Let $\{h_1, \ldots, h_k\}$ be an orthonormal system in $H$. Let $(x_n^*)_{n=1}^\infty$ be a sequence of unit vectors in $E^*$ which is norming for the linear span of $\{Sh_1, \ldots, Sh_k\}$. For all $M \geqslant 1$ we have, by the Fatou lemma,

$$\begin{aligned}\mathbb{E} \sup_{m=1,\ldots,M} \Big|\Big\langle \sum_{j=1}^k \gamma_j\, Sh_j, x_m^* \Big\rangle\Big|^p &\leqslant \liminf_{n\to\infty} \mathbb{E} \sup_{m=1,\ldots,M} \Big|\Big\langle \sum_{j=1}^k \gamma_j\, S_n h_j, x_m^* \Big\rangle\Big|^p \\ &\leqslant \liminf_{n\to\infty} \|S_n\|^p_{\gamma_p^\infty(H,E)}.\end{aligned}$$

Taking the limit $M \to \infty$ we obtain, by the monotone convergence theorem,

$$\mathbb{E}\Big\| \sum_{j=1}^k \gamma_j\, Sh_j \Big\|^p \leqslant \liminf_{n\to\infty} \|S_n\|^p_{\gamma_p^\infty(H,E)}$$

and the proposition follows. □

The next result shows that the class of $\gamma$-summing operators enjoys a certain ideal property:

**Proposition 5.4 (Ideal property I).** *Let $S \in \gamma^\infty(H,E)$. If $H'$ is another Hilbert space and $E'$ another Banach space, then for all $T \in \mathscr{L}(H',H)$ and $U \in \mathscr{L}(E,E')$ we have $UST \in \gamma^\infty(H',E')$ and for all $1 \leqslant p < \infty$ we have*

$$\|UST\|_{\gamma_p^\infty(H',E')} \leqslant \|U\|\, \|S\|_{\gamma_p^\infty(H,E)} \|T\|.$$

*Proof.* It suffices to prove that $ST \in \gamma^\infty(H',E)$ and $\|ST\|_{\gamma^\infty(H',E)} \leqslant \|S\|_{\gamma^\infty(H,E)} \|T\|$, the assertions concerning $U$ being trivial.

Let $\{h_1', \ldots, h_k'\}$ be an orthonormal system in $H'$. We denote by $\widetilde{H}'$ and $\widetilde{H}$ the spans in $H'$ and $H$ of $\{h_1', \ldots, h_k'\}$ and $\{Th_1', \ldots, Th_k'\}$, respectively. Let $\widetilde{E}$ be the span in $E$ of $\{STh_1', \ldots, STh_k'\}$. Then $S$ and $T$ restrict to operators $\widetilde{S} : \widetilde{H} \to \widetilde{E}$ and $\widetilde{T} : \widetilde{H'} \to \widetilde{H}$.

Let $\{h_1, \ldots, h_N\}$ be an orthonormal basis for $\widetilde{H}$. For all $x^* \in \widetilde{E}^*$ we have

$$\sum_{j=1}^k \langle \widetilde{S}\widetilde{T}h_j', x^*\rangle^2 = \|\widetilde{T}^*\widetilde{S}^*x^*\|^2_{\widetilde{H}'} \leqslant \|\widetilde{T}^*\|^2\, \|\widetilde{S}^*x^*\|^2_{\widetilde{H}} = \|\widetilde{T}\|^2 \sum_{n=1}^N \langle \widetilde{S}h_n, x^*\rangle^2.$$

Hence, by Theorem 3.9,

$$\mathbb{E}\Big\| \sum_{j=1}^k \gamma_j\, STh_j' \Big\|^p \leqslant \|T\|^p\, \mathbb{E}\Big\| \sum_{n=1}^N \gamma_n\, Sh_n \Big\|^p \leqslant \|T\|^p\, \|S\|^p_{\gamma_p^\infty(H,E)}$$

and the result follows. □

As a corollary we observe that we may ignore the kernel of $S$:

**Corollary 5.5.** *If $S \in \gamma^\infty(H,E)$ and $H_0$ is a closed subspace of $H$ containing $(\ker S)^\perp$, then the restriction $S_0$ of $S$ to $H_0$ belongs to $\gamma^\infty(H_0,E)$ and for all $1 \leqslant p < \infty$,*

$$\|S_0\|_{\gamma_p^\infty(H_0,E)} = \|S\|_{\gamma_p^\infty(H,E)}.$$

*Proof.* The only nontrivial thing to prove is the inequality $\|S\|_{\gamma_p^\infty(H,E)} \leqslant \|S_0\|_{\gamma_p^\infty(H_0,E)}$. Let $P_0$ be the orthonormal projection of $H$ onto $H_0$. Then $S = S_0 P_0$ and the desired inequality follows from Proposition 5.4. □

We are now in a position to prove the following characterisation of $\gamma$-summing operators in terms of orthonormal bases. We formulate the result for separable infinite-dimensional spaces; for finite-dimensional spaces the same result holds with a slightly simpler proof.

**Proposition 5.6.** *If $H$ is separable and $(h_n)_{n=1}^\infty$ is an orthonormal basis for $H$, then an operator $S \in \mathscr{L}(H,E)$ belongs to $\gamma^\infty(H,E)$ if and only if for some (equivalently, for all) $1 \leqslant p < \infty$,*

$$\sup_{N \geqslant 1} \mathbb{E}\Big\| \sum_{n=1}^N \gamma_n \, Sh_n \Big\|^p < \infty.$$

*In this case,*

$$\|S\|^p_{\gamma_p^\infty(H,E)} = \sup_{N \geqslant 1} \mathbb{E}\Big\| \sum_{n=1}^N \gamma_n \, Sh_n \Big\|^p.$$

*Proof.* Let $\{h_1', \dots, h_k'\}$ be an orthonormal system in $H$. For $K \geqslant 1$ let $P_K$ denote the orthogonal projection onto the span of $\{h_1, \dots, h_K\}$. For all $x^* \in E^*$ and $K \geqslant k$ we have

$$\sum_{j=1}^k \langle SP_K h_j', x^* \rangle^2 \leqslant \|P_K S^* x^*\|^2 = \sum_{n=1}^K \langle Sh_n, x^* \rangle^2.$$

Let $1 \leqslant p < \infty$. From Theorem 3.9 it follows that

$$\mathbb{E}\Big\| \sum_{j=1}^k \gamma_j \, SP_K h_j' \Big\|^p \leqslant \mathbb{E}\Big\| \sum_{n=1}^K \gamma_n \, Sh_n \Big\|^p \leqslant \sup_{N \geqslant 1} \mathbb{E}\Big\| \sum_{n=1}^N \gamma_n \, Sh_n \Big\|^p.$$

Hence by Fatou's lemma,

$$\mathbb{E}\Big\| \sum_{j=1}^k \gamma_j \, Sh_j' \Big\|^p \leqslant \liminf_{K \to \infty} \mathbb{E}\Big\| \sum_{j=1}^k \gamma_j \, SP_K h_j' \Big\|^p \leqslant \sup_{N \geqslant 1} \mathbb{E}\Big\| \sum_{n=1}^N \gamma_n \, Sh_n \Big\|^p.$$

It follows that

$$\|S\|^p_{\gamma_p^\infty(H,E)} \leqslant \sup_{N \geqslant 1} \mathbb{E}\Big\| \sum_{n=1}^N \gamma_n \, Sh_n \Big\|^p.$$

The converse inequality trivially holds and the proof is complete. □

## 5.2 $\gamma$-Radonifying operators

When $h$ is an element of a Hilbert space $H$ and $x$ an element of a Banach space $E$, we denote by $h \otimes x$ the operator in $\mathscr{L}(H,E)$ defined by

$$(h \otimes x)h' := [h, h']x, \qquad h' \in H.$$

An operator in $\mathscr{L}(H,E)$ is said to be of *finite rank* if it is a linear combination of operators of the above form. It is a trivial observation that every finite rank operator from $H$ to $E$ belongs to $\gamma^\infty(H,E)$. In fact we have:

**Lemma 5.7.** *If $S = \sum_{n=1}^N h_n \otimes x_n$ is a finite rank operator with $h_1, \dots, h_N$ orthonormal in $H$ and $x_1, \dots, x_N \in E$ arbitrary, then $S \in \gamma^\infty(H,E)$ and for all $1 \leqslant p < \infty$ we have*

$$\|S\|^p_{\gamma^\infty_p(H,E)} = \mathbb{E}\Big\| \sum_{n=1}^N \gamma_n \, x_n \Big\|^p.$$

*Proof.* Testing on $h_1, \dots, h_N$ gives the inequality '$\geqslant$'. To prove the inequality '$\leqslant$' let $P$ be the orthogonal projection from $H$ onto the span $\widetilde{H}$ of $\{h_1, \dots, h_N\}$ and define $\widetilde{S} \in \mathscr{L}(\widetilde{H}, E)$ by $\widetilde{S} = SP^*$. The inequality then follows from Proposition 5.4 applied to $S = \widetilde{S}P$, and Proposition 5.6 applied to $\widetilde{S}$. $\square$

In view of this observation the following definition makes sense.

**Definition 5.8.** *The space $\gamma(H,E)$ is defined as the closure in $\gamma^\infty(H,E)$ of all finite rank operators. The operators in $\gamma(H,E)$ are called* $\gamma$-radonifying.

By definition, $\gamma(H,E)$ is a Banach space with respect to the norm inherited from $\gamma^\infty(H,E)$. For notational simplicity, for $R \in \gamma(H,E)$ we shall write

$$\|R\|_{\gamma(H,E)} := \|R\|_{\gamma^\infty(H,E)}$$

and more generally $\|R\|_{\gamma_p(H,E)} := \|R\|_{\gamma_p^\infty(H,E)}$ for $1 \leqslant p < \infty$.

A bounded operator is *compact* if the image of the unit ball is relatively compact. Every $\gamma$-radonifying operator $R$ is compact: if $\lim_{n\to\infty} \|R_n - R\|_{\gamma(H,E)} = 0$ with each $R_n$ of finite rank, then $\lim_{n\to\infty} \|R_n - R\| = 0$ and the claim follows since each $R_n$ is compact. Here we use that the uniform limit of a sequence of compact operators is compact.

Without proof we mention the following theorem, which rephrases a famous result due to Hoffmann-Jorgensen and Kwapień on the almost sure convergence of random sums whose partial sums are almost surely bounded.

**Theorem 5.9 (Hoffmann-Jorgensen and Kwapień).** *Let $H$ be an infinite-dimensional Hilbert space. For a Banach space $E$ the following assertions are equivalent:*

(1) *$\gamma^\infty(H,E) = \gamma(H,E)$;*
(2) *$E$ does not contain a closed subspace isomorphic to $c_0$.*

An explicit example of an operator which is $\gamma$-summing but not $\gamma$-radonifying is the multiplication operator $R : \ell^2 \to c_0$ defined by

$$R\big((\alpha_n)_{n=1}^\infty\big) := (\alpha_n/\sqrt{\log(n+1)})_{n=1}^\infty$$

The proof of this statement depends on some subtle estimates for Gaussian sums and is omitted.

As an immediate consequence of Definition 5.8, every $R \in \gamma(H,E)$ is 'supported' on a separable closed subspace of $H$:

**Proposition 5.10.** *If $R \in \gamma(H,E)$, then $(\ker(R))^\perp$ is separable.*

*Proof.* Suppose that $R = \lim_{n\to\infty} R_n$ in $\gamma(H,E)$ with each $R_n$ of finite rank, say $R_n h = \sum_{j=1}^{k_n} [h, h_{jn}] x_{jn}$. Let $H_0$ denote the closed linear span of all vectors $h_{jn}$, $n \geqslant 1$, $1 \leqslant j \leqslant k_n$. Then $H_0$ is separable and if $h \perp H_0$, then $R_n h = 0$ for all $n \geqslant 1$ and consequently $Rh = 0$. □

The ideal property of $\gamma^\infty(H,E)$ carries over to $\gamma(H,E)$:

**Proposition 5.11 (Ideal property II).** *Let $R \in \gamma(H,E)$. If $H'$ is another Hilbert space and $E'$ another Banach space, then for all $T \in \mathscr{L}(H',H)$ and $U \in \mathscr{L}(E,E')$ we have $URT \in \gamma(H',E')$ and for all $1 \leqslant p < \infty$ we have*

$$\|URT\|_{\gamma_p(H',E')} \leqslant \|U\|\,\|R\|_{\gamma_p(H,E)}\|T\|.$$

*Proof.* If $R$ is of finite rank, then also $URT$ is of finite rank. Moreover if $\lim_{n\to\infty} R_n = R$ in $\gamma_p^\infty(H,E)$, then $\|U(R-R_n)T\|_{\gamma_p^\infty(H',E')} \leqslant \|U\|\,\|R - R_n\|_{\gamma_p^\infty(H,E)}\|T\|$ and therefore $URT \in \gamma(H',E')$. The estimate follows from the corresponding estimate for the $\gamma$-summing norms. □

We mention a simple but useful application.

**Proposition 5.12 (Convergence by right multiplication).** *If $H_1$ and $H_2$ are Hilbert spaces and $S_1, S_2, \dots$ and $S$ are operators in $\mathscr{L}(H_1,H_2)$ satisfying $S^* h = \lim_{n\to\infty} S_n^* h$ for all $h \in H_2$, then for all $R \in \gamma(H_2,E)$ we have $\lim_{n\to\infty} RS_n = RS$ in $\gamma(H_1,E)$.*

*Proof.* The uniform boundedness principle implies that $\sup_{n\geqslant 1}\|S_n\| < \infty$. Hence, by the estimate $\|RT\|_{\gamma(H_1,E)} \leqslant \|R\|_{\gamma(H_2,E)}\|T\|$ for $T \in \mathscr{L}(H_1,H_2)$, it suffices to consider finite rank operators $R \in \gamma(H_2,E)$. Fix such an operator, say $R = \sum_{m=1}^M h_m' \otimes x_m$, and let $(h_j)_{j=1}^k$ be orthonormal in $H_1$. Then, by the triangle inequality in $L^2(\Omega;E)$,

$$
\begin{aligned}
\Big(\mathbb{E}\Big\|\sum_{j=1}^{k}\gamma_j R(S-S_n)h_j\Big\|^2\Big)^{\frac12} &= \Big(\mathbb{E}\Big\|\sum_{m=1}^{M}\sum_{j=1}^{k}\gamma_j[S^*h'_m - S_n^*h'_m, h_j]x_m\Big\|^2\Big)^{\frac12} \\
&\leqslant \sum_{m=1}^{M}\Big(\mathbb{E}\Big|\sum_{j=1}^{k}\gamma_j[S^*h'_m - S_n^*h'_m, h_j]\Big|^2\Big)^{\frac12}\|x_m\| \\
&= \sum_{m=1}^{M}\Big(\sum_{j=1}^{k}|[S^*h'_m - S_n^*h'_m, h_j]|^2\Big)^{\frac12}\|x_m\| \\
&\leqslant \sum_{m=1}^{M}\|S^*h'_m - S_n^*h'_m\|\|x_m\|.
\end{aligned}
$$

Hence,

$$\|R(S-S_n)\|_{\gamma(H_1,E)} \leqslant \sum_{m=1}^{M}\|S^*h'_m - S_n^*h'_m\|\|x_m\|,$$

and by assumption the right hand side tends to zero as $n\to\infty$. □

Here is a simple illustration:

*Example 5.13.* Consider an operator $R\in\gamma(H,E)$ and let $(h_n)_{n=1}^\infty$ be an orthonormal basis for $(\ker(R))^\perp$. Let $P_n$ denote the orthogonal projection in $H$ onto the span of $\{h_1,\dots,h_n\}$. Then $\lim_{n\to\infty} RP_n = R$ in $\gamma(H,E)$.

**Proposition 5.14 (Measurability).** *Let $(A,\mathscr{A},\mu)$ be a $\sigma$-finite measure space and $H$ a separable Hilbert space. For a function $\Phi : A\to\gamma(H,E)$ define $\Phi h : A\to E$ by $(\Phi h)(\xi) := \Phi(\xi)h$ for $h\in H$. The following assertions are equivalent:*

(1) *$\Phi$ is strongly $\mu$-measurable;*
(2) *$\Phi h$ is strongly $\mu$-measurable for all $h\in H$.*

*Proof.* It suffices to prove that (2) implies (1). If $(h_n)_{n=1}^\infty$ is an orthonormal basis for $H$, then with the notations of the Example 5.13 for all $\xi\in A$ we have

$$\Phi(\xi) = \lim_{n\to\infty}\Phi(\xi)P_n = \lim_{n\to\infty}\sum_{j=1}^{n}[\,\cdot\,,h_j]\Phi(\xi)h_j,$$

with convergence in the norm of $\gamma(H,E)$. □

We proceed with the main result of this section which states, loosely speaking, that an operator is $\gamma$-radonifying if and only if it maps orthonormal sequences into $\gamma$-summable sequences.

**Theorem 5.15.** *If $H$ is separable, then for an operator $R\in\mathscr{L}(H,E)$ the following assertions are equivalent:*

(1) *$R\in\gamma(H,E)$;*

(2) *for all orthonormal bases* $(h_n)_{n=1}^\infty$ *in* $H$ *and all* $1 \leqslant p < \infty$ *the sum* $\sum_{n=1}^\infty \gamma_n \, Rh_n$ *converges in* $L^p(\Omega; E)$*;*
(3) *for some orthonormal basis* $(h_n)_{n=1}^\infty$ *in* $H$ *and some* $1 \leqslant p < \infty$ *the sum* $\sum_{n=1}^\infty \gamma_n \, Rh_n$ *converges in* $L^p(\Omega; E)$*.*

*In this situation, the sums in* (2) *and* (3) *converge almost surely and define an* $E$*-valued Gaussian random variable with covariance operator* $RR^*$*. For all orthonormal bases* $(h_n)_{n=1}^\infty$ *of* $H$ *and* $1 \leqslant p < \infty$ *we have*

$$\|R\|_{\gamma_p(H,E)}^p = \mathbb{E}\Big\| \sum_{n=1}^\infty \gamma_n \, Rh_n \Big\|^p.$$

*Proof.* (1)⇒(2): Fix $R \in \gamma(H,E)$ and $1 \leqslant p < \infty$, and let $(h_n)_{n=1}^\infty$ be an orthonormal basis of $H$. Let $P_n$ denote the orthogonal projection in $H$ onto the linear span of $\{h_1, \dots, h_n\}$. By Proposition 5.12 we have $\lim_{n\to\infty} RP_n = R$ in $\gamma(H,E)$, and by Proposition 5.6 for all $m < n$ we have

$$\mathbb{E}\Big\| \sum_{j=m+1}^n \gamma_j Rh_j \Big\|^p = \|RP_n - RP_m\|_{\gamma_p(H,E)}^p.$$

Since the right-hand side tends to 0 as $m,n \to \infty$, this proves the convergence of the sum $\sum_{n=1}^\infty \gamma_n Rh_n$ in $L^p(\Omega; E)$.

(2)⇒(3): This implication is trivial.

(3)⇒(1): With the notations as before, by Proposition 5.6 we have

$$\lim_{m,n\to\infty} \|RP_n - RP_m\|_{\gamma_p(H,E)}^p = \lim_{m,n\to\infty} \mathbb{E}\Big\| \sum_{j=m+1}^n \gamma_j Rh_j \Big\|^p = 0.$$

It follows that $(RP_n)_{n=1}^\infty$ is a Cauchy sequence in $\gamma(H,E)$. Its limit equals $R$, since $\lim_{n\to\infty} RP_n h = Rh$ for all $h \in H$.

This proves the equivalence of (1), (2), (3) as well as the final identity. The almost sure convergence in (2) and (3) follows from the Itô-Nisio theorem. □

We are now ready to characterise Gaussian covariance operators in terms of $\gamma$-radonifying operators.

**Theorem 5.16.** *Suppose* $Q \in \mathscr{L}(E^*, E)$ *and* $R \in \mathscr{L}(H,E)$ *satisfy* $Q = RR^*$*. The following assertions are equivalent:*

(1) $Q$ *is a Gaussian covariance operator;*
(2) $R \in \gamma(H,E)$*.*

*If* $X$ *is an* $E$*-valued random variable with covariance operator* $Q$*, then*

$$\mathbb{E}\|X\|^p = \|R\|_{\gamma_p(H,E)}^p, \quad 1 \leqslant p < \infty.$$

*Proof.* (1)$\Rightarrow$(2): Let $X$ be $E$-valued Gaussian with covariance $Q$. By Theorem 4.12 the Hilbert space $H_X$ is separable, and from the identities $\mathbb{E}\langle X, x^*\rangle^2 = \langle Qx^*, x^*\rangle = \|R^*x^*\|^2$ it follows that the mapping $j_X : \langle X, x^*\rangle \mapsto R^*x^*$ extends uniquely to an isometry from $H_X$ onto $\widetilde{H} := \overline{\mathrm{ran}(R^*)}$.

Let $(\gamma_n)_{n=1}^\infty$ be an orthonormal basis of $H_X$ and put $h_n := j_X\gamma_n$. Then $(h_n)_{n=1}^\infty$ is an orthonormal basis of $\widetilde{H}$. By the Karhunen-Loève theorem (Theorem 4.12) we have $X = \sum_{n=1}^\infty \gamma_n i_X\gamma_n$, where $i_X : H_X \to E$ is given by

$$i_X\langle X, x^*\rangle = Qx^* = RR^*x^* = Rj_X\langle X, x^*\rangle.$$

It follows that $i_X = Rj_X$, and therefore

$$\sum_{n=1}^\infty \gamma_n Rh_n = \sum_{n=1}^\infty \gamma_n Rj_X\gamma_n = \sum_{n=1}^\infty \gamma_n i_X\gamma_n = X. \tag{5.1}$$

Let $\widetilde{R}$ denote the restriction of $R$ to $\widetilde{H}$. By the implication (3)$\Rightarrow$(1) of Theorem 5.15, we have proved that $\widetilde{R} \in \gamma(\widetilde{H}, E)$. Since $R = 0$ on $\widetilde{H}^\perp = \ker(R)$, we have $R = \widetilde{R}P$ where $P$ is the orthogonal projection from $H$ onto $\widetilde{H}$. From Proposition 5.11 we infer that $R \in \gamma(H, E)$.

(2)$\Rightarrow$(1): Using Proposition 5.10, let $(h_n)_{n=1}^\infty$ be an orthonormal basis of the separable Hilbert space $\widetilde{H} = (\ker(R))^\perp$. The $E$-valued random variable $X := \sum_{n=1}^\infty \gamma_n Rh_n$ is Gaussian and has covariance operator $RR^* = Q$.

The final identity follows from (5.1) and Theorem 5.15. $\square$

We continue with a domination result for $\gamma$-radonifying operators.

**Theorem 5.17 (Domination).** *Let $H_1$ and $H_2$ be Hilbert spaces and let $R_1 \in \mathscr{L}(H_1, E)$ and $R_2 \in \mathscr{L}(H_2, E)$. If*

$$\|R_1^*x^*\| \leqslant \|R_2^*x^*\| \qquad \forall x^* \in E^*,$$

*then $R_2 \in \gamma(H_2, E)$ implies $R_1 \in \gamma(H_1, E)$ and for all $1 \leqslant p < \infty$ we have*

$$\|R_1\|_{\gamma_p(H_1,E)} \leqslant \|R_2\|_{\gamma_p(H_2,E)}.$$

*Proof.* Put $\widetilde{H}_1 = \overline{\mathrm{ran}(R_1^*)}$ and $\widetilde{H}_2 = \overline{\mathrm{ran}(R_2^*)}$. By assumption, the mapping $j : R_2^*x^* \mapsto R_1^*x^*$ extends to a contraction from $\widetilde{H}_2$ to $\widetilde{H}_1$. For all $h_1 \in \widetilde{H}_1$ and $x^* \in E^*$ we have $\langle R_2 j^*h_1, x^*\rangle = [h_1, jR_2^*x^*] = [h_1, R_1^*x^*] = \langle R_1h_1, x^*\rangle$. Hence $R_2j^*P = R_1$, where $P$ is the orthogonal projection of $H_1$ onto $\widetilde{H}_1$, and the result follows from Proposition 5.11. $\square$

**Corollary 5.18 (Covariance domination).** *Let $X_1$ and $X_2$ be $E$-valued Gaussian random variables satisfying*

$$\mathbb{E}\langle X_1, x^*\rangle^2 \leqslant \mathbb{E}\langle X_2, x^*\rangle^2 \qquad \forall x^* \in E^*.$$

*Then, for all $1 \leqslant p < \infty$,*

$$\mathbb{E}\|X_1\|^p \leqslant \mathbb{E}\|X_2\|^p.$$

*Proof.* Combine Theorems 5.16 and 5.17. $\square$

## 5.3 Examples of $\gamma$-radonifying operators

For certain range spaces, a complete characterisation of $\gamma$-radonifying operators can be given in non-probabilistic terms. The simplest example occurs when the range space is a Hilbert space.

**Theorem 5.19 (Operators into Hilbert spaces).** *If $E$ is a Hilbert space, then $R \in \gamma(H, E)$ if and only if $R \in \mathscr{L}_2(H, E)$, and in this case we have*

$$\|R\|_{\gamma(H,E)} = \|R\|_{\mathscr{L}_2(H,E)}.$$

Here, $\mathscr{L}_2(H, E)$ denotes the space of all *Hilbert-Schmidt operators* from $H$ to $E$, that is, completion of the space of all finite rank operators $R \in \mathscr{L}(H, E)$ with respect to the norm

$$\|R\|^2_{\mathscr{L}_2(H,E)} := \sum_{n=1}^{N} \|x_n\|^2,$$

where $R = \sum_{n=1}^{N} h_n \otimes x$ with the $h_1, \dots, h_N$ orthonomal in $H$.

*Proof.* This is trivial, since for $R = \sum_{n=1}^{N} h_n \otimes x$ with $h_1, \dots, h_N$ orthonomal in $H$ we have

$$\|R\|^2_{\gamma(H,E)} = \mathbb{E}\Big\| \sum_{n=1}^{N} \gamma_n x_n \Big\|^2 = \sum_{n=1}^{N} \|x_n\|^2 = \|R\|^2_{\mathscr{L}_2(H,E)}.$$

$\square$

In what follows we shall use the notation $A \eqsim_p B$ to express the fact that there exist constants $0 < c \leqslant C < \infty$, depending only on $p$, such that $cA \leqslant B \leqslant CA$. The notation $A \lesssim_p B$ has a similar meaning.

The next result shows that an operator from a separable Hilbert space into an $L^p$-space is $\gamma$-radonifying if and only if it satisfies a *square function estimate.*

**Theorem 5.20 (Operators into $L^p$-spaces).** *Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space, let $H$ be a separable Hilbert space, and let $1 \leqslant p < \infty$. For an operator $R \in \mathscr{L}(H, L^p(A))$ the following assertions are equivalent:*

(1) *$R \in \gamma(H, L^p(A))$;*
(2) *For all orthonormal bases $(h_n)_{n=1}^{\infty}$ of $H$ the function $\big(\sum_{n=1}^{\infty} |Rh_n|^2\big)^{\frac{1}{2}}$ belongs to $L^p(A)$;*
(3) *For some orthonormal basis $(h_n)_{n=1}^{\infty}$ of $H$ the function $\big(\sum_{n=1}^{\infty} |Rh_n|^2\big)^{\frac{1}{2}}$ belongs to $L^p(A)$.*

*In this case we have $\|R\|_{\gamma(H,L^p(A))} \eqsim_p \big\|\big(\sum_{n=1}^{\infty} |Rh_n|^2\big)^{\frac{1}{2}}\big\|$.*

*Proof.* Applying the identity $\sum_{n=M}^N |c_n|^2 = \mathbb{E}|\sum_{n=M}^N c_n\gamma_n|^2$ with $c_n = f_n(\xi)$, $\xi \in A$, then applying the scalar Kahane-Khintchine inequality, then Fubini's theorem, and finally the Kahane-Khintchine inequality in $L^p(A)$, for all $M \leqslant N$ and $f_M, \dots, f_N \in L^p(A)$ we obtain

$$\begin{aligned}\Big\|\Big(\sum_{n=M}^N |f_n|^2\Big)^{\frac12}\Big\|_p &= \Big\|\Big(\mathbb{E}\Big|\sum_{n=M}^N \gamma_n f_n\Big|^2\Big)^{\frac12}\Big\|_p \eqsim_p \Big\|\Big(\mathbb{E}\Big|\sum_{n=M}^N \gamma_n f_n\Big|^p\Big)^{\frac1p}\Big\|_p \\ &= \Big(\mathbb{E}\Big\|\sum_{n=M}^N \gamma_n f_n\Big\|_p^p\Big)^{\frac1p} \eqsim_p \Big(\mathbb{E}\Big\|\sum_{n=M}^N \gamma_n f_n\Big\|_p^2\Big)^{\frac12}.\end{aligned}$$

The equivalences (1)$\Leftrightarrow$(2), (1)$\Leftrightarrow$(3), and the final two-sided estimate now follow by taking $f_n := Rh_n$, where $(h_n)_{n=1}^\infty$ is an orthonormal basis of $H$. $\square$

Here is a neat application:

**Corollary 5.21.** *Let $(A, \mathscr{A}, \mu)$ be a finite measure space and $H$ a separable Hilbert space. For all $T \in \mathscr{L}(H, L^\infty(A))$ and $1 \leqslant p < \infty$ we have $T \in \gamma(H, L^p(A))$ and*

$$\|T\|_{\gamma(H,L^p(A))} \lesssim_p \|T\|_{\mathscr{L}(H,L^\infty(A))}.$$

*Proof.* Let $(h_n)_{n=1}^\infty$ be an orthonormal basis of $H$. Fixing $N \geqslant 1$ and $c \in \mathbb{R}^N$, for $\mu$-almost all $\xi \in A$ we have

$$\begin{aligned}\Big|\sum_{n=1}^N c_n(Th_n)(\xi)\Big| &\leqslant \Big\|\sum_{n=1}^N c_n Th_n\Big\|_\infty \\ &\leqslant \|T\|_{\mathscr{L}(H,L^\infty(A))}\Big\|\sum_{n=1}^N c_n h_n\Big\|_H = \|T\|_{\mathscr{L}(H,L^\infty(A))}\|c\|.\end{aligned}$$

Taking the supremum over a countable dense set in the unit ball of $\mathbb{R}^N$ we obtain the following estimate, valid for $\mu$-almost all $\xi \in A$:

$$\Big(\sum_{n=1}^N |(Th_n)(\xi)|^2\Big)^{\frac12} \leqslant \|T\|_{\mathscr{L}(H,L^\infty(A))}.$$

Now apply Theorem 5.20. $\square$

Every $f \in L^p(A; H)$ defines a bounded operator $R_f \in \mathscr{L}(H, L^p(A))$ by putting

$$(R_f h)(\xi) := [f(\xi), h], \qquad \xi \in A,\ h \in H.$$

The next result shows that $R_f \in \gamma(H, L^p(A))$, and that every $R \in \gamma(H, L^p(A))$ is of this form; this gives an alternative description of $\gamma(H, L^p(A))$. For later use it will be useful to formulate this result in a more slightly more general form. The isomorphism $\gamma(H, L^p(A)) \simeq L^p(A; H)$ is obtained in the special case $E = \mathbb{R}$ in the next theorem.

**Theorem 5.22 ($\gamma$-Fubini isomorphism).** *Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space, let $H$ be a Hilbert space, and let $1 \leqslant p < \infty$. The mapping $U : L^p(A; \gamma(H,E)) \to \mathscr{L}(H, L^p(A;E))$ defined by*

$$((Uf)h)(\xi) := f(\xi)h, \qquad \xi \in A,\ h \in H,$$

*defines an isometry $U$ from $L^p(A;\gamma_p(H,E))$ onto $\gamma_p(H, L^p(A;E))$.*

*Proof.* Let $f \in L^p(A;\gamma_p(H,E))$ be a simple function of the form $f = \sum_{m=1}^M 1_{A_m} \otimes U_m$, where the operators $U_m$ are of the form $\sum_{n=1}^N h_n \otimes x_{mn}$ for some orthonormal system $\{h_1, \dots, h_N\}$ in $H$. Let $\widetilde{H}$ be the span of $\{h_1, \dots, h_N\}$. Using Corollary 5.5, Lemma 5.7, and Fubini's theorem we obtain

$$\begin{aligned}
\|Uf\|_{\gamma_p(H,L^p(A;E))} &= \|Uf\|_{\gamma_p(\widetilde{H},L^p(A;E))} \\
&= \Big(\mathbb{E}\Big\|\sum_{n=1}^N \gamma_n\,(Uf)h_n\Big\|^p_{L^p(A;E)}\Big)^{\frac1p} = \Big(\int_A \mathbb{E}\Big\|\sum_{n=1}^N \gamma_n\, fh_n\Big\|^p\,d\mu\Big)^{\frac1p} \\
&= \Big(\int_A \|f\|^p_{\gamma_p(\widetilde{H},E)}\,d\mu\Big)^{\frac1p} = \Big(\int_A \|f\|^p_{\gamma_p(H,E)}\,d\mu\Big)^{\frac1p} \\
&= \|f\|_{L^p(A;\gamma_p(H,E))}.
\end{aligned}$$

Since the simple functions $f$ of the above form are dense, these estimates imply that $U$ extends to an isomorphism of $L^p(A;\gamma_p(H,E))$ onto a closed subspace of $\gamma_p(H, L^p(A;E))$. To show that this operator is surjective it is enough to show that its range is dense. But

$$U\Big(\sum_{n=1}^N 1_{A_n} \otimes \Big(\sum_{k=1}^K h_k \otimes x_{kn}\Big)\Big) = \sum_{k=1}^K h_k \otimes \Big(\sum_{n=1}^N 1_{A_n} \otimes x_{kn}\Big),$$

for all $A_n \in \mathscr{A}$ with $\mu(A_n) < \infty$, orthonormal $h_1, \dots, h_K \in H$, and arbitrary $x_{kn} \in E$. The elements on the right hand side are dense in $\gamma_p(H, L^p(A;E))$. $\square$

The final example is important in the theory of Brownian motion.

**Theorem 5.23 (Indefinite integration).** *The operator $I_T : L^2(0,T) \to C[0,T]$ defined by*

$$(I_T f)(t) := \int_0^t f(s)\,ds, \qquad f \in L^2(0,T),\ t \in [0,T],$$

*is $\gamma$-radonifying.*

A proof is outlined in Exercise 5.

## 5.4 Exercises

1. Let $1 \leqslant p < \infty$. Determine for which scalar sequences $a = (a_n)_{n=1}^{\infty}$ the diagonal operator $u_n \mapsto a_n u_n$ defines a $\gamma$-radonifying operator from $\ell^2$ to $\ell^p$. Here $u_n = (0, \dots, 0, 1, 0, \dots)$, with the '1' in the $n$-th entry, is the $n$-th unit vector of $\ell^p$.

   *Hint:* Apply Theorem 5.20.

2. Let $(h_n)_{n=1}^{\infty}$ be a *Hilbert sequence* in a Hilbert space $H$, that is, there exists a constant $C \geqslant 0$ such that for all scalars $\alpha_1, \dots, \alpha_N$,

$$\Big\| \sum_{n=1}^{N} \alpha_n h_n \Big\| \leqslant C \Big( \sum_{n=1}^{N} |\alpha_n|^2 \Big)^{\frac{1}{2}}.$$

   Show that if $R \in \gamma(H, E)$, then $\sum_{n=1}^{\infty} \gamma_n R h_n$ converges in $L^2(\Omega; E)$ and

$$\mathbb{E} \Big\| \sum_{n=1}^{\infty} \gamma_n R h_n \Big\|^2 \leqslant C^2 \|R\|_{\gamma(H,E)}^2.$$

3. (!) Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space and define $\Phi : A \to \gamma(H, E)$ by $\Phi := \phi \otimes U$, where $\phi \in L^2(A)$ and $U \in \gamma(H, E)$. Prove that the operator $R_\Phi : L^2(A; H) \to E$,

$$R_\Phi f := \int_A \Phi(\xi) f(\xi) \, d\mu(\xi) = \int_A \phi(\xi) U f(\xi) \, d\mu(\xi)$$

   belongs to $\gamma(L^2(A; H), E)$ with norm

$$\|R_\Phi\|_{\gamma(L^2(A;H),E)} = \|\phi\|_{L^2(A)} \|U\|_{\gamma(H,E)}.$$

4. (!) Let again $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space. For $\mu$-simple functions $\phi : A \to \gamma(H, E)$ we define $R_\phi : L^2(A; H) \to E$ by

$$R_\phi f := \int_A \phi(\xi) f(\xi) \, d\mu(\xi).$$

   By the previous exercise, $R_\phi \in \gamma(L^2(A; H), E)$.
   a) Prove that if $E$ has type 2, then the mapping $\phi \mapsto R_\phi$ has a unique extension to a continuous embedding $L^2(A; \gamma(H, E)) \hookrightarrow \gamma(L^2(A; H), E)$.
   *Hint:* Consider simple functions whose values are finite rank operators.
   b) Prove the following converse for $H = \mathbb{R}$ and $A = (0, 1)$: if the mapping $\phi \mapsto R_\phi$, defined for simple functions $\phi : (0, 1) \to E$, extends to a bounded operator $R$ from $L^2(0, 1; E)$ to $\gamma(L^2(0, 1), E)$, then $E$ has type 2.
   *Hint:* Consider step functions.

Examples of Banach space with type 2 are Hilbert spaces and $L^p$-spaces for $2 \leqslant p < \infty$ (see Exercise 3.4).

*Remark:* The following 'dual' result also holds, with a similar proof: if $E$ has cotype 2, then the mapping $R_\Phi \mapsto \Phi$ is well defined and has a unique extension to a continuous embedding $\gamma(L^2(A;H),E) \hookrightarrow L^2(A;\gamma(H,E))$. Conversely, if the mapping $R_\phi \mapsto \phi$ extends to a continuous embedding $\gamma(L^2(0,1),E) \hookrightarrow L^2(0,1;E)$, then $E$ has cotype 2.

5. We present a proof of Theorem 5.23 due to CIESIELSKI. Another proof will be outlined in the next lecture.
Without loss of generality we take $T = 1$ and set $I_T = I_1 =: I$.
a) Let $\gamma$ be a standard Gaussian variable. Prove that

$$\mathbb{P}\{|\gamma| > t\} \leqslant \frac{2}{t\sqrt{2\pi}} e^{-\frac{1}{2}t^2}.$$

b) Let $(\gamma_n)_{n=1}^\infty$ be a Gaussian sequence. Use a) and the Borel-Cantelli lemma to prove that for any $\alpha > 1$, almost surely we have

$$|\gamma_n| \leqslant \sqrt{2\alpha \log(n+1)}$$

for all but at most finitely many $n \geqslant 1$.

The *Haar basis* of $L^2(0,1)$ is defined by $h_1 \equiv 1$ and $h_n := \phi_{jk}$ for $n \geqslant 2$, where $n = 2^j + k$ with $j = 0,1,2,\dots$ and $k = 1,\dots,2^j$, and

$$\phi_{jk} = 2^{j/2} 1_{\left(\frac{k-1}{2^j}, \frac{k-1/2}{2^j}\right)} - 2^{j/2} 1_{\left(\frac{k-1/2}{2^j}, \frac{k}{2^j}\right)}.$$

c) Prove that $(h_n)_{n=1}^\infty$ is an orthonormal basis for $L^2(0,1)$.
d) Prove that, almost surely, the sum $\sum_{n=1}^\infty \gamma_n (Ih_n)(t)$ converges absolutely and uniformly with respect to $t \in [0,1]$.
*Hint:* Use b) together with the observation that for all $j \geqslant 0$ and $t \in [0,1]$, we have $I\phi_{jk}(t) = 0$ for all but at most one $k \in \{1,\dots,2^j\}$ and that for this $k$ we have $0 \leqslant I\phi_{jk}(t) \leqslant 2^{-j/2-1}$.
e) Combine d) with Theorem 5.15 and the final assertion of the Itô-Nisio theorem to deduce that $I$ is $\gamma$-radonifying from $L^2(0,1)$ to $C[0,1]$.

**Notes.** The class of $\gamma$-summing operators was introduced by LINDE and PIETSCH [70]. A detailed study of $\gamma$-summing operators is presented in DIESTEL, JARCHOW, TONGE [35, Chapter 12]. The notion of a $\gamma$-radonifying operator is older and has its origins in the work of GROSS [43]. Frequently $H$ is assumed to be separable and the equivalent conditions (2) and (3) of Theorem 5.15 are taken as the definition of a $\gamma$-radonifying operator.

To explain the name '$\gamma$-radonifying', let us first introduce some terminology. A probability measure $\mu$ on a topological space $E$ is called a *Radon measure* if for all Borel sets $B \subseteq E$ and $\varepsilon > 0$ there exists a compact subset $K \subseteq B$ such that $\mu(B \setminus K) < \varepsilon$. If $\mu$ is the distribution of a random variable with values in a Banach space $E$, then $\mu$ is a Radon measure on $E$; this can be deduced from Proposition 2.3 and some additional thought. Now Theorem 5.16 can be interpreted as saying that *a bounded operator* $T : H \to E$ *is* $\gamma$*-radonifying if and only if it maps the finitely additive standard Gaussian measure* $\gamma_H$ (see the discussion in the Notes of Lecture 4) *to a Radon measure* $\mu$ *on* $E$ (viz., the Gaussian measure $\mu$ with covariance operator $TT^*$).

In some sense, the class of $\gamma$-radonifying operators is the Gaussian analogue of the class of $p$-absolutely summing operators, a fact with indicates its importance from the point of view of Banach space theory. The intermediate notion of $p$-radonifying operators has been studied thoroughly by the French school. We refer to Vakhania, Tarieladze, Chobanyan [105] for more information and references to the literature.

The $\gamma$-Fatou lemma is essentially due to Kalton and Weis [58]. The authors used $\gamma$-radonifying norms to extend certain results in spectral theory involving square functions to the Banach space-valued setting. Propositions 5.12 and 5.14 are taken from [82].

Corollary 5.18 can be improved as follows: if $X$ and $Y$ are $E$-valued Gaussian random variables satisfying $\mathbb{E}\langle X, x^*\rangle^2 \leqslant \mathbb{E}\langle Y, x^*\rangle^2$ for all $x^* \in E^*$ and $C \subseteq E$ is closed, convex, and symmetric, then

$$\mathbb{P}\{X \notin C\} \leqslant \mathbb{P}\{Y \notin C\}. \tag{5.2}$$

This result is due to Anderson [2].

Without proof we mention the following result, essentially due to Neidhardt [85], which can be proved using Prokhorov's theorem (Theorem 2.19) and a Anderson's inequality (see the Notes of Lecture 4):

**Theorem 5.24 ($\gamma$-Dominated convergence).** *Suppose* $(T_n)_{n=1}^\infty$ *is a sequence in* $\mathscr{L}(H,E)$ *and assume that there exist* $R \in \gamma(H,E)$ *and* $T \in \mathscr{L}(H,E)$ *such that for all* $x^* \in E^*$ *we have:*

(1) $\|T_n^* x^*\| \leqslant \|R^* x^*\|$,
(2) $\lim_{n\to\infty} T_n^* x^* = T^* x^*$ *in* $H$.

*Then* $T \in \gamma(H,E)$ *and* $\lim_{n\to\infty} T_n = T$ *in the norm of* $\gamma(H,E)$.

The main idea is as follows. If $\mathscr{X}$ is a family $E$-valued Gaussian random variables whose covariances are dominated by $R$ in the sense of (1), then by using Anderson's inequality (5.2) it can be shown that $\mathscr{X}$ is uniformly tight, and Prokhorov's theorem can be applied.

The square function characterisation of $\gamma$-radonifying operators into $L^p$-spaces of Theorem 5.20 is taken from [83]. For $p = 2$, Corollary 5.21 asserts that if $(A, \mathscr{A}, \mu)$ is a finite measure space, then every bounded operator from $H$ to $L^2(A)$ which factors through $L^\infty(A)$ is Hilbert-Schmidt. In its present form, the corollary was suggested to us by Haase. A related result is contained in [83]; see also [106, Lemma 8.7.2]. The $\gamma$-Fubini isomorphism is taken from [82].

Exercise 2 is from [44]. Exercise 4 goes back to HOFFMANN-JORGENSEN and PISIER [49] and ROSIŃSKI and SUCHANECKI [96]. In its present form it was noted in [84]. From KWAPIEŃ's theorem (see the notes of Lecture 3 and Exercise 4) we deduce that the mapping $\phi \mapsto R_\phi$ induces an isomorphism of Banach spaces

$$L^2(0,1;E) \simeq \gamma(L^2(0,1),E)$$

if and only if $E$ is isomorphic to a Hilbert space.

The proof of Theorem 5.23 sketched in Exercise 5 is due to CIESIELSKI. He used the uniform convergence of the sum $\sum_{n=1}^\infty \gamma_n I_T h_n$ to give an elementary proof that a Brownian motion admits a version with continuous trajectories; we return to this point in the lext lecture. According to Theorem 5.16, the operator $I_T I_T^*$ is the covariance of a Gaussian measure $w$ on $C[0,T]$, the so-called *Wiener measure*. A straightforward computation shows that

$$\langle I_T I_T^* \delta_s, \delta_t \rangle = \int_{C[0,T]} f(s)f(t)\,dw(f) = \min\{s,t\}, \quad s,t \in [0,T].$$

Here $\delta_s$ and $\delta_t$ are the Dirac measures concentrated at $s$ and $t$. We refer to the textbook of STEELE [99] for a discussion of CIESIELKI's result as well as some of its ramifications.

# 6

# Stochastic integration I: the Wiener integral

The hard work in the previous lectures will pay off in this lecture, which is devoted to stochastic integration. In view of future applications to stochastic Cauchy problems we shall consider a setting where the integrands take values in the space of operators $\mathscr{L}(H,E)$, where $H$ is a Hilbert space and $E$ a Banach space, and the integrator is a $H$-cylindrical Brownian motion on a probability space $(\Omega,\mathscr{F},\mathbb{P})$. It is advisable, however, to keep in mind the special case $H=\mathbb{R}$ which concerns the stochastic integration of $E$-valued functions with respect to a real-valued Brownian motion (cf. Corollary 6.18).

In this lecture we only consider stochastic integrals of functions $\Phi : (0,T) \to \mathscr{L}(H,E)$; such integrals are sometimes called Wiener integrals. The more delicate problem of stochastic integration of stochastic processes $\Phi : (0,T)\times\Omega \to \mathscr{L}(H,E)$ will be considered later on in this course. The theory developed in the present lecture suffices for applications to linear stochastic evolution equations with additive noise, which is the topic of the next couple of lectures.

## 6.1 Brownian motion

An $E$-valued *stochastic process* (briefly, an $E$-valued *process*) indexed by a set $I$ is a family of $E$-valued random variables $(X(i))_{i\in I}$ defined on some underlying probability space $(\Omega,\mathscr{F},\mathbb{P})$.

**Definition 6.1.** *An $E$-valued process $(X(i))_{i\in I}$ is called* Gaussian *if for all $N \geqslant 1$ and $i_1,\dots,i_N \in I$ the $E^N$-valued random variable $(X(i_1),\dots,X(i_N))$ is Gaussian.*

### 6.1.1 Brownian motion

We start with the definition.

**Definition 6.2.** *A real-valued process* $(W(t))_{t\in[0,T]}$ *is called a* Brownian motion *if it enjoys the following properties:*

(i) $W(0)=0$ *almost surely;*
(ii) $W(t)-W(s)$ *is Gaussian with variance* $t-s$ *for all* $0\leqslant s\leqslant t\leqslant T$*;*
(iii) $W(t)-W(s)$ *is independent of* $\{W(r):\ 0\leqslant r\leqslant s\}$ *for all* $0\leqslant s\leqslant t\leqslant T$*.*

In some texts, Brownian motions are called *Wiener processes*.

**Proposition 6.3.** *Every Brownian motion is a Gaussian process.*

*Proof.* Fix $t_1,\dots,t_N\in[0,T]$. By independence, the $\mathbb{R}^N$-valued random variable $(W(t_1),\, W(t_2)-W(t_1),\, \dots, W(t_N)-W(t_{N-1}))$ is Gaussian, and the random variable $(W(t_1),\dots,W(t_N))$ is obtained from it under the linear transformation $(\rho_1,\dots,\rho_N)\mapsto(\rho_1,\rho_1+\rho_2,\dots,\rho_1+\dots+\rho_N)$. □

Here is a simple way to recognise Brownian motions:

**Proposition 6.4.** *A real-valued Gaussian process* $(W(t))_{t\in[0,T]}$ *is a Brownian motion if and only if*

$$\mathbb{E}(W(s)W(t))=\min\{s,t\}\qquad\forall 0\leqslant s,t\leqslant T.$$

*Proof.* Let us first prove the 'if' part. Property (i) follows from $\mathbb{E}(W(0))^2=0$. To prove (ii) let $0\leqslant s\leqslant t\leqslant T$. Then

$$\mathbb{E}(W(t)-W(s))^2=t-2\min\{s,t\}+s=t-s.$$

For (iii) we must prove that $W(t)-W(s)$ is independent of $(W(r_1),\dots,W(r_N))$ whenever $0\leqslant r_1,\dots,r_N\leqslant s\leqslant t\leqslant T$ (cf. Definition 3.4). Noting that $(W(r_1),\dots,W(r_N))$ is the image of $(W(r_1),W(r_2)-W(r_1),\dots,W(r_N)-W(r_{N-1}))$ under a linear transformation, it suffices to prove that $W(t)-W(s)$ is independent of $(W(r_1),W(r_2)-W(r_1),\dots,W(r_N)-W(r_{N-1}))$. For this, in turn, it is enough to check that the random variables $W(r_1),W(r_2)-W(r_1),\dots,W(r_N)-W(r_{N-1}),W(t)-W(s)$ are independent. By Proposition 4.10, all we have to check is their orthogonality in $L^2(\Omega)$. But this follows from a simple computation using $\mathbb{E}(W(s)W(t))=\min\{s,t\}$.

To prove the 'only if' part let $(W(t))_{t\in[0,T]}$ be a Brownian motion. Then for all $0\leqslant s\leqslant t\leqslant T$,

$$\begin{aligned}2\mathbb{E}(W(s)W(t))&=\mathbb{E}W(s)^2+\mathbb{E}W(t)^2-\mathbb{E}(W(t)-W(s))^2\\&=s+t-(t-s)=2s=2\min\{s,t\}.\end{aligned}$$ □

In order to prove the existence of Brownian motions it will be helpful to introduce the notion of an isonormal process.

Let $\mathscr{H}$ be a Hilbert space with inner product $[\cdot,\cdot]$.

**Definition 6.5.** *An* $\mathscr{H}$-isonormal process *on* $\Omega$ *is a mapping* $\mathscr{W} : \mathscr{H} \to L^2(\Omega)$ *with the following two properties:*

(i) *For all* $h \in \mathscr{H}$ *the random variable* $\mathscr{W}h$ *is Gaussian;*
(ii) *For all* $h_1, h_2 \in \mathscr{H}$ *we have* $\mathbb{E}(\mathscr{W}h_1 \cdot \mathscr{W}h_2) = [h_1, h_2]$.

From (ii) it follows that for all scalars $c_1, c_2$ and all $h_1, h_2 \in \mathscr{H}$ one has

$$\mathbb{E}(\mathscr{W}(c_1 h_1 + c_2 h_2) - (c_1 \mathscr{W}(h_1) + c_2 \mathscr{W}(h_2)))^2 = 0.$$

As a consequence, $\mathscr{H}$-isonormal processes are linear. By linearity we have $\sum_{n=1}^N c_n \mathscr{W}h_n = \mathscr{W}(\sum_{n=1}^N c_n h_n)$, which shows that for all $h_1, \ldots, h_N \in \mathscr{H}$ the $\mathbb{R}^N$-valued random variable $(\mathscr{W}h_1, \ldots, \mathscr{W}h_N)$ is Gaussian. Stated differently, $(\mathscr{W}h)_{h\in H}$ is a Gaussian process.

*Example 6.6.* If $\mathscr{H}$ is a separable Hilbert space with orthonormal basis $(h_n)_{n=1}^\infty$ and $(\gamma_n)_{n=1}^\infty$ is a Gaussian sequence, then $\mathscr{W}h := \sum_{n=1}^\infty \gamma_n [h, h_n]$ defines an $\mathscr{H}$-isonormal process $\mathscr{W}$. The verification is an easy exercise.

The next theorem provides the existence of Brownian motions:

**Theorem 6.7.** *If* $\mathscr{W}$ *is an* $L^2(0,T)$*-isonormal process, then* $W(t) := \mathscr{W} 1_{[0,t]}$ *defines a Brownian motion on* $[0,T]$.

*Proof.* By the observation preceding Example 6.6, $(W(t))_{t\in[0,T]}$ is a Gaussian process. Since it satisfies $\mathbb{E}(W(s)W(t)) = [1_{[0,s]}, 1_{[0,t]}]_{L^2(0,T)} = \min\{s,t\}$, it is a Brownian motion by Proposition 6.4. □

The Brownian motion constructed in Theorem 6.7 is given explicitly by

$$W(t) = \sum_{n=1}^\infty \gamma_n [h_n, 1_{[0,t]}] = \sum_{n=1}^\infty \gamma_n \int_0^t h_n(s)\, ds, \tag{6.1}$$

where $(\gamma_n)_{n=1}^\infty$ is a Gaussian sequence and $(h_n)_{n=1}^\infty$ is an orthonormal basis for $L^2(0,T)$. This formula gives a profound connection between Brownian motions and the integration operator $I_T : L^2(0,T) \to C[0,T]$ of Theorem 5.23. We return to this point in Exercise 2.

So far, we have never worried about the distinction between a pointwise defined random variable $X : \Omega \to E$ and its equivalence class modulo null sets. When considering stochastic processes $(X(i))_{i\in I}$, however, one is often interested in properties of the *trajectories* $i \mapsto X(i,\omega) := (X(i))(\omega)$, where $\omega \in \Omega$. Of course these are well-defined only if the $X(i)$ are defined pointwise. Since random variables are often given only as equivalence classes (for instance, when they are constructed as elements of $L^p(\Omega; E)$), one is confronted with the problem of selecting, for each $i \in I$, a pointwise defined representative of $X(i)$. The question then arises whether these representatives can be chosen in a way that the trajectories have 'good' properties.

This discussion leads naturally to the following definition.

**Definition 6.8.** *Two (pointwise defined) processes* $X = (X(i))_{i\in I}$ *and* $\widetilde{X} = (\widetilde{X}(i))_{i\in I}$ *are* versions *of each other if for all* $i \in I$ *we have* $X(i) = \widetilde{X}(i)$ *almost surely.*

Stated differently, $X$ and $\widetilde{X}$ are versions of each other if and only if $X(i)$ and $\widetilde{X}(i)$ define the same equivalence class for each $i \in I$. From now we shall tacitly assume that processes are always pointwise defined.

The next result, due to Kolmogorov, gives a sufficient condition for the existence of a (Hölder) continuous version of an $E$-valued process $(X(t))_{t\in[0,T]}$.

**Theorem 6.9 (Kolmogorov).** *Let* $(X(t))_{t\in[0,T]}$ *be an* $E$*-valued process on* $\Omega$ *with the property that there exist real constants* $C \geqslant 0$, $\alpha > 0$, $\beta > 0$, *such that*

$$\mathbb{E}\|X(t) - X(s)\|^\alpha \leqslant C(t-s)^{1+\beta} \qquad \forall 0 \leqslant s \leqslant t \leqslant T.$$

*Then for all* $0 \leqslant \gamma < \frac{\beta}{\alpha}$, $X$ *has a version* $\widetilde{X}$ *with Hölder continuous trajectories of exponent* $\gamma$*, that is, for all* $\omega \in \Omega$ *there is a constant* $\widetilde{C}(\omega) \geqslant 0$ *such that*

$$\|\widetilde{X}(t,\omega) - \widetilde{X}(s,\omega)\| \leqslant \widetilde{C}(\omega)|t-s|^\gamma \qquad \forall 0 \leqslant s, t \leqslant T.$$

*Proof.* We may assume that $T = 1$ for notational simplicity. For $j = 0, 1, \dots$ put

$$Y_j := \sup_{0\leqslant k\leqslant 2^j-1} \|X_{(k+1)2^{-j}} - X_{k2^{-j}}\|.$$

Clearly,

$$\mathbb{E}Y_j^\alpha \leqslant \sum_{k=0}^{2^j-1} \mathbb{E}\|X_{(k+1)2^{-j}} - X_{k2^{-j}}\|^\alpha \leqslant 2^j \cdot C2^{-(1+\beta)j} = C2^{-\beta j}.$$

Set $D_j := \{k2^{-j} : k = 0, \dots, 2^j - 1\}$ and $D := \bigcup_{j=0}^\infty D_j$. Fix $j \geqslant 0$ and $s, t \in D$ satisfying $|t - s| \leqslant 2^{-j}$. For each $n \geqslant 0$ let $s_n$ and $t_n$ be the largest elements in $D_n$ such that $s_n \leqslant s$ and $t_n \leqslant t$. Then either $s_n = t_n$ or $|t_n - s_n| = 2^{-n}$. Similarly, $s_{n+1} - s_n$ and $t_{n+1} - t_n$ can only take the values 0 or $2^{-(n+1)}$. Moreover, eventually $s_n = s$ and $t_n = t$. Hence,

$$\begin{aligned}\|X_t - X_s\| &\leqslant \|X_{s_j} - X_{t_j}\| + \sum_{n=j}^\infty \|X_{t_{n+1}} - X_{t_n}\| + \sum_{n=j}^\infty \|X_{s_{n+1}} - X_{s_n}\| \\ &\leqslant Y_j + 2\sum_{n=j+1}^\infty Y_n \leqslant 2\sum_{n=j}^\infty Y_n,\end{aligned}$$

where all sums are actually finite. Fixing $0 \leqslant \gamma < \frac{\beta}{\alpha}$ we obtain

$$\begin{aligned}Z :&= \sup\{\|X_t - X_s\|/|t-s|^\gamma :\ s, t \in D,\ s \neq t\} \\ &\leqslant \sup_{j\geqslant 0}\Big\{2^{(j+1)\gamma} \sup_{2^{-(j+1)}<|t-s|\leqslant 2^{-j}} \|X_t - X_s\| :\ s, t \in D,\ s \neq t\Big\} \\ &\leqslant \sup_{j\geqslant 0}\Big(2^{(j+1)\gamma} \cdot 2\sum_{n=j}^\infty Y_n\Big) \leqslant 2^{\gamma+1}\sum_{n=0}^\infty 2^{\gamma n} Y_n.\end{aligned}$$

In case $\alpha \geqslant 1$, the triangle inequality in $L^\alpha(\Omega)$ gives

$$(\mathbb{E}Z^\alpha)^{\frac{1}{\alpha}} \leqslant 2^{\gamma+1} \sum_{n=0}^{\infty} 2^{\gamma n} (\mathbb{E}Y_n^\alpha)^{\frac{1}{\alpha}} \leqslant 2^{\gamma+1} \sum_{n=0}^{\infty} 2^{\gamma n} (C 2^{-\beta n})^{\frac{1}{\alpha}},$$

which is finite since we assumed that $\gamma < \beta/\alpha$. For $0 < \alpha < 1$ we reason similarly, replacing the triangle inequality by the inequality $(\sum_{n=0}^{\infty} |c_n|)^\alpha \leqslant \sum_{n=0}^{\infty} |c_n|^\alpha$. In either case, it follows that $Z < \infty$ almost surely.

In particular, almost surely $X$ is uniformly continuous on $D$. On the set $\{Z < \infty\}$ we define $\widetilde{X}_t = \lim_{\substack{s \to t \\ s \in D}} X_s$ and on the remaining null set we set $\widetilde{X}_t := 0$. The process $\widetilde{X}$ thus obtained has Hölder continuous trajectories of exponent $\gamma$. By Fatou's lemma and the assumption of the theorem, for all $t \in [0,1]$ we have $\widetilde{X}_t = X_t$ almost surely. Therefore $\widetilde{X}$ is a version of $X$. □

**Corollary 6.10.** *Every Brownian motion has a version with Hölder continuous trajectories for any exponent* $\gamma < \frac{1}{2}$.

*Proof.* From $\mathbb{E}|W(t) - W(s)|^2 = |t - s|$ and Exercise 1 (or the Kahane-Khintchine inequality), for $k = 1, 2, \dots$ we obtain

$$\mathbb{E}|W(t) - W(s)|^{2k} = C_k |t - s|^k,$$

and the result follows from Kolmogorov's theorem upon letting $k \to \infty$. □

### 6.1.2 Cylindrical Brownian motion

**Definition 6.11.** *An* $L^2(0,T;H)$*-isonormal process is called an* $H$-cylindrical Brownian motion *on* $[0,T]$.

$H$-Cylindrical Brownian motions will be denoted by $W_H$. For $t \in [0,T]$ and $h \in H$ we put

$$W_H(t)h := W_H(1_{(0,t)} \otimes h).$$

For each fixed $h \in H$ the process $(W_H(t)h)_{h \in H}$ is a Brownian motion, which is standard if and only if $\|h\|_H = 1$.

*Example 6.12.* If $(W^{(n)})_{n=1}^{\infty}$ is a sequence of independent Brownian motions and $H$ is a separable Hilbert space with orthonormal basis $(h_n)_{n=1}^{\infty}$, then

$$W_H(t)h := \sum_{n=1}^{\infty} W^{(n)}(t)[h, h_n]$$

defines an $H$-cylindrical Brownian motion $(W_H(t))_{t \in [0,T]}$. The easy proof is left as an exercise.

*Remark 6.13.* Let $H = L^2(D)$, where $D$ is an open subset of $\mathbb{R}^d$. An $L^2(D)$-cylindrical Brownian motion provides the mathematical model for 'space-time white noise' on $[0,T] \times D$. This explains why $H$-cylindrical Brownian motions appear naturally in the context of stochastic partial differential equations. We will return to this in later lectures.

## 6.2 The stochastic Wiener integral

After these preliminaries we turn to the problem of defining a stochastic integral of suitable functions $\Phi : (0,T) \to \mathscr{L}(H,E)$ with respect to an $H$-cylindrical Brownian motion $W_H$.

For an $\mathscr{L}(H,E)$-valued step function of the form $\Phi = 1_{(a,b)} \otimes (h \otimes x)$ with $0 \leqslant a < b \leqslant T$ and $h \in H$, $x \in E$, we define the random variable $\int_0^T \Phi \, dW_H \in L^2(\Omega; E)$ by

$$\int_0^T \Phi \, dW_H := W_H(1_{(a,b)} \otimes h) \otimes x = (W_H(b)h - W_H(a)h) \otimes x$$

and extend this definition by linearity to step functions with values in the finite rank operators in $\mathscr{L}(H,E)$; such functions will be called *finite rank step functions*. In order to extend the stochastic integral to a broader class of $\mathscr{L}(H,E)$-valued functions, just as in the classical scalar-valued theory we shall compute its square expectation.

We make the preliminary observation that any step function $\Phi : (0,T) \to \mathscr{L}(H,E)$ uniquely defines a bounded operator $R_\Phi \in \mathscr{L}(L^2(0,T;H),E)$ by the formula

$$R_\Phi f := \int_0^T \Phi(t) f(t) \, dt, \qquad f \in L^2(0,T;H).$$

**Theorem 6.14 (Itô isometry).** *For all finite rank step functions $\Phi : (0,T) \to \mathscr{L}(H,E)$ we have $R_\Phi \in \gamma(L^2(0,T;H),E)$, the stochastic integral $\int_0^T \Phi \, dW_H$ is a Gaussian random variable, and*

$$\mathbb{E}\Big\| \int_0^T \Phi \, dW_H \Big\|^2 = \|R_\Phi\|^2_{\gamma(L^2(0,T;H),E)}.$$

*Proof.* Let $\Phi := \sum_{n=1}^N 1_{(t_{n-1},t_n)} \otimes U_n$ with $0 \leqslant t_0 < \cdots < t_N \leqslant T$ and the operators $U_n \in \mathscr{L}(H,E)$ of finite rank. It is an easy exercise in linear algebra to check that there is no loss of generality in assuming that $U_n = \sum_{j=1}^k h_j \otimes x_{jn}$, where the vectors $h_1, \dots, h_k \in H$ are orthonormal (and do not depend on $n$). Since $R_\Phi$ is of finite rank, it belongs to $\gamma(L^2(0,T;H),E)$.

Put $\phi_n := c_n 1_{(t_{n-1},t_n)}$, where the normalising constant $c_n := 1/\sqrt{t_n - t_{n-1}}$ assures that the functions $\phi_1, \dots, \phi_N$ are orthonormal in $L^2(0,T)$. The sequence $(\phi_n \otimes h_j)_{\substack{1 \leqslant j \leqslant k \\ 1 \leqslant n \leqslant N}}$ is orthonormal in $L^2(0,T;H)$, and from Lemma 5.7 we obtain that

$$
\begin{aligned}
\|R_\Phi\|^2_{\gamma(L^2(0,T;H);E)} &= \mathbb{E}\Big\|\sum_{j=1}^k\sum_{n=1}^N \gamma_{jn} R_\Phi(\phi_n\otimes h_j)\Big\|^2 \\
&= \mathbb{E}\Big\|\sum_{j=1}^k\sum_{n=1}^N \gamma_{jn}\int_0^T c_n 1_{(t_{n-1},t_n)}(t)U_n h_j\,dt\Big\|^2 \\
&= \mathbb{E}\Big\|\sum_{j=1}^k\sum_{n=1}^N \gamma_{jn}\sqrt{t_n-t_{n-1}}x_{jn}\Big\|^2,
\end{aligned}
$$

where $(\gamma_{jn})_{\substack{1\leqslant j\leqslant k\\ 1\leqslant n\leqslant N}}$ is a Gaussian sequence. On the other hand,

$$
\begin{aligned}
\mathbb{E}\Big\|\int_0^T \Phi\,dW_H\Big\|^2 &= \mathbb{E}\Big\|\sum_{j=1}^k\sum_{n=1}^N (W_H(t_n)h_j - W_H(t_{n-1})h_j)\otimes x_{jn}\Big\|^2 \\
&= \mathbb{E}\Big\|\sum_{j=1}^k\sum_{n=1}^N \frac{W_H(t_n)h_j - W_H(t_{n-1})h_j}{\sqrt{t_n-t_{n-1}}}\otimes\sqrt{t_n-t_{n-1}}x_{jn}\Big\|^2.
\end{aligned}
$$

Putting $\gamma'_{jn} := (W_H(t_n)h_j - W_H(t_{n-1})h_j)/\sqrt{t_n-t_{n-1}}$, the desired identity now follows since $(\gamma'_{jn})_{\substack{1\leqslant j\leqslant k\\ 1\leqslant n\leqslant N}}$ is a Gaussian sequence. □

As a consequence, the linear mapping $J_T^{W_H} : R_\Phi \mapsto \int_0^T \Phi\,dW_H$ uniquely extends to an isometric embedding

$$J_T^{W_H} : \gamma(L^2(0,T;H),E) \to L^2(\Omega;E).$$

Accordingly, the stochastic integral of an operator $R\in\gamma(L^2(0,T;H),E)$ can be defined as $J_T^{W_H}(R)$. In order for this to be useful we need a way to recognise those $\mathscr{L}(H,E)$-valued functions which 'represent' an operator in $\gamma(L^2(0,T;H),E)$. To this problem we turn next.

For a function $\Phi:(0,T)\to\mathscr{L}(H,E)$ and elements $h\in H$ and $x^*\in E^*$ we define $\Phi h:(0,T)\to E$ and $\Phi^*x^*:(0,T)\to H$ by $(\Phi h)(t) := \Phi(t)h$ and $(\Phi^*x^*)(t) := \Phi^*(t)x^*$ (where of course $\Phi^*(t) := (\Phi(t))^*$).

**Definition 6.15.** *A function $\Phi:(0,T)\to\mathscr{L}(H,E)$ is said to be* stochastically integrable *with respect to the $H$-cylindrical Brownian motion $W_H$ if there exists a sequence of finite rank step functions $\Phi_n:(0,T)\to\mathscr{L}(H,E)$ such that:*

(i) *for all $h\in H$ we have $\lim_{n\to\infty}\Phi_n h = \Phi h$ in measure;*
(ii) *there exists an $E$-valued random variable $X$ such that $\lim_{n\to\infty}\int_0^T\Phi_n\,dW_H = X$ in probability.*

*The* stochastic integral *of a stochastically integrable function $\Phi:(0,T)\to\mathscr{L}(H,E)$ is then defined as the limit in probability*

$$\int_0^T \Phi\,dW_H := \lim_{n\to\infty} \int_0^T \Phi_n\,dW_H.$$

Three remarks are in order.

(a) Condition (i) means $\lim_{n\to\infty} |\{t \in (0,T) : \|\Phi_n(t)h - \Phi(t)h\| > r\}| = 0$ for all $h \in H$ and $r > 0$, where $|B|$ denotes the Lebesgue measure of $B$.
(b) The stochastic integral is well defined in the sense that it is independent of the approximating sequence.
(c) From Theorem 4.15 it follows that the convergence in probability in condition (ii) is equivalent to convergence in $L^p(\Omega; E)$ for some (all) $1 \leqslant p < \infty$.

In the special case $E = \mathbb{R}$ we may identify $\mathscr{L}(H, \mathbb{R}) = H^*$ with $H$ by the Riesz representation theorem. Under this identification, Theorem 6.14 reduces to the statement that the stochastic integral of a step function $\phi : (0,T) \to H$ satisfies

$$\mathbb{E}\Big\| \int_0^T \phi\,dW_H \Big\|^2 = \|\phi\|^2_{L^2(0,T;H)}. \tag{6.2}$$

From this it is immediate that a strongly measurable function $\phi : (0,T) \to H$ is stochastically integrable with respect to $W_H$ if and only if $\phi \in L^2(0,T;H)$, and the isometry (6.2) extends to functions $\phi \in L^2(0,T;H)$.

**Definition 6.16.** *A function $\Phi : (0,T) \to \mathscr{L}(H,E)$ is called $H$-*strongly measurable *if for each $h \in H$ the function $\Phi h : (0,T) \to E$ is strongly measurable.*

By Theorem 6.14 and a limiting argument, we see that if a function $\Phi$ is stochastically integrable with respect to $W_H$, then the integral operator $R_\Phi$ associated with $\Phi$ is well-defined and $\gamma$-radonifying. Interestingly, the converse is true as well. These two statements are contained in the next theorem, which is the main result of this lecture.

**Theorem 6.17.** *Let $W_H$ be an $H$-cylindrical Brownian motion. For an $H$-strongly measurable function $\Phi : (0,T) \to \mathscr{L}(H,E)$ the following assertions are equivalent:*

(1) *$\Phi$ is stochastically integrable with respect to $W_H$;*
(2) *$\Phi^* x^* \in L^2(0,T;H)$ for all $x^* \in E^*$, and there exists an $E$-valued random variable $X$ such that for all $x^* \in E^*$, almost surely we have*

$$\langle X, x^* \rangle = \int_0^T \Phi^* x^*\,dW_H;$$

(3) *$\Phi^* x^* \in L^2(0,T;H)$ for all $x^* \in E^*$, and there exists an operator $R \in \gamma(L^2(0,T;H), E)$ such that for all $f \in L^2(0,T;H)$ and $x^* \in E^*$ we have*

$$\langle Rf, x^* \rangle = \int_0^T \langle \Phi(t) f(t), x^* \rangle\,dt.$$

*If these equivalent conditions are satisfied, the random variable $X$ and the operator $R$ are uniquely determined, we have $X = \int_0^T \Phi\,dW_H$ almost surely, and*

$$\mathbb{E}\Big\| \int_0^T \Phi\,dW_H \Big\|^2 = \|R\|^2_{\gamma(L^2(0,T;H),E)}.$$

In the situation of (3) we say that $\Phi$ *represents* the operator $R$. Note that condition (3) does not depend on the particular choice of $W_H$.

*Proof.* We shall prove the implications (1)⇒(2)⇒(4)⇒(3)⇒(1), where

(4) $\Phi^* x^* \in L^2(0,T;H)$ for all $x^* \in E^*$, and there exists a $\gamma$-radonifying operator $\widetilde{R}$ from a Hilbert space $\widetilde{H}$ to $E$ such that for all $x^* \in E^*$ we have

$$\|\Phi^* x^*\|_{L^2(0,T;H)} \leqslant \|\widetilde{R}^* x^*\|_{\widetilde{H}}.$$

(1)⇒(2): Let $(\Phi_n)_{n=1}^\infty$ be an approximating sequence of finite rank step functions for $\Phi$ and take $X := \int_0^T \Phi\,dW_H$. As we have already observed, $\lim_{n\to\infty} \int_0^T \Phi_n\,dW_H = X$ in $L^2(\Omega;E)$. Hence,

$$\lim_{n\to\infty} \int_0^T \Phi_n^* x^*\,dW_H = \lim_{n\to\infty} \Big\langle \int_0^T \Phi_n\,dW_H, x^* \Big\rangle = \langle X, x^* \rangle$$

in $L^2(\Omega)$, where the first identity is verified by writing out the definitions. By the special case of the Itô isometry contained in (6.2), the sequence $(\Phi_n^* x^*)_{n=1}^\infty$ is Cauchy in $L^2(0,T;H)$. Let $f$ be its limit. Since $\lim_{n\to\infty} \langle \Phi_n h, x^* \rangle = \langle \Phi h, x^* \rangle$ in measure, it follows that $f = \Phi^* x^*$ in $L^2(0,T;H)$. Once more by (6.2),

$$\lim_{n\to\infty} \int_0^T \Phi_n^* x^*\,dW_H = \int_0^T \Phi^* x^*\,dW_H.$$

(2)⇒(4): Let $i_X \in \gamma(H_X, E)$ be defined by (4.3). Then, by (6.2),

$$\int_0^T \|\Phi^*(t) x^*\|^2\,dt = \mathbb{E}\Big\| \int_0^T \Phi^* x^*\,dW_H \Big\|^2 = \mathbb{E}\langle X, x^* \rangle^2 = \|i_X^* x^*\|^2.$$

(4)⇒(3): The formula

$$(R_\Phi f)(x^*) := \int_0^T [f(t), \Phi^*(t) x^*]\,dt, \qquad f \in L^2(0,T;H),\ x^* \in E^*,$$

defines a bounded operator $R_\Phi$ from $L^2(0,T;H)$ to $E^{**}$. Once we know that $R_\Phi$ maps $L^2(0,T;H)$ into $E$, Theorem 5.17 shows that $R_\Phi \in \gamma(L^2(0,T;H),E)$.

For the proof that $R_\Phi$ takes values in $E$ we invoke Theorem 1.20. By assumption, for all $h \in H$ the function $\Phi h$ is strongly measurable and the functions $\langle \Phi h, x^* \rangle = [h, \Phi^* x^*]$ are square integrable. It follows that $\Phi h$ is Pettis integrable. Therefore, for step functions $f$, the element $R_\Phi f \in E^{**}$ is given by the Pettis integral $\int_0^T \Phi(t)f\,dt$ in $E$. Thus, $R_\Phi f \in E$ for all step functions $f : (0,T) \to H$. Since these functions are dense in $L^2(0,T;H)$, a limiting argument implies that $R_\Phi f \in E$ for all $f \in L^2(0,T;H)$.

(3)⇒(1): We split the proof into three steps.

*Step 1* - We begin by constructing an $E$-valued random variable $X$, which will turn out later to be the stochastic integral $\int_0^T \Phi\,dW_H$.

By Proposition 5.10 there is a separable closed subspace $\mathscr{H}_0$ of $L^2(0,T;H)$ such that $Rf = 0$ for all $f \in \mathscr{H}_0^\perp$. Choose a separable closed subspace $H_0$ of $H$ such that $\mathscr{H}_0 \subseteq L^2(0,T;H_0)$. Note that the range of $R^*$ is contained in $\mathscr{H}_0$, hence in $L^2(0,T;H_0)$.

Let $(f_m)_{m=1}^\infty$ and $(h_n)_{n=1}^\infty$ be orthonormal bases for $L^2(0,T)$ and $H_0$, respectively. The functions $\phi_{mn} := f_m \otimes h_n$ define an orthonormal basis $(\phi_{mn})_{m,n=1}^\infty$ for $L^2(0,T;H_0)$. By (6.2) the random variables $\gamma_{mn} := \int_0^T \phi_{mn}\,dW_H$ are standard Gaussian, and the linearity of the stochastic integral implies that they are jointly Gaussian. The orthonormality of the $\phi_{mn}$ implies that the $\gamma_{mn}$ are orthonormal in $L^2(\Omega)$, and therefore independent by Proposition 4.10. Thus we have shown that $(\gamma_{mn})_{m,n=1}^\infty$ is a Gaussian sequence.

Put

$$X := \sum_{m,n=1}^\infty \gamma_{mn} R\phi_{mn}.$$

This sum converges in $L^2(\Omega;E)$ by Theorem 5.15. Moreover, the identity $\langle R\phi_{mn}, x^*\rangle = [\Phi^* x^*, \phi_{mn}]_{L^2(0,T;H_0)}$ implies $\Phi^* x^* = R^* x^* \in L^2(0,T;H_0)$ and

$$\begin{aligned}\langle X, x^*\rangle &= \sum_{m,n=1}^\infty \int_0^T \langle R\phi_{mn}, x^*\rangle \phi_{mn}\,dW_H \\ &= \int_0^T \sum_{m,n=1}^\infty \langle R\phi_{mn}, x^*\rangle \phi_{mn}\,dW_H = \int_0^T \Phi^* x^*\,dW_H,\end{aligned} \tag{6.3}$$

where the second identity follows from $L^2(0,T;H)$-convergence and (6.2).

*Step 2* - Define the operators $\Phi_k(t) \in \mathscr{L}(H,E)$ by

$$\Phi_k(t)h := \sum_{j=1}^{2^k} 1_{(\frac{(j-1)T}{2^k}, \frac{jT}{2^k})}(t) R U_{jk} h,$$

where $U_{jk} \in \mathscr{L}(H, L^2(0,T;H))$ is given by $U_{jk}h := \frac{2^k}{T} 1_{(\frac{(j-1)T}{2^k}, \frac{jT}{2^k})} \otimes h$. Note that $RU_{jk} \in \gamma(H,E)$ by the ideal property. Hence, each $\Phi_k$ is an $\gamma(H,E)$-valued step function. The identity

$$\langle \Phi_k(t)h, x^*\rangle = \sum_{j=1}^{2^k} 1_{(\frac{(j-1)T}{2^k}, \frac{jT}{2^k})} \frac{2^k}{T} \int_{\frac{(j-1)T}{2^k}}^{\frac{jT}{2^k}} \langle \Phi(t)h, x^*\rangle\,dt$$

shows that $\Phi_k$ is obtained from $\Phi$ by averaging. We will show that

(i) $\lim_{k\to\infty} \Phi_k h = \Phi h$ in measure for all $h \in H$,

(ii) $\lim_{k\to\infty} \int_0^T \Phi_k \, dW_H = X$ in probability, where $X$ is as in Step 1.

To prove (i) fix $h \in H$ and assume that $\|h\| = 1$. To get around the difficulty that we cannot be sure that $\Phi h \in L^2(0,T;E)$ we do a trunction argument.

Fix an arbitrary $\varepsilon > 0$. For $r > 0$ define $S^{(r)} \in \mathscr{L}(L^2(0,T;H))$ by $S^{(r)} f := 1_{\{\|\Phi(t)h\|\leqslant r\}} f$. Using Proposition 5.12 choose $r_0 > 0$ so large that

$$\|R - RS^{(r_0)}\|_{\gamma(L^2(0,T;H),E)} < \varepsilon, \quad |\{t \in (0,T) : \ \|\Phi(t)h - f^{(r_0)}(t)\| > \varepsilon\}| < \varepsilon,$$

where $f^{(r_0)}(t) := 1_{\{\|\Phi(t)h\|\leqslant r_0\}} \Phi(t)h$. Since $f^{(r_0)} \in L^2(0,T;E)$, by the properties of averaging operators (see Exercise 3) we have

$$f^{(r_0)} = \lim_{k\to\infty} \sum_{j=1}^{2^k} 1_{(\frac{(j-1)T}{2^k}, \frac{jT}{2^k})} \frac{2^k}{T} \int_{\frac{(j-1)T}{2^k}}^{\frac{jT}{2^k}} f^{(r_0)}(t)\, dt = \lim_{k\to\infty} f_k^{(r_0)} \tag{6.4}$$

in $L^2(0,T;E)$, where $f_k^{(r_0)} := \sum_{j=1}^{2^k} 1_{(\frac{(j-1)T}{2^k}, \frac{jT}{2^k})} RS^{(r_0)} U_{jk} h$.

If $s \in (\frac{(j-1)T}{2^k}, \frac{jT}{2^k})$, then

$$\|f_k^{(r_0)}(s) - \Phi_k(s)h\| = \|RS^{(r_0)}U_{jk}h - RU_{jk}h\| \leqslant \|R - RS^{(r_0)}\|_{\gamma(L^2(0,T;H),E)} < \varepsilon.$$

Hence,

$$\begin{aligned}
&|\{t \in (0,T) : \ \|\Phi(t)h - \Phi_k(t)h\| > 3\varepsilon\}| \\
&\quad \leqslant \varepsilon + |\{t \in (0,T) : \ \|f^{(r_0)}(t) - f_k^{(r_0)}(t)\| > \varepsilon\}| \\
&\qquad + |\{t \in (0,T) : \ \|f_k^{(r_0)}(t) - \Phi_k(t)h\| > \varepsilon\}| \\
&\quad = \varepsilon + |\{t \in (0,T) : \ \|f^{(r_0)}(t) - f_k^{(r_0)}\| > \varepsilon\}|.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, (i) follows from (6.4) by letting $k \to \infty$.

We continue with the proof of (ii). Put

$$X_1 = \int_0^T \Phi_1 \, dW_H, \qquad X_n = \int_0^T (\Phi_n - \Phi_{n-1})\, dW_H \quad \text{for} \quad n \geqslant 2.$$

We claim that the random variables $X_n$ are independent. By the linearity of the stochastic integral, the random variables $X_n$ are jointly Gaussian and therefore by Proposition 4.10 it suffices to check that $\mathbb{E}\langle X_m, x^*\rangle\langle X_n, y^*\rangle = 0$ for $m \neq n$ and $x^*, y^* \in E^*$. By (6.2) and linearity, the expectation equals

$$\int_0^T [\Phi_m^*(t)x^* - \Phi_{m-1}^*(t)x^*, \ \Phi_n^*(t)y^* - \Phi_{n-1}^*(t)y^*]\, dt$$

using the convention that $\Phi_0 = 0$. By a direct computation using the properties of the averaging operators, this expression equals 0.

Put $S_N := \sum_{n=1}^N X_n = \int_0^T \Phi_N \, dW_H$. By (6.3), (6.2), and the properties of averaging operators, for all $x^* \in E^*$ we have

$$\lim_{N\to\infty} \mathbb{E}\langle X - S_N, x^*\rangle^2 = \lim_{N\to\infty} \|\Phi^* x^* - \Phi_N^* x^*\|^2_{L^2(0,T;H)} = 0$$

and therefore $\lim_{N\to\infty}\langle S_N, x^*\rangle = \langle X, x^*\rangle$ in probability. The Itô-Nisio theorem implies that $\lim_{N\to\infty} S_N = X$ in probability.

*Step 3* – So far, we have found a sequence of $\gamma(H,E)$-valued step functions $(\Phi_n)_{n=1}^\infty$ with the convergence properties as required in Definition 6.15. To conclude the proof we approximate the values of the functions $\Phi_n$ by finite rank operators. □

**Corollary 6.18.** *A strongly measurable function $\phi : (0,T) \to E$ is stochastically integrable with respect to a real-valued Brownian motion if and only if $\phi$ represents an operator $R \in \gamma(L^2(0,T), E)$.*

As an application of Theorem 6.17 we have the following domination criterion for stochastic integrability.

**Theorem 6.19.** *Suppose that $\Phi_1, \Phi_2 : (0,T) \to \mathscr{L}(H,E)$ are $H$-strongly measurable functions, and assume that $\Phi_2$ stochastically integrable with respect to the $H$-cylindrical Brownian motion $W_H$. If*

$$\int_0^T \|\Phi_1^*(t)x^*\|^2\,dt \leqslant \int_0^T \|\Phi_2^*(t)x^*\|^2\,dt \qquad \forall x^* \in E^*,$$

*then $\Phi_1$ is stochastically integrable with respect to $W_H$, and for all $1 \leqslant p < \infty$ we have*

$$\mathbb{E}\Big\|\int_0^T \Phi_1\,dW_H\Big\|^p \leqslant \mathbb{E}\Big\|\int_0^T \Phi_2\,dW_H\Big\|^p.$$

*Proof.* First note that by Theorem 6.17, for all $x^* \in E^*$ the function $\Phi_2^* x^*$ belongs to $L^2(0,T;H)$. By (4) in the proof of Theorem 6.17, $\Phi_2$ represents an operator $R_2 \in \gamma(L^2(0,T;H),E)$. In view of $R_2^* x^* = \Phi_2^* x^*$ we have

$$\int_0^T \|\Phi_2^*(t)x^*\|^2\,dt = \|R_2^* x^*\|^2_{L^2(0,T;H)}.$$

Let $R_{\Phi_1} \in \gamma(L^2(0,T;H),E)$ denote the operator representing $\Phi_1$ whose existence is assured by Theorem 6.17 (3). The first assertion follows by applying Theorem 6.17 to $\Phi_1$ and the $L^p$-inequality follows from Corollary 5.18. □

## 6.3 Exercises

1. (!) Let $\gamma$ be a Gaussian random variable with variance $\mathbb{E}\gamma^2 = q$. Compute $\mathbb{E}\gamma^{2k}$, $k = 1, 2, \ldots$
   *Hint:* Express $\mathbb{E}\gamma^{2k+2}$ in terms of $\mathbb{E}\gamma^{2k}$.

2. In view of the identity (6.1), Theorem 5.23 provides another proof of the existence of a continuous version for Brownian motions. In this exercise we show that in the converse direction Theorem 5.23 can be deduced from the path continuity of Brownian motions.
   Let $W$ be a Brownian motion and let $\widetilde{W}$ be a version of it with continuous trajectories.
   a) Use the Pettis measurability theorem to prove that the function $X_T : \Omega \to C[0,T]$ defined by $(X_T(\omega))(t) := \widetilde{W}(t,\omega)$ is strongly measurable.
   *Hint:* The Dirac measures span a norming subspace in $(C[0,T])^*$.
   b) Show that the random variable $X_T$ is Gaussian.
   c) Show that the covariance operator $Q_T$ of $X_T$ is given by $Q_T = I_T I_T^*$, where $I_T : L^2(0,T) \to C[0,T]$ is the integration operator of Theorem 5.23, and deduce from this that $I_T$ is $\gamma$-radonifying.

3. Fix $1 \leqslant p < \infty$. For $n = 0,1,2,\ldots$ define the linear operators $A_n : L^p(0,T;E) \to L^p(0,T;E)$ by
$$A_n f := \sum_{j=1}^{2^n} 1_{(\frac{(j-1)T}{2^n}, \frac{jT}{2^n})} \otimes x_{jn},$$
   where
$$x_{jn} := \frac{2^n}{T} \int_{\frac{(j-1)T}{2^n}}^{\frac{jT}{2^n}} f(t)\,dt.$$
   a) Show that each $A_n$ is bounded and satisfies $\|A_n\| = 1$.
   b) Show that $\lim_{n\to\infty} A_n f = f$ in $L^p(0,T;E)$ for all $f \in L^p(0,T;E)$.
   *Hint:* What happens if $f$ is a dyadic step function?
   c) Prove the assertion involving averaging operators in Step 3 (ii) of the proof of (3)$\Rightarrow$(1) of Theorem 6.17.

4. Let the function $\Phi : (0,T) \to \mathscr{L}(H,E)$ be stochastically integrable with respect to $W_H$.
   a) Show that for all $t \in [0,T]$ the restriction of $\Phi|_{(0,t)}$ is stochastically integrable on $(0,t)$ with respect to $W_H$, that $1_{(0,t)}\Phi$ is stochastically integrable on $(0,T)$ with respect to $W_H$, and that almost surely
$$\int_0^t \Phi\,dW_H = \int_0^T 1_{(0,t)}\Phi\,dW_H.$$
   We consider the $E$-valued process $X$, where $X_t = \int_0^t \Phi\,dW_H$ for $t \in [0,T]$.
   b) Show that $X$ is a Gaussian process.
   c) Show that $X$ has trajectories in $L^p(0,T;E)$ almost surely for every $1 \leqslant p < \infty$.
   *Hint:* Prove the stronger statement that $\mathbb{E}\int_0^T \|X(t)\|^p\,dt < \infty$.

*Remark:* Using martingale techniques it can be shown that $X$ has a continuous version. We return to this later on.

5. Suppose that $\Phi : (0,T) \to \mathscr{L}(H,E)$ is stochastically integrable with respect to the $H$-cylindrical Brownian motion $W_H$.
   a) Show that for each $h \in H$ function $\Phi h$ is stochastically integrable with respect to the Brownian motions $W_H h$.
   b) Prove the following series expansion: if $H$ is separable, then for any orthonormal basis $(h_n)_{n=1}^\infty$ of $H$ we have
   $$\int_0^T \Phi \, dW_H = \sum_{n=1}^\infty \int_0^T \Phi h_n \, dW_H h_n,$$
   with convergence almost surely and in $L^p(\Omega; E)$ for all $1 \leqslant p < \infty$.
   *Hint:* First consider the functions $\Phi P_n$, where $P_n$ is the orthogonal projection in $H$ onto the span of $\{h_1, \dots, h_n\}$.

**Notes.** The notion of Brownian motion has its origin in the observations by the botanist Brown (1773-1858) who observed that small particles suspended in a fluid display random movements. The first rigorous mathematical treatment was given by Wiener in the 1920s.

The proof of Theorem 6.9 is taken from Revuz and Yor [94]. Its Corollary 6.10 is nearly optimal in the following sense: almost surely, one has
$$\limsup_{\delta \downarrow 0} \max_{|t-s| \leqslant \delta} \frac{|W(t) - W(s)|}{\sqrt{2|t-s| \ln(1/|t-s|)}} = 1.$$
This is a classical result of Lévy. In particular it shows that almost surely the paths of a Brownian motion are nowhere Hölder continuous of exponent $\frac{1}{2}$. For proofs and further results on Brownian motion we refer to Karatzas and Shreve [59] and Revuz and Yor [94]. A more recent result of Ciesielski [23] asserts that almost surely, the trajectories of Brownian motions belong to the Besov space $B^{\frac{1}{2}}_{p,\infty}$ for all $1 \leqslant p < \infty$. This result was extended to Banach space-valued Brownian motions, with a simpler proof, by Hytönen and Veraar [51].

For accessible introductions to the classical (scalar-valued) theory of stochastic integration we refer to the books by Chung and Williams [22], Kuo [63], Oksendal [86], and Steele [99]. For scalar-valued functions, the isometry of Theorem 6.14 goes back to Wiener and was generalised to stochastic processes in the fundamental work of Itô.

By combining the observation on Kwapień's theorem in the Notes of the previous lecture with Corollary 6.18 we obtain that the following assertions are equivalent for a Banach space $E$:

(1) the space of strongly measurable $E$-valued functions $f : (0,T) \to E$ which are stochastically integrable with respect to Brownian motion equals $L^2(0,T;E)$;

(2) the space $E$ is isomorphic to a Hilbert space.

An explicit example of a uniformly bounded function $f : (0,1) \to l^p$ for $1 \leqslant p < 2$ which fails to be stochastically integrable was constructed in an early stage of the theory by Yor [110]. Further examples along this line were constructed Rosiński and Suchanecki [96] who also proved (for $H = \mathbb{R}$) the equivalence (1)⇔(2) of Theorem 6.17. Step 3 of the proof of (3)⇒(1) in Theorem 6.17 is a variation of their argument. In its present formulation, Theorem 6.17 can be found in [84]; a preliminary version was obtained in [16] by using different methods. The idea in Step 2 of the proof of (3)⇒(1) is taken from [84]. The implication (3)⇒(1) can alternatively be derived from variant of Theorem 5.24. This is the approach taken in [84], where also Theorems 6.14, 6.19, and the result of Exercise 5 were obtained.

In the Hilbert space literature, the series expansions of Example 6.12 and Exercise 5 are often taken as the starting point for defining the stochastic integral; see for instance the monograph of Da Prato and Zabczyk [27].

A more probabilistic approach to the theory of stochastic integration in Banach spaces is taken by Métivier and Pellaumail [76].

# 7

# Semigroups of linear operators

Having developed the probabilistic tools needed for our study of stochastic evolution equations, in this lecture we turn to the theory of $C_0$-semigroups. We review their basic properties and show how semigroups are used to solve the (deterministic) inhomogeneous abstract Cauchy problem

$$u'(t) = Au(t) + f(t).$$

Here $A$ generates a $C_0$-semigroup on $E$ and the forcing term $f$ is a locally integrable $E$-valued function. As we shall see in the next lecture, the techniques for solving this problem by means of so-called weak and strong solutions can be extended to stochastic abstract Cauchy problems with additive noise, the main difference being that Bochner integrals are replaced by the stochastic integrals introduced in the previous lecture. Heuristically, the reason why this works is that the noise can be viewed as a 'random' forcing term.

## 7.1 $C_0$-semigroups

Linear equations of mathematical physics can often be cast in the abstract form

$$\begin{cases} u'(t) = Au(t), & t \in [0,T], \\ u(0) = x, \end{cases} \tag{ACP}$$

where $A$ is a linear, usually unbounded, operator defined on a linear subspace $\mathscr{D}(A)$, the *domain* of $A$, of a Banach space $E$. Typically, $E$ is a Banach space of functions suited for the particular problem and $A$ is a partial differential operator. The abstract initial value problem (ACP) is referred to as the *abstract Cauchy problem* associated with $A$.

*Example 7.1.* Let $D$ be an open domain in $\mathbb{R}^d$ with topological boundary $\partial D$. On $D$ we consider the heat equation

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,\xi) = \Delta u(t,\xi), & t \in [0,T],\ \xi \in D; \\ u(t,\xi) = 0, & t \in [0,T],\ \xi \in \partial D; \\ u(0,\xi) = u_0(\xi), & \xi \in D. \end{cases} \tag{7.1}$$

For initial values $x = u_0 \in L^p(D)$ with $1 \leqslant p < \infty$, this problem can be rewritten in the abstract form (ACP) by taking $E = L^p(D)$ and defining $A$ by

$$\begin{aligned} \mathscr{D}(A) &:= \{f \in W^{2,p}(D) :\ f|_{\partial D} \equiv 0\} = W^{2,p}(D) \cap W_0^{1,p}(D), \\ Af &:= \Delta f, \quad f \in \mathscr{D}(A). \end{aligned}$$

Here, $W^{k,p}(D)$ is the Sobolev space of all $f \in L^p(D)$ whose weak partial derivatives up to order $k$ exist and belong to $L^p(D)$, $W_0^{k,p}(D)$ is the closure in $W^{k,p}(D)$ of all test functions $f \in C_c^\infty(D)$, and $\Delta = \sum_{j=1}^d \frac{\partial^2}{\partial \xi_j^2}$ is the Laplacian. Note how the boundary condition is built into the definition of $A$ by the specification of its domain.

The idea is now that instead of looking for a solution $u : [0,T] \times D \to \mathbb{R}$ of (7.1) one looks for a solution $u : [0,T] \to L^p(D)$ of (ACP). To get an idea how this may be done we first take a look at the much simpler case where $E = \mathbb{R}^d$ and $A : \mathscr{D}(A) = E \to E$ is represented by a $(d \times d)$-matrix. In that case, the unique solution of (ACP) is given by

$$u(t) = e^{tA} u_0, \qquad t \in [0,T],$$

where $e^{tA} = \sum_{n=0}^\infty \frac{t^n A^n}{n!}$. The matrices $e^{tA}$ may be thought of as 'solution operators' mapping the initial value $u_0$ to the solution $e^{tA}u_0$ at time $t$. Clearly, $e^{0A} = I$, $e^{tA}e^{sA} = e^{(t+s)A}$, and $t \mapsto e^{tA}$ is continuous. We generalise these properties to infinite dimensions as follows.

Let $E$ be a real or complex Banach space.

**Definition 7.2.** *A family $S = \{S(t)\}_{t \geqslant 0}$ of bounded linear operators acting on a Banach space $E$ is called a $C_0$*-semigroup *if the following three properties are satisfied:*

S1. $S(0) = I$;
S2. $S(t)S(s) = S(t+s)$ *for all* $t, s \geqslant 0$;
S3. $\lim_{t \downarrow 0} \|S(t)x - x\| = 0$ *for all* $x \in E$.

*The* infinitesimal generator*, or briefly the* generator*, of $S$ is the linear operator $A$ with domain $\mathscr{D}(A)$ defined by*

$$\mathscr{D}(A) = \{x \in E :\ \lim_{t \downarrow 0} \frac{1}{t}(S(t)x - x) \ \textit{exists}\},$$

$$Ax = \lim_{t \downarrow 0} \frac{1}{t}(S(t)x - x), \quad x \in \mathscr{D}(A).$$

We shall frequently use the trivial observation that if $A$ generates the $C_0$-semigroup $(S(t))_{t\geqslant 0}$, then $A-\mu$ generates the $C_0$-semigroup $(e^{-\mu t}S(t))_{t\geqslant 0}$.

The next two propositions collect some elementary properties of $C_0$-semigroups and their generators.

**Proposition 7.3.** *Let $S$ be a $C_0$-semigroup on $E$. There exist constants $M \geqslant 1$ and $\mu \in \mathbb{R}$ such that $\|S(t)\| \leqslant Me^{\mu t}$ for all $t \geqslant 0$.*

*Proof.* There exists a number $\delta > 0$ such that $\sup_{t\in[0,\delta]} \|S(t)\| =: \sigma < \infty$. Indeed, otherwise we could find a sequence $t_n \downarrow 0$ such that $\lim_{n\to\infty} \|S(t_n)\| = \infty$. By the uniform boundedness theorem, this implies the existence of an $x \in E$ such that $\sup_{n\geqslant 1} \|S(t_n)x\| = \infty$, contradicting the strong continuity assumption (S3). This proves the claim. By the semigroup property (S2), for $t \in [(k-1)\delta, k\delta]$ it follows that $\|S(t)\| \leqslant \sigma^k \leqslant \sigma^{(t+1)/\delta}$, where the second inequality uses that $\sigma \geqslant 1$ by (S1). This proves the proposition, with $M = \sigma^{\frac{1}{d}}$ and $\mu = \frac{1}{d}\ln\sigma$. ∎

**Proposition 7.4.** *Let $S$ be a $C_0$-semigroup on $E$ with generator $A$.*

(1) *For all $x \in E$ the orbit $t \mapsto S(t)x$ is continuous for $t \geqslant 0$.*
(2) *For all $x \in \mathscr{D}(A)$ and $t \geqslant 0$ we have $S(t)x \in \mathscr{D}(A)$ and $AS(t)x = S(t)Ax$.*
(3) *For all $x \in E$ we have $\int_0^t S(s)x\,ds \in \mathscr{D}(A)$ and*
$$A\int_0^t S(s)x\,ds = S(t)x - x.$$
*If $x \in \mathscr{D}(A)$, then both sides are equal to $\int_0^t S(s)Ax\,ds$.*
(4) *The generator $A$ is a closed and densely defined operator.*
(5) *For all $x \in \mathscr{D}(A)$ the orbit $t \mapsto S(t)x$ is continuously differentiable for $t \geqslant 0$ and*
$$\frac{d}{dt}S(t)x = AS(t)x = S(t)Ax, \quad t \geqslant 0.$$

*Proof.* (1): The right continuity of $t \mapsto S(t)x$ follows from the right continuity at $t = 0$ (S3) and the semigroup property (S2). For the left continuity, observe that
$$\|S(t)x - S(t-h)x\| \leqslant \|S(t-h)\|\|S(h)x - x\| \leqslant \sup_{s\in[0,t]} \|S(s)\|\|S(h)x - x\|,$$
where the supremum is finite by Proposition 7.3.

(2): This follows from the semigroup property:
$$\lim_{h\downarrow 0} \frac{1}{h}(S(t+h)x - S(t)x) = S(t)\lim_{h\downarrow 0}\frac{1}{h}(S(h)x - x) = S(t)Ax.$$

(3): The first identity follows from

$$
\begin{aligned}
\lim_{h\downarrow 0}\frac{1}{h}(S(h)-I)\int_0^t S(s)x\,ds &= \lim_{h\downarrow 0}\frac{1}{h}\Big(\int_0^t S(s+h)x\,ds - \int_0^t S(s)x\,ds\Big)\\
&= \lim_{h\downarrow 0}\frac{1}{h}\Big(\int_t^{t+h} S(s)x\,ds - \int_0^h S(s)x\,ds\Big)\\
&= S(t)x - x,
\end{aligned}
$$

where we used the continuity of $t\mapsto S(t)x$. The identity for $x\in\mathscr{D}(A)$ will follow from the second part of the proof of (4).

(4): Denseness of $\mathscr{D}(A)$ follows from the first part of (3), since by (1) we have $\lim_{t\downarrow 0}\frac{1}{t}\int_0^t S(s)x\,ds = x$.

To prove that $A$ is closed we must check that the graph $\mathscr{G}(A)=\{(x,Ax): x\in\mathscr{D}(A)\}$ is closed in $E\times E$. Suppose that $(x_n)_{n=1}^\infty$ is a sequence in $\mathscr{D}(A)$ such that $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} Ax_n = y$ in $E$. We must show that $x\in\mathscr{D}(A)$ and $Ax=y$. Using that $\lim_{t\downarrow 0}\frac{1}{t}(S(t)-I)S(s)x_n = S(s)Ax_n$ uniformly for $s\in[0,h]$, we obtain

$$
\begin{aligned}
\frac{1}{h}(S(h)x-x) &= \lim_{n\to\infty}\frac{1}{h}(S(h)x_n - x_n)\\
&= \lim_{n\to\infty}\frac{1}{h}\Big(A\int_0^h S(s)x_n\,ds\Big)\\
&= \lim_{n\to\infty}\frac{1}{h}\lim_{t\downarrow 0}\frac{1}{t}(S(t)-I)\int_0^h S(s)x_n\,ds\\
&= \lim_{n\to\infty}\frac{1}{h}\lim_{t\downarrow 0}\int_0^h \frac{1}{t}(S(t)-I)S(s)x_n\,ds\\
&= \lim_{n\to\infty}\frac{1}{h}\int_0^h S(s)Ax_n\,ds\\
&= \frac{1}{h}\int_0^h S(s)y\,ds.
\end{aligned}
$$

Passing to the limit for $h\downarrow 0$ this gives $x\in\mathscr{D}(A)$ and $Ax=y$. The above identities also prove the second part of (3).

(5): This follows from (1), (2), and the definition of $A$. □

In hindsight, the second part of (3) is a special case of Hille's theorem. However, our proof of the closedness of $A$ already gave the result in this particular case.

**Definition 7.5.** *A* classical solution *of (ACP) is a continuous function $u:[0,T]\to E$ which belongs to $C^1((0,T];E)\cap C((0,T];\mathscr{D}(A))$ and satisfies $u(0)=x$ and $u'(t)=Au(t)$ for all $t\in(0,T]$.*

Here $\mathscr{D}(A)$ is regarded as a Banach space endowed with the graph norm.

**Corollary 7.6.** *For initial values $x\in\mathscr{D}(A)$ the problem (ACP) has a unique classical solution, which is given by $u(t)=S(t)x$.*

*Proof.* Part (5) of the proposition proves that $t \mapsto u(t) = S(t)x$ is a classical solution. Suppose that $t \mapsto v(t)$ is another classical solution. It is easy to check that the function $s \mapsto S(t-s)v(s)$ is continuous on $[0,t]$ and continuously differentiable on $(0,t)$ with derivative

$$\frac{d}{ds}S(t-s)v(s) = -AS(t-s)v(s) + S(t-s)v'(s) = 0$$

where we used that $v$ is a classical solution. Thus, $s \mapsto S(t-s)v(s)$ is constant on every interval $[0,t]$. Since $v(0) = x$ it follows that $v(t) = S(t-t)v(t) = S(t-0)v(0) = S(t)x = u(t)$. □

Note that for $x \in \mathscr{D}(A)$ the orbit $t \mapsto S(t)x$ even belongs to $C^1([0,T];E) \cap C([0,T];\mathscr{D}(A))$. The reason for defining classical solutions as we did above is that there exist important classes of $C_0$-semigroups which have the property that $t \mapsto S(t)x$ is a classical solution not only for $x \in \mathscr{D}(A)$, but for all $x \in E$. An example is the class of analytic $C_0$-semigroups which will be studied later on in this course.

**Definition 7.7.** *Let $T$ be a linear operator with domain $\mathscr{D}(T)$ on a complex Banach space $E$. The* resolvent set *of $T$ is the set $\varrho(T)$ consisting of all $\lambda \in \mathbb{C}$ for which there exists a (necessarily unique) bounded linear operator $R(\lambda, T)$ on $E$ such that*

(i) *$R(\lambda,T)(\lambda - T)x = x$ for all $x \in \mathscr{D}(T)$;*
(ii) *$R(\lambda,T)x \in \mathscr{D}(T)$ and $(\lambda - T)R(\lambda,T)x = x$ for all $x \in E$.*

*The* spectrum *of $T$ is the complement $\sigma(T) := \mathbb{C} \setminus \varrho(T)$.*

We call $R(\lambda,T) = (\lambda - T)^{-1}$ the *resolvent* of $T$ at $\lambda$. It is routine to check the *resolvent identity*: for all $\lambda_1, \lambda_2 \in \varrho(T)$ we have

$$R(\lambda_1,T) - R(\lambda_2,T) = (\lambda_2 - \lambda_1)R(\lambda_1,T)R(\lambda_2,T).$$

When $T$ is an operator on a real Banach space we put $\varrho(T) := \varrho(T_{\mathbb{C}})$ and $\sigma(T) := \sigma(T_{\mathbb{C}})$, where $T_{\mathbb{C}}$ is the complexification of $T$ (see Exercise 1).

In the next two lemmas, $A$ is the generator of a $C_0$-semigroup $S$ on a Banach space $E$ (in the case of a real Banach space, all formulas involving complex numbers should be interpreted in terms of complexifications). We fix constants $M \geqslant 1$ and $\mu \in \mathbb{R}$ such that $\|S(t)\| \leqslant Me^{\mu t}$ for all $t \geqslant 0$.

**Proposition 7.8.** *We have $\{\lambda \in \mathbb{C} : \ \mathrm{Re}\lambda > \mu\} \subseteq \varrho(A)$ and on this set the resolvent of $A$ is given by*

$$R(\lambda,A)x = \int_0^\infty e^{-\lambda t}S(t)x\,dt, \quad x \in E.$$

*As a consequence, for $\mathrm{Re}\lambda > \mu$ we have*

$$\|R(\lambda,A)\| \leqslant \frac{M}{\mathrm{Re}\lambda - \mu}.$$

*Proof.* Fix $x \in E$ and define $R_\lambda x := \int_0^\infty e^{-\lambda t} S(t)x\,dt$. From a straightforward computation using the semigroup property we obtain the identity

$$\lim_{h \downarrow 0} \frac{1}{h}(S(h) - I)R_\lambda x = \lambda R_\lambda x - x$$

from which it follows that $R_\lambda x \in \mathscr{D}(A)$ and $AR_\lambda x = \lambda R_\lambda x - x$. This shows that the bounded operator $R_\lambda$ is a right inverse for $\lambda - A$.

Integrating by parts and using that $\frac{d}{dt}S(t)x = S(t)Ax$ for $x \in \mathscr{D}(A)$ we obtain

$$\lambda \int_0^T e^{-\lambda t} S(t)x\,dt = -e^{-\lambda T} S(T)x + x + \int_0^T e^{-\lambda t} S(t)Ax\,dt.$$

Since $\mathrm{Re}\lambda > \mu$, sending $T \to \infty$ gives $\lambda R_\lambda x = x + R_\lambda Ax$. This shows that $R_\lambda$ is also a left inverse. □

**Lemma 7.9.** *For all $x \in E$ we have $\lim_{\lambda \to \infty} \lambda R(\lambda, A)x = x$.*

*Proof.* First we prove this for $x \in \mathscr{D}(A)$ by using the resolvent identity. Pick $\lambda' > \mu$. Writing $(\lambda' - A)x =: y$ we have

$$\lambda R(\lambda, A)x - x = \frac{\lambda}{\lambda - \lambda'}(R(\lambda', A)y - R(\lambda, A)y) - R(\lambda', A)y$$

Passing to the limit $\lambda \to \infty$ the right hand side tends to 0. This gives the result for $x \in \mathscr{D}(A)$. By the estimate of Proposition 7.8, the operators $\lambda R(\lambda, A)$ are uniformly bounded for $\lambda \geqslant \mu_0 > \mu$. Therefore the result for $x \in E$ follows by density. □

This lemma self-improves to $\lim_{\lambda \to \infty} \lambda^n R(\lambda, A)^n x = x$, which shows that $\mathscr{D}(A^n)$ is dense in $E$ for all $n \geqslant 1$.

## 7.2 Duality

For the discussion of the inhomogeneous Cauchy problem in the next section we need some preliminary material on duality of densely defined linear operators.

Let $E_1$ and $E_2$ be Banach spaces. To keep track of domains it will be useful to *define* a linear operator $A$ with domain $\mathscr{D}(A)$ from $E_1$ to $E_2$ as a pair $(A, \mathscr{D}(A))$, where $\mathscr{D}(A)$ is a linear subspace of $E_1$ and $A : \mathscr{D}(A) \to E_2$ is a linear mapping. If $(A, \mathscr{D}(A))$ is densely defined, that is, if $\mathscr{D}(A)$ is dense in $E_1$, we may define a linear operator $(A^*, \mathscr{D}(A^*))$ from $E_2^*$ to $E_1^*$ in the following way. Define $\mathscr{D}(A^*)$ to be the set of all $x_2^* \in E_2^*$ with the property that there exists an element $x_1^* \in E_1^*$ such that

$$\langle x, x_1^* \rangle = \langle Ax, x_2^* \rangle, \qquad \forall x \in \mathscr{D}(A).$$

Since $\mathscr{D}(A)$ is dense in $E_1$, the element $x_1^* \in E_1^*$ (if it exists) is unique and we set

$$A^* x_2^* := x_1^*, \quad x_2^* \in \mathscr{D}(A^*).$$

**Definition 7.10.** *Let $(A, \mathscr{D}(A))$ be a densely defined linear operator. The operator $(A^*, \mathscr{D}(A^*))$ is called the* adjoint *of $(A, \mathscr{D}(A))$.*

In order to discuss the properties of $A^*$ in a systematic way it is helpful to consider the topology on the dual space $E^*$ induced by the elements of a Banach space $E$, the so-called weak$^*$-topology.

**Definition 7.11.** *The* weak$^*$-topology *on $E^*$ is the topology generated by all sets of the form*

$$\{x^* \in E^* : \ |\langle x, y^* - x^* \rangle| < \varepsilon\}$$

*where $x \in E$, $y^* \in E^*$, and $\varepsilon > 0$.*

It is easily checked that the mappings $x^* \mapsto \langle x, y^* - x^* \rangle$ are continuous with respect to the weak$^*$-topology, and that the weak$^*$-topology is the coarsest topology on $E^*$ with this property.

**Lemma 7.12.** *Let $V$ be a non-empty subset of $E$. The* annihilator

$$V^\perp := \{x^* \in E^* : \ \langle v, x^* \rangle = 0 \ \text{ for all } \ v \in V\}$$

*is weak$^*$-closed.*

*Proof.* Let $y^* \notin V^\perp$ be arbitrary. By assumption there exists $v \in V$ such that $\langle v, y^* \rangle \neq 0$. The set

$$U := \Big\{x^* \in E^* : \ |\langle v, y^* - x^* \rangle| < \frac{1}{2} |\langle v, y^* \rangle| \Big\}$$

is weak$^*$-open, contains $y^*$, and is disjoint from $V^\perp$. It follows that the complement of $V^\perp$ is weak$^*$-open. ☐

It is an exercise in linear algebra to check that a linear subspace $F$ of $E^*$ is weak$^*$-dense if and only if it separates the points of $E$, that is, whenever $x \neq y$ in $E$ there is an $x^* \in F$ such that $\langle x, x^* \rangle \neq \langle y, x^* \rangle$. This fact is not really needed however. Whenever we say that a subspace $F$ of $E^*$ is weak$^*$-dense, what we shall actually use is that $F$ separates the points of $E$ and all formulations could be adapted accordingly.

**Proposition 7.13.** *Let $E_1$ and $E_2$ be Banach spaces and let $(A, \mathscr{D}(A))$ be a densely defined linear operator from $E_1$ to $E_2$.*

(1) *The adjoint $(A^*, \mathscr{D}(A^*))$ is weak$^*$-closed from $E_2^*$ to $E_1^*$, that is, the graph of $A^*$ is weak$^*$-closed in $E_2^* \times E_1^*$.*
(2) *If $(A, \mathscr{D}(A))$ is also closed, then $(A^*, \mathscr{D}(A^*))$ is weak$^*$-densely defined, that is, the domain of $A^*$ is weak$^*$-dense in $E_2^*$.*

*Proof.* We start with the preliminary remark that if $E$ and $F$ are Banach spaces, then the pairing

$$\langle (x, y), (x^*, y^*) \rangle := \langle x, x^* \rangle + \langle y, y^* \rangle$$

allows us to identify $E^* \times F^*$ with the dual of $E \times F$.

(1): Let $\mathscr{G}(A^*) = \{(x_1^*, A^* x_1^*) : x_1^* \in \mathscr{D}(A^*)\}$ be the graph of $A^*$ in $E_2^* \times E_1^*$. By definition of $\mathscr{D}(A^*)$ we have $(x_2^*, x_1^*) \in \mathscr{G}(A^*)$ if and only if

$$\langle (-Ax_1, x_1), (x_2^*, x_1^*) \rangle = 0, \quad \forall x_1 \in \mathscr{D}(A).$$

In other words, $\mathscr{G}(A^*)$ is the annihilator of $\rho(\mathscr{G}(A))$, where $\rho : E_1 \times E_2 \to E_2 \times E_1$ is defined by $\rho(x_1, x_2) = (-x_2, x_1)$. By Lemma 7.12, $\mathscr{G}(A^*)$ is weak$^*$-closed. This proves that $A^*$ is weak$^*$-closed.

(2): Now assume that $(A, \mathscr{D}(A))$ is also closed. We will show that $\mathscr{D}(A^*)$ separates the points of $E_2$. Suppose $x_2 \neq y_2$ in $E_2$. Then $(0, x_2 - y_2)$ is a non-zero element of $E_1 \times E_2$ which does not belong to $\mathscr{G}(A)$. Since $\mathscr{G}(A)$ is closed, by the Hahn-Banach theorem there exists an element $(x_1^*, x_2^*) \in (\mathscr{G}(A))^\perp$ such that

$$\langle (0, x_2 - y_2), (x_1^*, x_2^*) \rangle = \langle x_2 - y_2, x_2^* \rangle \neq 0.$$

To finish the proof we check that $x_2^* \in \mathscr{D}(A^*)$. For all $x_1 \in \mathscr{D}(A)$ we have $(x_1, Ax_1) \in \mathscr{G}(A)$ and therefore

$$0 = \langle (x_1, Ax_1), (x_1^*, x_2^*) \rangle = \langle x_1, x_1^* \rangle + \langle Ax_1, x_2^* \rangle.$$

But this means that $x_2^* \in \mathscr{D}(A^*)$ and $A^* x_2^* = -x_1^*$. □

The following simple result 'dualises' the definition of $\mathscr{D}(A^*)$.

**Proposition 7.14.** *Let $(A, \mathscr{D}(A))$ be a closed and densely defined linear operator from $E_1$ to $E_2$. If $x_1 \in E_1$ and $x_2 \in E_2$ are such that $\langle x_2, x_2^* \rangle = \langle x_1, A^* x_2^* \rangle$ for all $x_2^* \in \mathscr{D}(A^*)$, then $x_1 \in \mathscr{D}(A)$ and $Ax_1 = x_2$.*

*Proof.* We must prove that $(x_1, x_2) \in \mathscr{G}(A)$. Since $\mathscr{G}(A)$ is closed in $E_1 \times E_2$, by the Hahn-Banach theorem it suffices to check that $\langle (x_1, x_2), (x_1^*, x_2^*) \rangle = 0$ for all $(x_1^*, x_2^*) \in (\mathscr{G}(A))^\perp$.

Fix an arbitrary $(x_1^*, x_2^*) \in (\mathscr{G}(A))^\perp$. As in the second part of the previous proof we have $x_2^* \in \mathscr{D}(A^*)$ and $A^* x_2^* = -x_1^*$. Hence,

$$\langle (x_1, x_2), (x_1^*, x_2^*) \rangle = \langle x_1, -A^* x_2^* \rangle + \langle x_2, x_2^* \rangle = 0. \qquad \square$$

## 7.3 The abstract Cauchy problem

We now take a look at the inhomogeneous abstract Cauchy problem

$$\begin{cases} u'(t) = Au(t) + f(t), & t \in [0,T], \\ u(0) = x \end{cases} \tag{IACP}$$

with initial value $x \in E$. We assume that $A$ generates a $C_0$-semigroup $S$ on $E$ and take $f \in L^1(0,T;E)$.

Adapting the notion of a classical solution to the problem (IACP) leads to the so-called problem of maximal regularity. Instead of going into this, we refer to the Notes for more information and introduce here two alternative notions of solutions in terms of the integrated equation.

**Definition 7.15.** *A* strong solution *of* (IACP) *is a function $u \in L^1(0,T;E)$ such that for all $t \in [0,T]$ we have $\int_0^t u(s)\,ds \in \mathscr{D}(A)$ and*

$$u(t) = x + A\int_0^t u(s)\,ds + \int_0^t f(s)\,ds.$$

*A* weak solution *of* (IACP) *is a function $u \in L^1(0,T;E)$ such that for all $t \in [0,T]$ and $x^* \in \mathscr{D}(A^*)$ we have*

$$\langle u(t), x^*\rangle = \langle x, x^*\rangle + \int_0^t \langle u(s), A^*x^*\rangle\,ds + \int_0^t \langle f(s), x^*\rangle\,ds.$$

As an immediate consequence of Proposition 7.14 we make the following observation:

**Proposition 7.16.** *Every weak solution of* (IACP) *is a strong solution.*

Of course the converse holds trivially. We proceed with an existence and uniqueness result for strong solutions of (IACP).

**Theorem 7.17.** *For all $x \in E$ and $f \in L^1(0,T;E)$ the problem* (IACP) *admits a unique strong solution $u$, which is given by the convolution formula*

$$u(t) = S(t)x + \int_0^t S(t-s)f(s)\,ds. \tag{7.2}$$

*If $f \in L^p(0,T;E)$ with $1 \leqslant p < \infty$, then $u \in L^p(0,T;E)$.*

*Proof.* For the existence part, by Proposition 7.16 it suffices to show that (IACP) admits a weak solution. It is an easy consequence of Proposition 7.4 (3) that $u$ is a weak solution corresponding to the initial value $x$ if and only if $t \mapsto u(t) - S(t)x$ is a weak solution corresponding to the initial value 0. Therefore, without loss of generality we may assume that $x = 0$.

Let $u$ be given by (7.2). Then $u \in L^1(0,T;E)$; if $f \in L^p(0,T;E)$, then $u \in L^p(0,T;E)$. By Fubini's theorem and Proposition 7.4 (3), for all $t \in [0,T]$ and $x^* \in \mathscr{D}(A^*)$ we have

$$\begin{aligned}\int_0^t \langle u(s), A^*x^*\rangle\,ds &= \int_0^t \int_0^s \langle f(r), S^*(s-r)A^*x^*\rangle\,dr\,ds\\ &= \int_0^t \int_r^t \langle f(r), S^*(s-r)A^*x^*\rangle\,ds\,dr\\ &= \int_0^t \langle f(r), S^*(t-r)x^* - x^*\rangle\,dr\\ &= \langle u(t), x^*\rangle - \int_0^t \langle f(r), x^*\rangle\,dr.\end{aligned}$$

To prove uniqueness, suppose that $u$ and $\widetilde{u}$ are strong solutions of (IACP). Then $v := u - \widetilde{u}$ is integrable and satisfies $v(t) = A\int_0^t v(s)\,ds$ for all $t \in [0,T]$. Put

$$w(t) := \int_0^t \int_0^s v(r)\,dr\,ds.$$

By the fundamental theorem of calculus, $w$ is continuously differentiable on $[0,T]$, and using Hille's theorem we see that $w(t) \in \mathscr{D}(A)$ and

$$w'(t) = \int_0^t v(s)\,ds = \int_0^t A\int_0^s v(r)\,dr\,ds = Aw(t).$$

Fix $t \in [0,T]$ and put $g(s) := S(t-s)w(s)$. Then $g$ is continuously differentiable on $[0,t]$ with derivative

$$g'(s) = -AS(t-s)w(s) + S(t-s)w'(s) = 0.$$

It follows that $g$ is constant on $[0,t]$. Hence

$$w(t) = g(t) = g(0) = S(t)w(0) = 0.$$

We have shown that $\int_0^t \int_0^s v(r)\,dr\,ds = 0$ for all $t \in [0,T]$. It follows that $v = 0$ almost everywhere. □

## 7.4 Examples of $C_0$-semigroups

In this section we collect, without proofs, a number of important examples of $C_0$-semigroups. We encourage the reader to formulate the corresponding initial value problems; cf. Example 7.1. References to the literature are given in the Notes.

*Example 7.18 (Multiplication semigroup).* Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space and let $m : A \to \mathbb{R}$ be $\mu$-measurable. If $\operatorname{ess\,sup}_{\xi\in A} f(\xi) < \infty$, then the formula

$$S(t)f(\xi) := e^{tm(\xi)}f(\xi)$$

defines a $C_0$-semigroup on $L^p(A)$ for $1 \leqslant p < \infty$. The domain of its generator $A$ consists of all $f \in L^p(A)$ such that $mf \in L^p(A)$, and for $f \in \mathscr{D}(A)$ we have $Af = mf$.

*Example 7.19 (Translation semigroup).* On the space $L^p(\mathbb{R}_+)$, $1 \leqslant p < \infty$, the formula

$$(S(t)f)(\xi) := f(\xi + t)$$

defines a $C_0$-semigroup $S$. The domain of its generator $A$ consists of all $f \in L^p(\mathbb{R})$ whose weak derivative $f'$ exists and belongs to $L^p(\mathbb{R})$, and for $f \in \mathscr{D}(A)$ we have $Af = f'$.

These two examples represent perhaps the simplest constructions of $C_0$-semigroups and can be extended in various ways. We continue with two examples involving the Laplace operator.

*Example 7.20 (Heat semigroup).* On $L^p(\mathbb{R}^d)$, $1 \leqslant p < \infty$, the formula

$$(S(t)f)(\xi) := \frac{1}{\sqrt{(4\pi t)^n}} \int_{\mathbb{R}^d} f(\eta) \exp\Big(-\frac{|\xi - \eta|^2}{4t}\Big)\, d\eta$$

defines a $C_0$-semigroup. Its generator $A$ is given by $\mathscr{D}(A) = W^{2,p}(\mathbb{R}^d)$ and $Af = \Delta f$.

*Example 7.21 (Heat semigroup on bounded domains with Dirichlet boundary conditions).* Let $D$ be a bounded domain in $\mathbb{R}^d$ with $C^2$-boundary $\partial D$. On the space $L^p(D)$ with $1 \leqslant p < \infty$, the *Dirichlet Laplacian* is the operator $A$ defined by

$$\begin{aligned} \mathscr{D}(A) &:= W^{2,p}(D) \cap W_0^{1,p}(D), \\ Af &:= \Delta f \quad \text{for} \quad f \in \mathscr{D}(A). \end{aligned}$$

See Example 7.1. This operator is the generator of a $C_0$-semigroup on $L^p(D)$.

The previous two examples admit far-reaching generalisations to more general second order elliptic operators, and also different kinds of boundary conditions can be allowed.

We continue with two examples of operators generating a *$C_0$-group*. These are defined in the same way as $C_0$-semigroups, except that the index set is now the whole real line.

*Example 7.22 (Wave group).* On the space $W^{1,2}(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ we consider the operator $A$ defined by

$$\begin{aligned} \mathscr{D}(A) &:= W^{2,2}(\mathbb{R}^d) \times W^{1,2}(\mathbb{R}^d), \\ A(f_1, f_2) &:= (f_2, \Delta f_1) \quad \text{for} \quad (f_1, f_2) \in \mathscr{D}(A). \end{aligned}$$

This operator is the generator of a $C_0$-group on $W^{1,2}(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ which is associated with the wave equation $u''(t) = \Delta u$, written as a system $u' = v$, $v' = \Delta u$.

*Example 7.23 (Unitary $C_0$-groups on Hilbert spaces).* If $A$ is a self-adjoint operator on a complex Hilbert space $H$, then $iA$ is the generator of a $C_0$-group $S$ of unitary operators on $H$. This classical result of Stone is of fundamental importance in quantum mechanics. By the spectral theorem for self-adjoint operators, this example can be viewed as a special case of Example 7.18.

## 7.5 Exercises

1. Suppose that $E$ is a real Banach space. The product $E \times E$ can be given the structure of a complex vector space by introducing a complex scalar multiplication as follows:
$$(a+ib)(x,y) := (ax - by, bx + ay).$$
The idea is, of course, to think of the pair $(x,y) \in E \times E$ as if it were $x + iy$. The resulting complex vector space is denoted by $E_{\mathbb{C}}$.
   a) Prove that the formula
$$\|(x,y)\| := \sup_{\theta\in[0,2\pi]} \|(\cos\theta)x + (\sin\theta)y\|$$
defines a norm on $E_{\mathbb{C}}$ which turns $E_{\mathbb{C}}$ into a complex Banach space.
   b) Check that this norm on $E_{\mathbb{C}}$ extends the norm of $E$ in the sense that for all $x \in E$,
$$\|(x,0)\| = \|(0,x)\| = \|x\|.$$
   c) Check that for all $x, y \in E$ we have $\|(x,y)\| = \|(x,-y)\|$.
   d) Show that if $T$ is a (real-)linear bounded operator on $E$, then $T$ extends to a bounded (complex-)linear operator $T_{\mathbb{C}}$ on $E_{\mathbb{C}}$ by putting
$$T_{\mathbb{C}}(x,y) := (Tx, Ty),$$
and check that $\|T_{\mathbb{C}}\| = \|T\|$.

   A norm on $E_{\mathbb{C}}$ with the properties b), c), d) is called a *complexification* of the norm of $E$. The norm introduced in a) is by no means the unique complexification of the norm of $E$, and in concrete examples there is often a more natural choice.
   e) Show that any two complex norms on $E_{\mathbb{C}}$ which satisfy b) and c) are equivalent.

   By e), the spectrum of $T_{\mathbb{C}}$ is independent of the particular complexification chosen.

2. In this exercise we prove some properties of resolvents. We assume that $(T, \mathscr{D}(T))$ is a linear operator from $E$ to $E$ with resolvent set $\varrho(T)$.
   a) Prove that if $\varrho(T) \neq \varnothing$, then $T$ is closed.
   b) Prove the resolvent identity: for all $\lambda_1, \lambda_2 \in \varrho(T)$ we have
$$R(\lambda_1, T) - R(\lambda_2, T) = (\lambda_2 - \lambda_1) R(\lambda_1, T) R(\lambda_2, T).$$
   c) Prove that $\varrho(T)$ is an open subset of $\mathbb{C}$.
   d) Prove that
$$\lim_{\lambda\to\mu} \frac{R(\lambda,T) - R(\mu,T)}{\lambda - \mu} = -R(\mu,T)^2$$
with convergence in the operator norm.

e) Prove that if $T$ is closed and densely defined, then $\varrho(T^*) = \varrho(T)$ and

$$R(\lambda, T^*) = R(\lambda, T)^*, \quad \lambda \in \varrho(T) = \varrho(T^*).$$

f) Show that every closed subset of $\mathbb{C}$ is the spectrum of a suitable closed operator $T$.

3. Let $S$ be a $C_0$-semigroup on $E$ which is *uniformly bounded*, that is, $\sup_{t\geqslant 0} \|S(t)\| < \infty$. We show that there exists an equivalent norm $|\!|\!|\cdot|\!|\!|$ on $E$ such that $S$ is a *contraction semigroup* with respect to $|\!|\!|\cdot|\!|\!|$, that is, $|\!|\!|S(t)|\!|\!| \leqslant 1$ for all $t \geqslant 0$.
   a) Show that $|\!|\!|x|\!|\!| := \sup_{t\geqslant 0} \|S(t)x\|$ defines an equivalent norm on $E$.
   b) Show that $S$ is a contraction semigroup with respect to $|\!|\!|\cdot|\!|\!|$.

4. Let $S$ be a $C_0$-semigroup on $E$ with generator $A$, and suppose that $\|S(t)\| \leqslant Me^{\mu t}$ for all $t \geqslant 0$. Prove that

$$\|(R(\lambda, A))^k\| \leqslant M/(\mathrm{Re}\lambda - \mu)^k, \quad \mathrm{Re}\lambda > \mu,\ k = 1, 2, \dots.$$

*Hint:* By considering $A - \mu$ instead of $A$ we may assume that $\mu = 0$. In that situation observe that $|\!|\!|R(\lambda, A)|\!|\!| \leqslant 1/\mathrm{Re}\lambda$.

*Remark:* A celebrated theorem of Hille and Yosida asserts that the converse holds as well. We refer to the Notes for more information.

5. Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space. Suppose $f : A \to E^*$ is a function such that $\xi \mapsto \langle x, f(\xi)\rangle$ belongs to $L^1(A)$ for all $x \in E$.
   a) Show that the map $S : E \to L^1(A)$ defined by $Sx := \langle x, f\rangle$ is closed.
   b) Conclude from this that the formula

$$\langle x, x^*\rangle := \int_A \langle x, f\rangle \, d\mu$$

defines a bounded linear functional $x^* \in E^*$.

The functional $x^*$ is called the *weak$^*$-integral* of $f$ with respect to $\mu$, notation:

$$x^* =: \mathrm{weak}^* \int_A f \, d\mu.$$

   c) Show that the weak$^*$-integral commutes with adjoints of bounded operators on $E$.
   d) Show that if $f$ is an $E^*$-valued Bochner integrable function, then the Bochner integral and the weak$^*$-integral of $f$ agree.

Now suppose that $A$ generates a $C_0$-semigroup on $E$ and put $S^*(t) := (S(t))^*$ for $t \geqslant 0$.
   e) Prove the following dual version of the identities in Proposition 7.4 (3): for all $x^* \in E^*$ and $t \geqslant 0$ we have $\mathrm{weak}^* \int_0^t S^*(s)x^* \, ds \in \mathscr{D}(A^*)$ and

$$A^*\Big(\mathrm{weak}^* \int_0^t S^*(s)x^* \, ds\Big) = S^*(t)x^* - x^*.$$

If $x^* \in \mathscr{D}(A^*)$, then both sides are equal to $\mathrm{weak}^* \int_0^t S^*(s)A^*x^* \, ds$.

**Notes.** Excellent recent introductions to the theory of $C_0$-semigroups include the monographs by Arendt, Batty, Hieber, Neubrander [3], Davies [29], Engel and Nagel [38], Goldstein [41], Pazy [89]. For a discussion of the examples in Section 7.4 we refer to these sources. Their monumental 1957 treatise of Hille and Phillips[48] is freely available on-line (`http://www.ams.org/online_bks/coll31/`).

Due to limitations of space and time we have chosen not to discuss the two basic generation theorems of semigroup theorem. The first of these, the *Hille-Yosida theorem*, reads as follows.

**Theorem 7.24 (Hille-Yosida theorem).** *For a densely defined operator $A$ on a Banach space $E$ and constants $M \geqslant 1$ and $\mu \in \mathbb{R}$, the following assertions are equivalent:*

(1) *$A$ generates a $C_0$-semigroup on $E$ satisfying $\|S(t)\| \leqslant Me^{\mu t}$ for all $t \geqslant 0$;*
(2) *$\{\lambda \in \mathbb{C} :\ \lambda > \mu\} \subseteq \varrho(A)$ and $\|(R(\lambda, A))^k\| \leqslant M/(\mathrm{Re}\lambda - \mu)^k$ for all $\lambda > \mu$ and $k = 1, 2, \dots$;*
(3) *$\{\lambda \in \mathbb{C} :\ \mathrm{Re}\lambda > \mu\} \subseteq \varrho(A)$ and $\|(R(\lambda, A))^k\| \leqslant M/(\mathrm{Re}\lambda - \mu)^k$ for all $\mathrm{Re}\lambda > \mu$ and $k = 1, 2, \dots$.*

For $C_0$-contraction semigroups, Theorem 7.24 was obtained independently and simultaneously by Hille [47] and Yosida [111]; the extension to arbitrary $C_0$-semigroups is due to Feller, Miyadera, Phillips. The easy implication (1)⇒(3) has been discussed in Exercise 4 and (3)⇒(2) is trivial; the difficult implication is (2)⇒(1).

In order to state the second generation theorem, the *Lumer-Phillips theorem*, for $x \in E$ define $\partial(x) := \{x^* \in E^* :\ \|x^*\| = \|x\|,\ \langle x, x^*\rangle = \|x\|\|x^*\|\}$. By the Hahn-Banach theorem, $\partial(x) \neq \varnothing$.

**Theorem 7.25 (Lumer-Phillips theorem).** *For a densely defined operator $A$ on a Banach space $E$ with $\varrho(A) \cap (0, \infty) \neq \varnothing$ the following assertions are equivalent:*

(1) *$A$ generates a $C_0$-contraction semigroup on $E$;*
(2) *For all $x \in \mathscr{D}(A)$ and $\lambda > 0$ we have $\|(\lambda - A)x\| \geqslant \lambda\|x\|$;*
(2) *For all $x \in \mathscr{D}(A)$ and all $x^* \in \partial(x)$ we have $\mathrm{Re}\langle Ax, x^*\rangle \leqslant 0$;*
(3) *For all $x \in \mathscr{D}(A)$ there exists $x^* \in \partial(x)$ such that $\mathrm{Re}\langle Ax, x^*\rangle \leqslant 0$.*

This theorem, as its name suggests, is due to Lumer and Phillips [71]. We shall return to it later in the context of analytic $C_0$-semigroups. A detailed account of Theorems 7.24 and 7.25 and their history is given in [38].

The terminology for the various notions of solutions is not entirely standard. Ours is suggested by that of Da Prato and Zabczyk [27] for solutions of stochastic evolution equations.

The results of Section 7.2 can be found in any introductory text on functional analysis.

Theorem 7.17 is due to Ball [5], who also proved the following converse: if (IACP) admits a unique weak solution for all $f \in L^1(0,T;E)$ and initial values $x \in E$, then $A$ is the generator of a $C_0$-semigroup on $E$.

The convolution formula (7.2) is often taken as the definition of a *mild solution.* Typical questions then revolve around proving regularity properties of mild solutions in terms of properties of the forcing function $f$ and the semigroup $S$. We refer to [89, Chapter 4] for some elementary results in this direction. For the treatment of certain classes of non-linear Cauchy problems it is of particular importance to know whether the mild solutions have *maximal $L^p$-regularity*, meaning that for all $f \in L^p(0,T;E)$ the solution $u$ belongs to $W^{1,p}(0,T;E) \cap L^p(0,T;\mathscr{D}(A))$. A necessary condition for this is that $S$ be analytic; it is a classical result that this condition is also sufficient in Hilbert spaces. For analytic $C_0$-semigroups on Banach spaces the maximal regularity problem has recently be settled by Kalton and Lancien [57] (who gave a counterexample in $L^p$-spaces $E$) and Weis [108] (who obtained necessary and sufficient conditions for maximal $L^p$-regularity in UMD Banach spaces $E$). We refer to the lectures by Kunstmann and Weis [61] for a detailed account of this problem and its history, as well as a number of non-trivial examples.

A systematic discussion of complexifications is given in Muñoz, Sarantopoulos, Tonge [79]. The reader is warned that not every complex Banach space is the complexification of some underlying real Banach space. The first (non-constructive) proof of this fact was given by Bourgain [11]. An explicit counterexample was found subsequently by Kalton [56].

# 8

# Linear equations with additive noise I

Let $A$ be the generator of a $C_0$-semigroup $S$ on $E$. In the previous lecture we have seen that the inhomogeneous abstract Cauchy problem

$$u'(t) = Au(t) + f(t), \quad u(0) = x,$$

is solved by the convolution formula

$$u(t) = S(t)x + \int_0^t S(t-s)f(s)\,ds,$$

We now turn to the stochastic analogue of this equation,

$$\begin{cases} dU(t) = AU(t)\,dt + B\,dW_H, \\ \ \ u(0) = x, \end{cases}$$

where $W_H$ is an $H$-cylindrical Brownian motion, and $B \in \mathscr{L}(H,E)$ is a bounded operator. In concrete examples, $W_H$ models space-time white noise and $B$ 'injects' this noise into the state space $E$. Reasoning by analogy, this equation should be solved by the stochastic convolution

$$U(t) = S(t)x + \int_0^t S(t-s)B\,dW_H.$$

We shall see that this is indeed correct, provided the $\mathscr{L}(H,E)$-valued function $S(\cdot)B$ is stochastically integrable with respect to $W_H$.

## 8.1 Stochastic preliminaries

In this section we collect several results which are needed in the proof of the main result of this lecture, Theorem 8.6. We begin with an integrations by parts formula.

**Lemma 8.1 (Integration by parts).** *For all $\phi \in C^1[0,T]$ and $h \in H$, almost surely the following identity holds:*

$$\int_0^T \phi'(t) W_H(t)h\,dt = \phi(T)W_H(T)h - \int_0^T \phi \otimes h\,dW_H.$$

Before we prove the lemma we clarify the meaning of the integral on the left hand side. Recalling that $W_H h$ is a Brownian motion, using Corollary 6.10 we select a version of $W_H h$ whose trajectories are continuous almost surely. Then the integral on the left hand side is well defined almost surely as a Lebesgue integral.

*Proof.* We may assume that $\phi'(0) = 0$; this somewhat simplifies the calculations below.

We begin by noting that $\int_0^T \phi \otimes h\,dW_H = \int_0^T \phi\,dW_H h$. For step functions this is clear from the definitions and the general case follows by approximation. Rescaling $h$ to unit length, it is therefore enough to prove the almost sure identity

$$\int_0^T \phi'(t) W(t)\,dt = \phi(T)W(T) - \int_0^T \phi\,dW$$

for functions $\phi \in C^1[0,T]$, where $W$ is a scalar Brownian motion. This identity is a special case of Itô's formula, but for those readers who are not familiar with it we shall give a self-contained argument (which is indeed nothing but the proof of Itô's formula in the special case considered here). Let

$$g := \sum_{n=1}^N c_n 1_{(t_{n-1},t_n]}, \quad G := \sum_{n=1}^N \sum_{m=1}^n c_m (t_m - t_{m-1}) 1_{(t_{n-1},t_n]}$$

with $c_1, \dots, c_N$ scalars and $0 = t_0 < \dots < t_N = T$. Then, almost surely,

$$\int_0^T g(t)W(t)\,dt = \sum_{n=1}^N c_n \int_{t_{n-1}}^{t_n} W(t)\,dt$$

and

$$\begin{aligned}
G(T)&W(T) - \int_0^T G\,dW \\
&= \sum_{m=1}^N c_m(t_m - t_{m-1})W(T) - \sum_{n=1}^N \sum_{m=1}^n c_m(t_m - t_{m-1})(W(t_n) - W(t_{n-1})) \\
&= \sum_{m=1}^N c_m(t_m - t_{m-1})W(T) - \sum_{m=1}^N \sum_{n=m}^N c_m(t_m - t_{m-1})(W(t_n) - W(t_{n-1})) \\
&= \sum_{m=1}^N c_m(t_m - t_{m-1})W(T) - \sum_{m=1}^N c_m(t_m - t_{m-1})(W(T) - W(t_{m-1})) \\
&= \sum_{m=1}^N c_m(t_m - t_{m-1})W(t_{m-1}).
\end{aligned}$$

Now let $\phi \in C^1[0,T]$ be given and put $g_k := \sum_{n=1}^{N_k} \phi'(t_{k,n-1})1_{(t_{k,n-1},t_{k,n}]}$, assuming that $\lim_{k\to\infty} \sup_{1\leqslant n\leqslant N_k}(t_{k,n} - t_{k,n-1}) = 0$. Then $\lim_{k\to\infty} g_k = \phi'$ uniformly on $(0,T]$. Defining the functions $G_k$ in terms of the $g_k$ as above, we have $\lim_{k\to\infty} G_k = \phi$ uniformly on $(0,T]$. The above computation gives the following identity, which almost surely holds for all $k$:

$$\begin{aligned}
&\Big|\int_0^T g_k(t)W(t)\,dt - \Big(G_k(T)W(T) - \int_0^T G_k\,dW\Big)\Big| \\
&= \Big|\sum_{n=1}^{N_k} \phi'(t_{k,n-1})\int_{t_{k,n-1}}^{t_{k,n}} W(t)\,dt - \sum_{n=1}^{N_k} \phi'(t_{k,n-1})(t_{k,n} - t_{k,n-1})W(t_{k,n-1})\Big|.
\end{aligned}$$

As $k \to \infty$, the left hand side tends to $|\int_0^T \phi'(t)W(t)\,dt - (\phi(T)W(T) - \int_0^T \phi\,dW)|$ in $L^2(\Omega)$ and hence in measure, whereas the right hand side tends to 0 almost surely by path continuity. This proves the lemma. □

We continue with a Fubini theorem for interchanging a Bochner integral and a stochastic integral of an $H$-valued function. In this context it is natural to impose an integrability condition which is $L^1$ with respect to the variable of Bochner integration and $L^2$ with respect to the variable of stochastic integration.

**Lemma 8.2 (Stochastic Fubini theorem).** *Let $\phi : (0,T) \times (0,T) \to H$ be a strongly measurable function satisfying*

$$\int_0^T \Big(\int_0^T \|\phi(s,t)\|_H^2\,dt\Big)^{\frac{1}{2}}\,ds < \infty.$$

(1) *$t \mapsto \phi(s,t)$ belongs to $L^2(0,T;H)$ for almost all $s \in (0,T)$, and the $L^2(\Omega)$-valued function $s \mapsto \int_0^T \phi(s,t)\,dW_H(t)$ belongs to $L^1(0,T;L^2(\Omega))$;*
(2) *$s \mapsto \phi(s,t)$ belongs to $L^1(0,T;H)$ for almost all $t \in (0,T)$, and the $H$-valued function $t \mapsto \int_0^T \phi(s,t)\,ds$ belongs to $L^2(0,T;H)$;*
(3) *in $L^2(\Omega)$ we have*

$$\int_0^T \Big(\int_0^T \phi(s,t)\,dW_H(t)\Big)\,ds = \int_0^T \Big(\int_0^T \phi(s,t)\,ds\Big)\,dW_H(t).$$

*Proof.* (1): By assumption we have $\phi \in L^1(0,T;L^2(0,T;H))$, and therefore (1) is an immediate consequence of the Itô isometry (6.2).

(2): We claim that for a step function $\phi : (0,T) \times (0,T) \to H$ we have

$$\|\phi\|_{L^2(0,T;L^1(0,T;H))} \leqslant \|\phi\|_{L^1(0,T;L^2(0,T;H))}.$$

It suffices to prove this for $T = 1$. If $\phi = \sum_{j=1}^M \sum_{k=1}^N 1_{(s_{j-1},s_j)}1_{(t_{k-1},t_k)} \otimes h_{jk}$, then

$$
\begin{aligned}
\|\phi\|^2_{L^2(0,T;L^1(0,T;H))} &= \sum_{k=1}^{N}(t_k - t_{k-1})\Big(\sum_{j=1}^{M}(s_j - s_{j-1})\|h_{jk}\|\Big)^2 \\
&= \sum_{i=1}^{M}\sum_{j=1}^{M}(s_i - s_{i-1})(s_j - s_{j-1})\sum_{k=1}^{N}(t_k - t_{k-1})\|h_{ik}\|\|h_{jk}\|
\end{aligned}
$$

and similarly

$$
\begin{aligned}
\|\phi\|^2_{L^1(0,T;L^2(0,T;H))} &= \Big(\sum_{j=1}^{M}(s_j - s_{j-1})\Big(\sum_{k=1}^{N}(t_k - t_{k-1})\|h_{jk}\|^2\Big)^{\frac12}\Big)^2 \\
&= \sum_{i=1}^{M}\sum_{j=1}^{M}(s_i - s_{i-1})(s_j - s_{j-1}) \\
&\quad \times \Big(\sum_{k=1}^{N}(t_k - t_{k-1})\|h_{ik}\|^2\Big)^{\frac12}\Big(\sum_{k=1}^{N}(t_k - t_{k-1})\|h_{jk}\|^2\Big)^{\frac12}.
\end{aligned}
$$

In view of the Cauchy-Schwarz inequality, this proves the claim. It follows that the identity mapping on step functions extends to a continuous embedding of $L^1(0,T;L^2(0,T;H))$ into $L^2(0,T;L^1(0,T;H))$. This gives (2).

(3): For step functions $\phi$ the identity follows by a trivial computation, and its extension to functions $\phi \in L^1(0,T;L^2(0,T;H))$ is obtained by approximation using (1) and (2). $\square$

## 8.2 Semigroup preliminaries

Let $A$ be the generator of a $C_0$-semigroup $S$ on $E$. Define

$$E^\odot := \overline{\mathscr{D}(A^*)},$$

the closure being taken with respect to the norm topology of $E^*$. Note that $E^\odot$ is a closed and weak*-dense subspace of $E^*$. We let $A^\odot$ be the *part* of $A^*$ in $E^\odot$, that is,

$$
\begin{aligned}
\mathscr{D}(A^\odot) &:= \{x^* \in \mathscr{D}(A^*) : \ A^*x^* \in E^\odot\}, \\
A^\odot x^* &:= A^*x^*, \qquad x^* \in \mathscr{D}(A^\odot).
\end{aligned}
$$

**Proposition 8.3.** *Let $A$ be the generator of a $C_0$-semigroup $S$ on $E$. The adjoint semigroup $S^*$ restricts to a $C_0$-semigroup $S^\odot$ on $E^\odot$ whose generator equals $A^\odot$.*

*Proof.* For $t \in [0,T]$, $x \in E$, and $x^* \in \mathscr{D}(A^*)$ we have

$$|\langle x, S^*(t)x^* - x^*\rangle| \leqslant \int_0^t |\langle x, S^*(s)A^*x^*\rangle|\, ds \leqslant t\|x\| \cdot \sup_{s\in[0,T]}\|S(s)\| \cdot \|A^*x^*\|.$$

Taking the supremum over all $x \in E$ of norm $\|x\| \leqslant 1$ gives

$$\limsup_{t\downarrow 0} \|S^*(t)x^* - x^*\| \leqslant \lim_{t\downarrow 0}\ t \cdot \sup_{s\in[0,T]} \|S(s)\| \cdot \|A^*x^*\| = 0.$$

Since $\mathscr{D}(A^*)$ is invariant under $S^*$ (by duality we have $A^*S^*(t)x^* = S^*(t)A^*x^*$ for $x^* \in \mathscr{D}(A^*)$ and $t \geqslant 0$) and $S^*(t)$ is uniformly bounded on $[0,T]$, it follows that $S^*$ restricts to a $C_0$-semigroup $S^\odot$ on $E^\odot$.

Let $B$ denote the generator of $S^\odot$. If $x^\odot \in \mathscr{D}(B)$, then for all $x \in \mathscr{D}(A)$ we have

$$\langle x, Bx^\odot\rangle = \lim_{t\downarrow 0} \frac{1}{t}\langle x, S^\odot(t)x^\odot - x^\odot\rangle = \lim_{t\downarrow 0} \frac{1}{t}\langle S(t)x - x, x^\odot\rangle = \langle Ax, x^\odot\rangle.$$

Hence $x^\odot \in \mathscr{D}(A^*)$ and $A^*x^\odot = Bx^\odot$. Since $Bx^\odot \in E^\odot$ it follows that $x^\odot \in \mathscr{D}(A^\odot)$ and $A^\odot x^\odot = A^*x^\odot = Bx^\odot$. Conversely, if $x^\odot \in \mathscr{D}(A^\odot)$, then $A^\odot x^\odot \in E^\odot$ and $s \mapsto S^\odot(s)A^\odot x^\odot$ is strongly continuous and, for all $x \in E$,

$$\Big|\Big\langle x, A^\odot x^\odot - \frac{1}{t}(S^\odot(t)x^\odot - x^\odot)\Big\rangle\Big| \leqslant \|x\|\Big\|A^\odot x^\odot - \frac{1}{t}\int_0^t S^\odot(s)A^\odot x^\odot\,ds\Big\|.$$

Hence,

$$\Big\|A^\odot x^\odot - \frac{1}{t}(S^\odot(t)x^\odot - x^\odot)\Big\| \leqslant \Big\|A^\odot x^\odot - \frac{1}{t}\int_0^t S^\odot(s)A^\odot x^\odot\,ds\Big\|.$$

Since $A^\odot x^\odot \in E^\odot$, the right hand side tends to 0 as $t \downarrow 0$ by strong continuity. This proves that $x^\odot \in \mathscr{D}(B)$ and $Bx^\odot = A^\odot x^\odot$. □

This proposition will be used in combination with the next approximation result.

**Lemma 8.4.** *For $k = 0, 1, 2, \ldots,$ linear combinations of the functions $\phi \otimes x$ with $\phi \in C^k[0,T]$ and $x \in E$ are dense in $C^k([0,T];E)$.*

*Proof.* We begin with the case $k = 0$, which is proved by a standard partition of unity argument. Let $f \in C([0,T];E)$ be arbitrary. Let $\varepsilon > 0$. Since $f$ is uniformly continuous we may choose $\delta > 0$ such that $\|f(t) - f(s)\| < \varepsilon$ whenever $|t - s| < \delta$. Let $I_1, \ldots, I_N$ be open intervals of length $< \delta$ covering $[0,T]$ and let $\phi_1, \ldots, \phi_N$ be a partition of unity with respect to this cover, that is, $0 \leqslant \phi_n \leqslant 1$, $\phi_n$ is supported in $I_n$, and $\sum_{n=1}^N \phi_n = 1$. Choose points $t_n \in [0,T] \cap I_n$, let $x_n := f(t_n)$, and put $f_\varepsilon := \sum_{n=1}^N \phi_n \otimes x_n$. Fix $t \in [0,T]$. If $t \in I_n$, then $|t - t_n| < \delta$ and therefore $\|f(t) - x_n\| < \varepsilon$. If $t \notin I_n$, then $\phi_n(t) = 0$. Hence, using that $f = \sum_{n=1}^N \phi_n f$,

$$\|f(t) - f_\varepsilon(t)\| \leqslant \sum_{n=1}^N \phi_n(t)\|f(t) - x_n\| \leqslant \varepsilon \sum_{n:\ t\in I_n} \phi_n(t) \leqslant \varepsilon.$$

This proves that $\|f - f_\varepsilon\| \leqslant \varepsilon$.

The general case is proved with induction on $k$. Suppose the lemma has been proved for $k = 0, \dots, l$ and let $f \in C^{l+1}([0,T];E)$ be arbitrary. Then $f' \in C^l([0,T];E)$ and therefore we can find functions $g_j \in C^l([0,T];E)$ of the form $g_j = \sum_{n=1}^{N_j} \phi_{jn} \otimes x_{jn}$ with $\phi_{jn} \in C^l[0,T]$ and $x_{jn} \in E$ such that $\lim_{j\to\infty} g_j = f'$ in $C^l([0,T];E)$. Let $\psi_{jn}(t) := \int_0^t \phi_{jn}(s)\,ds$, put $x_0 := f(0)$, and set

$$f_j := 1 \otimes x_0 + \sum_{n=1}^{N_j} \psi_{jn} \otimes x_{jn}.$$

Then $\lim_{j\to\infty} f_j = f$ in $C([0,T];E)$ and $\lim_{n\to\infty} f_j' = f'$ in $C^l([0,T];E)$, so $\lim_{j\to\infty} f_j = f$ in $C^{l+1}([0,T];E)$. $\square$

## 8.3 Existence and uniqueness: cylindrical Brownian motion

We consider the stochastic abstract Cauchy problem

$$\begin{cases} dU(t) = AU(t)\,dt + B\,dW_H(t), & t \in [0,T], \\ U(0) = x. \end{cases} \tag{SACP}$$

Here $A$ is the generator of a $C_0$-semigroup $\{S(t)\}_{t\geqslant 0}$ on $E$, $W_H$ is an $H$-cylindrical Brownian motion on $(\Omega, \mathscr{F}, \mathbb{P})$, and $B \in \mathscr{L}(H,E)$ is a given bounded operator.

An $E$-valued process $\{U(t)\}_{t\in[0,T]}$ will be called *strongly measurable* if it has a version which is strongly $\mathscr{B}([0,T]) \times \mathscr{F}$-measurable on $[0,T] \times \Omega$.

**Definition 8.5.** *A* weak solution *of the problem* (SACP) *is an $E$-valued process $\{U^x(t)\}_{t\in[0,T]}$ which has a strongly measurable version with the following properties:*

(i) *almost surely, the paths $t \mapsto U^x(t)$ are integrable;*
(ii) *for all $t \in [0,T]$ and $x^* \in \mathscr{D}(A^*)$ we have, almost surely,*

$$\langle U^x(t), x^*\rangle = \langle x, x^*\rangle + \int_0^t \langle U^x(s), A^*x^*\rangle\,ds + W_H(t)B^*x^*.$$

In order not to overburden notations, we do not distinguish notationally the process $\{U^x(t)\}_{t\in[0,T]}$ from its version with the properties (i) and (ii).

**Theorem 8.6.** *The following assertions are equivalent:*

(1) *the problem* (SACP) *has a weak solution $\{U^x(t)\}_{t\in[0,T]}$;*
(2) *$t \mapsto S(t)B$ is stochastically integrable on $(0,T)$ with respect to $W_H$.*

*In this situation, for every $t \in (0,T)$ the function $s \mapsto S(t-s)B$ is stochastically integrable on $(0,t)$ with respect to $W_H$ and almost surely we have*

$$U^x(t) = S(t)x + \int_0^t S(t-s)B\, dW_H(s). \tag{8.1}$$

*Proof.* We start by noting that $t \mapsto U^x(t)$ is a weak solution corresponding to the initial value $x$ if and only if $t \mapsto U^x(t) - S(t)x$ is a weak solution corresponding to the initial value 0. Without loss of generality we shall therefore assume that $x = 0$ and write $U(t) := U^0(t)$ for convenience.

(1) ⇒ (2): We will show first that for all $t \in [0,T]$ and $x^* \in \mathscr{D}(A^{\odot 2})$, almost surely we have

$$\langle U(t), x^* \rangle = \int_0^t B^* S^*(t-s)x^*\, dW_H(s). \tag{8.2}$$

Fix $t \in [0,T]$ and $x^\odot \in \mathscr{D}(A^\odot)$. By Fubini's theorem, almost surely the identity

$$\langle U(s), x^\odot \rangle = \int_0^s \langle U(r), A^\odot x^\odot \rangle\, dr + W_H(s)B^* x^\odot \tag{8.3}$$

holds for almost all $s \in (0,t)$; here we use that both terms on the right hand side are jointly measurable on $(0,t) \times \Omega$. In combination with Lemma 8.1 this gives, for any $C^1$-function $\phi : [0,t] \to \mathbb{R}$,

$$\begin{aligned}
&\int_0^t \phi'(s)\langle U(s), x^\odot \rangle\, ds \\
&\quad = \int_0^t \phi'(s)\Big(\int_0^s \langle U(r), A^\odot x^\odot \rangle\, dr\Big)\, ds + \int_0^t \phi'(s) W_H(s)B^* x^\odot\, ds \\
&\quad = \phi(t)\int_0^t \langle U(s), A^\odot x^\odot \rangle\, ds - \int_0^t \phi(s)\langle U(s), A^\odot x^\odot \rangle\, ds \\
&\qquad\qquad + \phi(t)W_H(t)B^* x^\odot - \int_0^t \phi(s)B^* x^\odot\, dW_H(s)
\end{aligned}$$

almost surely. Multiplying both sides of (8.3) with $\phi(t)$, putting $f := \phi \otimes x^\odot$ and rewriting, we obtain

$$\langle U(t), f(t) \rangle = \int_0^t \langle U(s), f'(s) + A^\odot f(s) \rangle\, ds + \int_0^t B^* f(s)\, dW_H(s) \tag{8.4}$$

almost surely. By Lemma 8.4 applied to the Banach space $\mathscr{D}(A^\odot)$, this identity extends to arbitrary functions $f \in C^1([0,t]; \mathscr{D}(A^\odot))$. In particular we may take $f(s) = S^\odot(t-s)x^\odot$, with $x^\odot \in \mathscr{D}(A^{\odot 2})$. For this choice of $f$, the identity (8.4) reduces to (8.2).

So far we have proved that (8.2) holds for functionals $x^* \in \mathscr{D}(A^{\odot 2})$. We shall prove next that (8.2) holds for functionals $x^* \in E^*$. Then the stochastic integrability of $s \mapsto S(t-s)B$ on $(0,t)$ follows from Theorem 6.17.

The extension of (8.2) from functionals $x^* \in \mathscr{D}(A^{\odot 2})$ to functionals $x^* \in E^*$ is not entirely straightforward since in general $\mathscr{D}(A^{\odot 2})$ is only weak*-dense in $E^*$. Let $x^* \in E^*$ be arbitrary and fixed, and let weak*-$\lim_{n\to\infty} x_n^* = x^*$ with all $x_n^* \in \mathscr{D}(A^{\odot 2})$ (for instance, take $x_n^* = \lambda_n^3 R(\lambda_n, A^*)^3 x^*$ with suitable $\lambda_n \to \infty$). By dominated convergence, for all $f \in L^2(0,t;H)$ we have

$$\lim_{n\to\infty} [f, B^*S^*(t-\cdot)x_n^*]_{L^2(0,t;H)} = [f, B^*S^*(t-\cdot)x^*]_{L^2(0,t;H)}.$$

It follows that for all $N \geqslant 1$, $B^*S^*(t-\cdot)x^*$ belongs to the weak closure in $L^2(0,t;H)$ of the tail sequence $(B^*S^*(t-\cdot)x_n^*)_{n=N}^\infty$. By the Hahn-Banach theorem, $B^*S^*(t-\cdot)x^*$ belongs to the strong closure in $L^2(0,t;H)$ of the convex hull of this sequence. It follows that there exist vectors $y_N^*$, belonging to the convex hull of $(x_n^*)_{n=N}^\infty$, such that

$$\|B^*S^*(t-\cdot)y_N^* - B^*S^*(t-\cdot)x^*\|_{L^2(0,t;H)} < \frac{1}{N}.$$

The isometry (6.2) implies that

$$\lim_{N\to\infty} \int_0^t B^*S^*(t-s)y_N^* \, dW_H(s) = \int_0^t B^*S^*(t-s)x^* \, dW_H(s)$$

in $L^2(\Omega)$. By passing to a subsequence and using that weak*-$\lim_{N\to\infty} y_N^* = x^*$ (this follows from the fact that we used the tail sequence $(x_n^*)_{n=N}^\infty$ to define $y_N^*$), we obtain

$$\begin{aligned}\langle U(t), x^*\rangle = \lim_{j\to\infty} \langle U(t), y_{N_j}^*\rangle &= \lim_{j\to\infty} \int_0^t B^*S^*(t-s)y_{N_j}^* \, dW_H(s) \\ &= \int_0^t B^*S^*(t-s)x^* \, dW_H(s)\end{aligned}$$

almost surely.

(2) $\Rightarrow$ (1): Suppose now that the function $t \mapsto S(t)B$ is stochastically integrable on $(0,T)$. This implies the stochastic integrability of $s \mapsto S(t-s)B$ on $(0,t)$ for all $t \in (0,T]$. We check that the process $U$ defined by the convolution (8.1) with $x = 0$ has a strongly measurable version which is a weak solution of the problem (SACP) with initial value $x = 0$.

To prove that $U$ has a strongly measurable version we argue as follows. As in the proof of Step 1 of Theorem 6.17 (3)$\Rightarrow$(1) we may assume that $H$ is separable. Then by Proposition 5.14 the $\gamma(L^2(0,T;H),E)$-valued function $t \mapsto R_t$ is strongly measurable, where $R_t$ is the integral operator associated with $s \mapsto 1_{(0,t)}(s)S(t-s)B$. By covariance domination, $\|R_t\|_{\gamma(L^2(0,T;H),E)} \leqslant \|R_T\|_{\gamma(L^2(0,T;H),E)}$. Applying the Itô isometry of Theorem 6.14 we see that $U$ defines an element of $L^\infty(0,T;L^2(\Omega;E))$. The existence of a strongly measurable version follows from this (cf. Example 1.21).

Fix $x^* \in \mathscr{D}(A^*)$ and $t \in [0,T]$. Then almost surely

$$\langle U(t), A^* x^* \rangle = \int_0^t B^* S^*(t-s) A^* x^* \, dW_H(s).$$

By the stochastic Fubini theorem applied to $\phi(s,t) := 1_{\{0 \leqslant s \leqslant t \leqslant T\}} B^* S^*(t-s)x^*$, the $L^2(\Omega)$-valued function $t \mapsto \langle U(t), A^* x^* \rangle$ is integrable on $(0,T)$ and

$$\begin{aligned}
\int_0^t \langle U(s), A^* x^* \rangle \, ds &= \int_0^t \int_0^s B^* S^*(s-r) A^* x^* \, dW_H(r) \, ds \\
&= \int_0^t \int_r^t B^* S^*(s-r) A^* x^* \, ds \, dW_H(r) \\
&= \int_0^t B^* S^*(t-r) x^* - B^* x^* \, dW_H(r) \\
&= \langle U(t), x^* \rangle - W_H(t) B^* x^*,
\end{aligned}$$

where all identities are understood in the sense of $L^2(\Omega)$. In particular the identities hold almost surely.

It remains to check that the trajectories of $U$ are integrable almost surely. Let $\mu_t$ be the distribution of $U(t)$ and let $Q_t$ be its covariance operator. We have

$$\langle Q_t x^*, x^* \rangle = \int_0^t \|B^* S^*(s) x^*\|_H^2 \, ds \leqslant \langle Q_T x^*, x^* \rangle = \langle R x^*, x^* \rangle.$$

Hence by Fubini's theorem and covariance domination, for arbitrary but fixed $1 \leqslant p < \infty$ we obtain

$$\mathbb{E} \int_0^T \|U(t)\|^p \, dt = \int_0^T \int_E \|x\|^p \, d\mu_t(x) \, dt \leqslant T \int_E \|x\|^p \, d\mu_T(x) < \infty.$$

This implies that almost all trajectories $t \mapsto U(t,\omega)$ belong to $L^p(0,T;E)$. $\square$

Note that theorem 8.6 contains the following uniqueness assertion: if $U^x$ and $\widetilde{U}^x$ are both weak solutions of (SACP), then $U^x$ and $\widetilde{U}^x$ are versions of each other: both $U^x(t)$ and $\widetilde{U}^x(t)$ equal the right hand side of (8.1) almost surely. This justifies us to speak of 'the' solution of (SACP).

Comparing the proof of Theorem 8.6 with that of Theorem 7.17 we observe that the existence proofs are essentially identical, whereas the uniqueness parts are very different. The reason is that the exceptional sets in the definition of a weak solution of the stochastic problem (SACP) depend on $t$ and $x^*$, which prevents us from applying Proposition 7.14 almost surely. Because of this it is no longer clear whether a weak solution is always a strong solution (cf. Proposition 7.16).

## 8.4 Existence and uniqueness: Brownian motion

Next we consider the problem (SACP) under the assumption that $B \in \gamma(H,E)$. In this situation the term '$B\, dW_H$' may be replaced by '$dW^B$', where $W^B$ is an $E$-valued Brownian motion canonically associated with $B$.

**Definition 8.7.** *An $E$-valued process $(W(t))_{t\in[0,T]}$ is called an $E$-*valued Brownian motion *if it enjoys the following properties:*

i) $W(0) = 0$ *almost surely;*
ii) $W(t-s)$ *and* $W(t) - W(s)$ *are identically distributed Gaussian random variables for all* $0 \leqslant s \leqslant t \leqslant T$*;*
iii) $W(t) - W(s)$ *is independent of* $\{W(r) :\ 0 \leqslant r \leqslant s\}$ *for all* $0 \leqslant s \leqslant t \leqslant T$.

**Proposition 8.8.** *Let $(W_H(t))_{t\in[0,T]}$ be an $H$-cylindrical Brownian motion and let $B \in \gamma(H,E)$. If $(h_n)_{n=1}^\infty$ is an orthonormal basis of* $(\ker(B))^\perp$*, then:*

(1) *the sum*

$$W^B(t) := \sum_{n=1}^\infty W_H(t)h_n \otimes Bh_n$$

*converges almost surely and in $L^p(\Omega;E)$, $1 \leqslant p < \infty$, for all $t \in [0,T]$;*
(2) *up to a null set, $W^B(t)$ is independent of the choice of the basis $(h_n)_{n=1}^\infty$;*
(3) *the process $(W^B(t))_{t\in[0,T]}$ defines an $E$-valued Brownian motion.*

The proof involves a straightforward application of Theorem 5.15, noting that for $0 \leqslant s \leqslant t \leqslant T$ the covariance operator of $W^B(t) - W^B(s)$ equals $(t-s)BB^*$.

This proposition shows that for operators $B \in \gamma(H,E)$ the problem (SACP) may be restated as

$$\begin{cases} dU(t) = AU(t)\,dt + dW^B(t), & t \in [0,T], \\ U(0) = x. \end{cases}$$

In the converse direction, every $E$-valued Brownian motion is of the form $W^B$ for canonical choices of $H$ and $B \in \gamma(H,E)$ (Exercise 2).

**Definition 8.9.** *Let $B \in \gamma(H,E)$. A* strong solution *of* (SACP) *is a strongly measurable $E$-valued process $(U^x(t))_{t\in[0,T]}$ with the following properties:*

*i) the trajectories of $U^x$ are integrable almost surely;*
*ii) for all $t \in [0,T]$, almost surely we have $\int_0^t U^x(s)\,ds \in \mathscr{D}(A)$ and*

$$U^x(t) = x + A\int_0^t U^x(s)\,ds + W^B(t).$$

**Theorem 8.10.** *Let $B \in \gamma(H,E)$. The following assertions are equivalent:*

(1) *the problem* (SACP) *has a strong solution;*
(2) *the problem* (SACP) *has a weak solution.*

*In this situation, the weak and strong solutions are versions of each other, and both are given by* (8.1).

*Proof.* We only need to prove that (2) implies (1). We may assume that $x = 0$. Let $U$ be a weak solution of (SACP) with initial value $x = 0$. Fix $t \in [0, T]$. We claim that the function $\Psi_t : (0, t) \to \mathscr{L}(H, E)$,

$$\Psi_t(r)h := \int_r^t S(s-r)Bh\, ds,$$

is stochastically integrable with respect to $W_H$ and

$$\int_0^t \Psi_t(r)\, dW_H(r) = \int_0^t U(s)\, ds. \tag{8.5}$$

To see this, note that for all $x^* \in E^*$ the stochastic Fubini theorem gives

$$\begin{aligned}\int_0^t \Psi_t^*(r)x^*\, dW_H(r) &= \int_0^t \int_r^t B^*S^*(s-r)x^*\, ds\, dW_H(r)\\ &= \int_0^t \int_0^s B^*S^*(s-r)x^*\, dW_H(r)\, ds = \int_0^t \langle U(s), x^*\rangle\, ds,\end{aligned}$$

where the last identity follows from the assumption that $U$ is a weak solution and therefore satisfies (8.1). The claim now follows from Theorem 6.17.

Also, from $\Psi_t(r)h \in \mathscr{D}(A)$ and $A\Psi_t(r)h = S(t-r)Bh - Bh$ it follows that $A\Psi_t : (0, t) \to \mathscr{L}(H, E)$ is stochastically integrable with respect to $W_H$ and

$$\int_0^t A\Psi_t(r)\, dW_H(r) = \int_0^t (S(t-r)B - B)\, dW_H(r) = U(t) - W^B(t),$$

where in the second identity we used that $W_H(t)B^*x^* = \langle W^B(t), x^*\rangle$.

Combining these facts it follows that $\Psi_t$ is stochastically integrable as a function from $(0, t)$ to $\mathscr{L}(H, \mathscr{D}(A))$. It follows that the left hand side of (8.5) defines a $\mathscr{D}(A)$-valued Gaussian random variable. Moreover, as $A$ is bounded from $\mathscr{D}(A)$ to $E$, almost surely we have

$$A\int_0^t U(s)\, ds = A\int_0^t \Psi_t(r)\, dW_H(r) = \int_0^t A\Psi_t(r)\, dW_H(r) = U(t) - W^B(t).$$

This shows that $U$ is a strong solution. □

We may now apply the result of Exercise 5.4 as follows:

**Corollary 8.11.** *Let $E$ have type* 2 *and assume that $B \in \gamma(H, E)$. Then the problem* (SACP) *has a unique strong solution, and this solution is given by the convolution* (8.1).

It can be shown that this solution has a version with continuous trajectories; this follows from the Da Prato-Kwapień-Zabczyk factorisation principle which will be discussed later on. It appears to be an open problem whether, in the more general situation of Theorems 8.6 and 8.10, a solution (if it exists) always has a continuous version.

## 8.5 Non-existence

In this section we present an example of a stochastic evolution equation driven by a rank one Brownian motion which has no (weak or strong) solution.

*Example 8.12.* Let $E = L^p(\mathbb{T})$, where $\mathbb{T}$ denotes the unit circle in the complex plane with its normalized Lebesgue measure. We let $A = d/d\theta$ denote the generator of the rotation (semi)group $S$ on $L^p(\mathbb{T})$, $S(t)f(\theta) = f(\theta + t \bmod 2\pi)$. Consider the stochastic Cauchy problem

$$\begin{cases} dU(t) = AU(t) + \phi\, dW, \quad t \in [0, 2\pi], \\ U(0) = 0, \end{cases} \tag{8.6}$$

where $W$ is a standard real Brownian motion and $\phi \in L^p(\mathbb{T})$ is a fixed element. This problem has a weak solution if and only if the operator $R := R_{2\pi} : L^2(\mathbb{T}) \to L^p(\mathbb{T})$ of Theorem 8.6 (with $T = 2\pi$) is $\gamma$-radonifying. Let $(h_n)_{n=1}^\infty$ be an orthonormal basis for $L^2(\mathbb{T})$. For all $N \geqslant M \geqslant 1$, by Fubini's theorem and the Khintchine inequality we have

$$\begin{aligned} \mathbb{E}\Big\| \sum_{n=M}^N \gamma_n R h_n \Big\|_{L^p(\mathbb{T})}^p &= \int_0^{2\pi} \mathbb{E}\Big| \sum_{n=M}^N \gamma_n R h_n(\theta)\Big|^p \, d\theta \\ &\eqsim_p \int_0^{2\pi} \Big( \sum_{n=M}^N |R h_n(\theta)|^2 \Big)^{\frac{p}{2}} d\theta = \Big\| \Big( \sum_{n=M}^N |R h_n|^2 \Big)^{\frac12} \Big\|_{L^p(\mathbb{T})}^p. \end{aligned}$$

Now,

$$\sum_{n=M}^N |R h_n(\theta)|^2 = \sum_{n=M}^N \Big| \int_0^{2\pi} h_n(t) \phi(\theta + t \bmod 2\pi)\, dt \Big|^2 = \sum_{n=M}^N \big| [h_n, \phi_\theta]_{L^2(\mathbb{T})} \big|^2$$

where $\phi_\theta(t) := \phi(\theta + t \bmod 2\pi)$. Via an application of the Kahane-Khintchine inequality we deduce that $R \in \gamma(L^2(\mathbb{T}), L^p(\mathbb{T}))$ if and only if $\phi \in L^2(\mathbb{T})$. In particular, for $p \in [1, 2)$ and $\phi \in L^p(\mathbb{T}) \setminus L^2(\mathbb{T})$ the resulting initial value problem has no weak solution.

It is not a coincidence that a nonexistence is obtained in the range $p \in [1, 2)$ only. Indeed, for $p \in [2, \infty)$ the space $L^p(\mathbb{T})$ has type 2, and therefore Corollary 8.11 guarantees the existence of a strong solution for (8.6).

## 8.6 Exercises

1. This exercise offers an alternative approach to the integration by parts formula of Lemma 8.1. The starting point is the fact that if $\mathscr{H}$ is a real

Hilbert space, $\phi : [0,T] \to \mathbb{R}$ is of bounded variation, and $\psi : [0,T] \to \mathscr{H}$ is continuous, then

$$\int_0^T \psi(t)\, d\phi(t) = \phi(T)\psi(T) - \int_0^T \phi(t)\, d\psi(t),$$

where both integrals are interpreted as Riemann-Stieltjes integrals in $\mathscr{H}$. Let $(W(t))_{t\in[0,T]}$ be a standard Brownian motion.

a) Show that the function $\psi : [0,t] \to L^2(\Omega)$, $\psi(t) := W(t)$, is continuous.
b) Deduce Lemma 8.1 from the above integration by parts formula.

2. Let $(W(t))_{t\in[0,T]}$ be an $E$-valued Brownian motion. Show that there exists a unique Gaussian covariance operator $Q \in \mathscr{L}(E^*, E)$ such that

$$\mathbb{E}\langle W(s), x^*\rangle\langle W(t), y^*\rangle = \min\{s,t\}\langle Qx^*, y^*\rangle$$

for all $0 \leqslant s,t \leqslant T$ and $x^*, y^* \in E^*$.
*Hint:* Consider $Q := Q_T/T$, where $Q_T$ is the covariance of $W(T)$.

3. We consider the problem (SACP) with initial value $x = 0$ and assume that it admits a weak solution $U$. Prove that $U$ is a Gaussian process with covariance

$$\mathbb{E}\langle U(s), x^*\rangle\langle U(t), y^*\rangle = \int_0^{\min\{s,t\}} [B^*S^*(s-r)x^*, B^*S^*(t-r)y^*]\, ds$$

for all $0 \leqslant s,t \leqslant T$ and $x^*, y^* \in E^*$.

4. We consider the problem (SACP) with initial value $x$ and assume that it admits a weak solution $U^x$.
   a) Prove that the solvability of the problem (SACP) is independent of the time $T$. More precisely, show that if (SACP) has a weak (resp. strong) solution on some interval $[0,T]$, then it has a weak (resp. strong) solution on every interval $[0,T]$.
   *Hint:* Use the semigroup property and Theorem 8.6.

   By a) and uniqueness, $U^x$ extends to a solution on $[0,\infty)$. For $f \in C_{\rm b}(E)$ and $t \geqslant 0$ we define the function $P(t)f : E \to \mathbb{R}$ by

$$P(t)f(x) := \mathbb{E}f(U^x(t)), \quad x \in E.$$

   b) Explain why for all $f \in C_{\rm b}(E)$ and $t \geqslant 0$ we have the identity

$$\mathbb{E}f(U^x(t)) = \int_E f(S(t)x + y)\, d\mu_t(y),$$

   where $\mu_t$ denotes the distribution of the random variable $U^0(t)$.
   c) Deduce that $P(t)f \in C_{\rm b}(E)$.

d) Prove the identity
$$\mu_{t+s} = \mu_t * S(t)\mu_s,$$
where $*$ denotes convolution and $S(t)\mu_s$ is the image measure of $\mu_s$ under the operator $S(t)$.
*Hint:* Use Fourier transforms and observe that for the covariances $Q_t$ of $U^0(t)$, $t \geqslant 0$, we have the identity
$$Q_{t+s} = Q_t + S(t)Q_sS^*(t).$$

e) Deduce that $P = (P(t))_{t\geqslant 0}$ is a semigroup of operators on $C_{\mathrm{b}}(E)$, in the sense that $P(0) = I$ and $P(t)P(s) = P(t+s)$ for all $t, s \geqslant 0$.

f) Prove that for all $x \in E$ and $f \in C_{\mathrm{b}}(E)$ we have
$$\lim_{t\to 0} P(t)f(x) = f(x)$$
uniformly on compact subsets $K$ of $E$.
*Hint:* By the remark in Exercise 6.4, the process
$$V^x(t) := S(t)x + \int_0^t S(s)B\,dW_H(s), \quad t \in [0,T],$$
has a continuous version (a proof will be given later in this course). Now use b) together with the observation that for each fixed $t \in [0,T]$ the random variables $U^x(t)$ and $V^x(t)$ are identically distributed.

*Remark:* By considering (real and imaginary parts of) trigonometric polynomials of the form $x \mapsto \exp(i\langle x, x^*\rangle)$ it is not hard to show that $P$ fails to be a $C_0$-semigroup on $C_{\mathrm{b}}(E)$ (and even on the closed subspace $UC_{\mathrm{b}}(E)$ of all bounded uniformly continuous functions) unless $A = 0$.

5. In addition to the assumptions of the previous exercise, let us assume that there exists a Borel probability measure $\mu_\infty$ on $E$ such that $\lim_{t\to\infty} \mu_t = \mu_\infty$ in the sense that
$$\lim_{t\to\infty} \int_E f(x)\,d\mu_t(x) = \int_E f(x)\,d\mu_\infty(x)$$
for all $f \in C_{\mathrm{b}}(E)$.

a) Prove the identity
$$\mu_\infty = \mu_t * S(t)\mu_\infty.$$

b) Prove that $\mu_\infty$ is an *invariant measure* in the sense that for all $f \in C_{\mathrm{b}}(E)$ and $t \geqslant 0$ we have
$$\int_E P(t)f\,d\mu_\infty = \int_E f\,d\mu_\infty.$$

c) Prove that $P$ extends to a $C_0$-semigroup of contractions on the space $L^p(E, \mu_\infty)$, $1 \leqslant p < \infty$.

**Notes.** The theory of (linear and non-linear) stochastic evolution equations in Hilbert spaces dates back to the 1970s and was developed extensively through the efforts of the Italian and Polish schools around Da Prato and Zabczyk. A comprehensive overview is given in the monographs [27, 28] by these two authors. Parts of the theory have been extended to (martingale-)type 2 spaces; we refer to the review paper by Brzeźniak [15] and the references given there.

The results of Section 8.3 are taken from [84] and generalise known Hilbert space results and improve the preliminary Banach space results of [16]. The proof of theorem Theorem 8.6 essentially follows the Hilbert space proof in [27]. The theory of adjoint semigroups was initiated by Phillips, who proved Proposition 8.3 and noted as a consequence that $E^{\odot} = E^*$ if $E$ is reflexive.

The equivalence of weak and strong solutions in the case where $B$ is $\gamma$-radonifying is taken from an unpublished note by Veraar.

The example in Section 8.5 is from [84]. Such examples cannot exist in Hilbert spaces, due to Corollary 8.11.

The semigroup $P$ of Exercises 4 and 5 is called the *Ornstein-Uhlenbeck semigroup* associated with $A$ and $B$. The literature on this class of semigroups is extensive, with contributions by many mathematicians. Using Itô's formula it can be shown that the infinitesimal generator $L$ of $P$ is given, on a suitable dense subspace of $\mathscr{D}(L)$ consisting of cylindrical functions, by

$$Lf(x) = \frac{1}{2}\mathrm{Tr}\,(BB^*D^2f(x)) + \langle Ax, Df(x)\rangle, \quad x \in \mathscr{D}(A),$$

where $D$ denotes the Fréchet derivative and Tr the trace. The first term in the right hand side is the 'diffusion part' corresponding to $B\,W_H$ and the second is the 'drift part' corresponding to $A$.

The clever argument in part f) of Exercise 4 is due to Veraar. A self-contained analytic proof can be found in see [42].

For a systematic account on invariant measures for stochastic evolution equations we refer to Da Prato and Zabczyk [28].

# 9

# $\gamma$-Boundedness

In this lecture we address the second topic in the paradigm sketched in the introduction of Lecture 5: a 'Gaussian' generalisation to a Banach space setting of the notion of uniform boundedness of families of operators in Hilbert spaces. Roughly speaking, a family of operators $\mathscr{T}$ is said to be '$\gamma$-bounded' if a Kahane contraction principle holds with scalars replaced by operators from $\mathscr{T}$. This makes $\gamma$-boundedness into a powerful tool for estimating Gaussian sums. Perhaps more important is the fact that there are numerous abstract methods to create $\gamma$-bounded families, which can be used to show that families of operators arising naturally in the context of parabolic PDEs (such as resolvents) and stochastic analysis (such as families of conditional expectation operators) are $\gamma$-bounded.

## 9.1 Randomised boundedness

Throughout this lecture $\varphi = (\varphi_n)_{n=1}^\infty$ denotes a sequence of independent symmetric real-valued random variables satisfying $\mathbb{E}\varphi_n^2 = 1$, $n \geqslant 1$. For instance, $\varphi$ could be a Rademacher sequence or a Gaussian sequence.

We begin with a simple observation.

**Proposition 9.1.** *Let $H_1$ and $H_2$ be Hilbert spaces. For a subset $\mathscr{T} \subseteq \mathscr{L}(H_1, H_2)$ and a constant $M \geqslant 0$ the following assertions are equivalent:*

(1) *$\mathscr{T}$ is uniformly bounded and $\sup_{T\in\mathscr{T}} \|T\| \leqslant M$;*
(2) *for all $N \geqslant 1$, all $T_1, \dots, T_N \in \mathscr{T}$, and all $x_1, \dots, x_N \in H_1$,*

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^N \varphi_n T_n x_n\Big\|^2\Big)^{\frac12} \leqslant M\Big(\mathbb{E}\Big\|\sum_{n=1}^N \varphi_n x_n\Big\|^2\Big)^{\frac12}.$$

*Proof.* For the proof of (1)$\Rightarrow$(2), write $\|h\|^2 = [h,h]$ and use that $\mathbb{E}\varphi_j\varphi_k = \delta_{jk}$. For the proof of (2)$\Rightarrow$(1), consider the case $N = 1$ in (2) to obtain $\|Th\| \leqslant M\|h\|$ for all $T \in \mathscr{T}$ and $h \in H_1$. $\square$

With Hilbert spaces replaced by Banach spaces the implication (1)⇒(2) does not hold in general. This motivates the following definition.

**Definition 9.2.** *Let $E_1$ and $E_2$ be Banach spaces. An operator family $\mathscr{T} \subseteq \mathscr{L}(E_1, E_2)$ is said to be* $\varphi$-bounded *if there exists a constant $M \geqslant 0$ such that*

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^{N} \varphi_n T_n x_n\Big\|^2\Big)^{\frac{1}{2}} \leqslant M\Big(\mathbb{E}\Big\|\sum_{n=1}^{N} \varphi_n x_n\Big\|^2\Big)^{\frac{1}{2}},$$

*for all $N \geqslant 1$, all $T_1, \dots, T_N \in \mathscr{T}$, and all $x_1, \dots, x_N \in E_1$. When $\varphi$ is a Rademacher sequence, a $\varphi$-bounded family is called* $R$-bounded*; when $\varphi$ is a Gaussian sequence the family is called* $\gamma$-bounded.

The least admissible constant $M$ is called the *$\varphi$-bound* of $\mathscr{T}$, notation: $\varphi(\mathscr{T})$. As in the Hilbert space case, every $\varphi$-bounded family $\mathscr{T}$ is uniformly bounded and we have

$$\sup_{T\in\mathscr{T}} \|T\| \leqslant \varphi(\mathscr{T}).$$

When $\varphi$ is a Rademacher sequence or a Gaussian sequence, the bound $\varphi(\mathscr{T})$ is denoted by $R(\mathscr{T})$ and $\gamma(\mathscr{T})$, respectively. In these two cases, the Kahane-Khintchine inequality shows that the exponent 2 in the definition may be replaced by any exponent $1 \leqslant p < \infty$; this only affects the numerical value of the bounds. $R$-bounds and $\gamma$-bounds relative to the $L^p$-norm will be denoted by $R_p(\mathscr{T})$ and $\gamma_p(\mathscr{T})$. As a rule, we will state our results relative to the $L^2$-norm, but frequently the results carry over to $L^p$-norms if we make this modification.

Proposition 9.3 below shows that every $R$-bounded family is $\gamma$-bounded, and Corollary 3.6 and Theorem 3.7 imply that the converse holds if $E_1$ has finite cotype.

**Proposition 9.3.** *Any $R$-bounded family $\mathscr{T}$ is $\varphi$-bounded and $\varphi(\mathscr{T}) \leqslant R(\mathscr{T})$.*

*Proof.* Let $(r_n')_{n=1}^\infty$ be a Rademacher sequence on an independent probability space $(\Omega', \mathscr{F}', \mathbb{P}')$. Then for all $T_1, \dots, T_N \in \mathscr{T}$ and $x_1, \dots, x_N \in E_1$, by randomising we obtain

$$\begin{aligned}\mathbb{E}\Big\|\sum_{n=1}^{N} \varphi_n T_n x_n\Big\|^2 &= \mathbb{E}\mathbb{E}'\Big\|\sum_{n=1}^{N} r_n'\varphi_n T_n x_n\Big\|^2 \\ &\leqslant R(\mathscr{T})^2\mathbb{E}\mathbb{E}'\Big\|\sum_{n=1}^{N} r_n'\varphi_n x_n\Big\|^2 = R(\mathscr{T})^2\mathbb{E}\Big\|\sum_{n=1}^{N} \varphi_n x_n\Big\|^2. \quad \square\end{aligned}$$

The proof of the next proposition is left as an exercise to the reader.

**Proposition 9.4.** *If $\mathscr{T} \subseteq \mathscr{L}(E_1, E_2)$ and $\mathscr{S} \subseteq \mathscr{L}(E_1, E_2)$ are $\varphi$-bounded, then the family $\mathscr{S} + \mathscr{T} = \{S + T : \ S \in \mathscr{S},\ T \in \mathscr{T}\}$ is $\varphi$-bounded in $\mathscr{L}(E_1, E_2)$ and*

$$\varphi(\mathscr{S} + \mathscr{T}) \leqslant \varphi(\mathscr{S}) + \varphi(\mathscr{T}).$$

*Likewise, if $\mathscr{T} \subseteq \mathscr{L}(E_1, E_2)$ and $\mathscr{S} \subseteq \mathscr{L}(E_2, E_3)$ are $\varphi$-bounded, then the family $\mathscr{S}\mathscr{T} = \{ST : \ S \in \mathscr{S}, \ T \in \mathscr{T}\}$ is $\varphi$-bounded in $\mathscr{L}(E_1, E_3)$ and*

$$\varphi(\mathscr{S}\mathscr{T}) \leqslant \varphi(\mathscr{S})\varphi(\mathscr{T}).$$

The *strong operator topology* of $\mathscr{L}(E_1, E_2)$ is the topology generated by all sets of the form

$$V(S, x, \varepsilon) := \{T \in \mathscr{L}(E_1, E_2) : \ \|Sx - Tx\| < \varepsilon\}$$

with given $S \in \mathscr{L}(E_1, E_2)$, $x \in E$, and $\varepsilon > 0$. Note that a set $O \subseteq \mathscr{L}(E_1, E_2)$ is open in this topology if and only if for all $S \in O$ there exist $x_1, \dots, x_k \in E_1$ and a number $\varepsilon > 0$ such that

$$\bigcap_{j=1}^{k} \{T \in \mathscr{L}(E_1, E_2) : \ \|Sx_j - Tx_j\| < \varepsilon\} \subseteq O.$$

It is an easy exercise to check that $\lim_{n\to\infty} T_n = T$ in the strong operator topology if and only if $\lim_{n\to\infty} T_n x = Tx$ for all $x \in E_1$.

**Proposition 9.5 (Strong closure).** *If $\mathscr{T} \subseteq \mathscr{L}(E_1, E_2)$ is $\varphi$-bounded, then its closure $\overline{\mathscr{T}}$ in the strong operator topology is $\varphi$-bounded and $\varphi(\overline{\mathscr{T}}) = \varphi(\mathscr{T})$.*

*Proof.* Let $\overline{T}_1, \dots, \overline{T}_N \in \overline{\mathscr{T}}$ and $x_1, \dots, x_N \in E_1$ be arbitrary. Given an $\varepsilon > 0$, choose operators $T_1, \dots, T_N \in \mathscr{T}$ such that $\|\overline{T}_n x_n - T_n x_n\| < 2^{-n}\varepsilon$, $n = 1, \dots, N$. Then, by the triangle inequality in $L^2(\Omega; E_2)$ applied twice,

$$\begin{aligned}
&\Big(\mathbb{E}\Big\| \sum_{n=1}^{N} \varphi_n \overline{T}_n x_n \Big\|^2\Big)^{\frac12} \\
&\quad \leqslant \Big(\mathbb{E}\Big\| \sum_{n=1}^{N} \varphi_n T_n x_n \Big\|^2\Big)^{\frac12} + \Big(\mathbb{E}\Big\| \sum_{n=1}^{N} \varphi_n (\overline{T}_n x_n - T_n x_n) \Big\|^2\Big)^{\frac12} \\
&\quad \leqslant \varphi(\mathscr{T})\Big(\mathbb{E}\Big\| \sum_{n=1}^{N} \varphi_n x_n \Big\|^2\Big)^{\frac12} + \sum_{n=1}^{N} \|\overline{T}_n x_n - T_n x_n\| \\
&\quad \leqslant \varphi(\mathscr{T})\Big(\mathbb{E}\Big\| \sum_{n=1}^{N} \varphi_n x_n \Big\|^2\Big)^{\frac12} + \varepsilon.
\end{aligned}$$

This proves that $\overline{\mathscr{T}}$ is $\varphi$-bounded with $\varphi(\overline{\mathscr{T}}) \leqslant \varphi(\mathscr{T})$. The converse inequality is trivial. ☐

The *absolute convex hull* of a set $V$, notation $\operatorname{abs\,conv}(V)$, is the set of all vectors of the form $\sum_{j=1}^{k} \lambda_j x_j$ with $\sum_{j=1}^{k} |\lambda_j| \leqslant 1$ and $x_j \in V$ for $j = 1, \dots, k$.

**Proposition 9.6 (Convex hull).** *If $\mathscr{T}$ is $\varphi$-bounded in $\mathscr{L}(E_1, E_2)$, then the convex hull and the absolute convex hull of $\mathscr{T}$ are $\varphi$-bounded in $\mathscr{L}(E_1, E_2)$ and $\varphi(\mathscr{T}) = \varphi(\mathrm{conv}(\mathscr{T})) = \varphi(\mathrm{abs\,conv}(\mathscr{T}))$.*

*Proof.* First we prove the statement for the convex hull. Choose $S_1, \dots S_n \in \mathrm{conv}(\mathscr{T})$ arbitrarily. Noting that

$$\mathrm{conv}(\mathscr{T}) \times \dots \times \mathrm{conv}(\mathscr{T}) = \mathrm{conv}(\mathscr{T} \times \dots \times \mathscr{T})$$

we can find $\lambda_1, \dots, \lambda_k \in [0,1]$ with $\sum_{j=1}^k \lambda_j = 1$ such that $S_n = \sum_{j=1}^k \lambda_j T_{jn}$ with $T_{jn} \in \mathscr{T}$ for all $j = 1, \dots, k$ and $n = 1, \dots, N$. Then, for all $x_1, \dots, x_N \in E_1$,

$$\begin{aligned}\Big(\mathbb{E}\Big\|\sum_{n=1}^N \varphi_n S_n x_n\Big\|^2\Big)^{\frac12} &\leqslant \sum_{j=1}^k \lambda_j \Big(\mathbb{E}\Big\|\sum_{n=1}^N \varphi_n T_{jn} x_n\Big\|^2\Big)^{\frac12} \\ &\leqslant \varphi(\mathscr{T}) \sum_{j=1}^k \lambda_j \Big(\mathbb{E}\Big\|\sum_{n=1}^N \varphi_n x_n\Big\|^2\Big)^{\frac12} = \varphi(\mathscr{T}) \Big(\mathbb{E}\Big\|\sum_{n=1}^N \varphi_n x_n\Big\|^2\Big)^{\frac12}.\end{aligned}$$

This proves the $\varphi$-boundedness of $\mathrm{conv}(\mathscr{T})$ with the estimate $\varphi(\mathrm{conv}(\mathscr{T})) \leqslant \varphi(\mathscr{T})$. The opposite inequality $\varphi(\mathscr{T}) \leqslant \varphi(\mathrm{conv}(\mathscr{T}))$ is trivial.

The result for the absolute convex hull follows by noting that this hull is contained in the convex hull of $\mathscr{T} \cup \{0\} \cup -\mathscr{T}$; the set $\mathscr{T} \cup \{0\} \cup -\mathscr{T}$ is $\varphi$-bounded with the same $\varphi$-bound as $\mathscr{T}$ (use Proposition 2.16 to add the zero operator and replace some of the $\varphi_n$ by $-\varphi_n$ in the random sums). $\square$

By combining Propositions 9.5 and 9.6 we obtain that the strongly closed absolutely convex hull of every $\varphi$-bounded set is $\varphi$-bounded. This may be used to show that $\varphi$-boundedness is preserved by taking integral means.

**Theorem 9.7 (Integral means I).** *Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space and let $\mathscr{T}$ be a $\varphi$-bounded subset of $\mathscr{L}(E_1, E_2)$. Suppose $f : A \to \mathscr{L}(E_1, E_2)$ is a function with the following properties:*

(i) *the function $\xi \mapsto f(\xi)x$ is strongly $\mu$-measurable for all $x \in E_1$;*
(ii) *we have $f(\xi) \in \mathscr{T}$ for $\mu$-almost all $\xi \in A$.*

*For $\phi \in L^1(A)$ define $T_f^\phi \in \mathscr{L}(E_1, E_2)$ by*

$$T_f^\phi x := \int_A \phi(\xi) f(\xi) x \, d\mu(\xi), \qquad x \in E_1,$$

*The family $\mathscr{T}_f^\phi := \{T_f^\phi : \ \|\phi\|_1 \leqslant 1\}$ is $\varphi$-bounded and $\varphi(\mathscr{T}_f^\phi) \leqslant \varphi(\mathscr{T})$.*

*Proof.* Since $\mathscr{T}$ is $\varphi$-bounded and therefore uniformly bounded, the integral defining $T_f^\phi x$ is well-defined as a Bochner integral in $E_2$ for every $x \in E_1$ and defines a bounded operator $T_f^\phi$ of norm $\|T_f^\phi\| \leqslant \|\phi\|_1 \sup_{T \in \mathscr{T}} \|T\|$.

To prove the $\varphi$-boundedness of the family $\mathscr{T}_f^\phi$ along with the estimate for its $\varphi$-bound it suffices to check that the family $\{T_f^\phi : \ \|\phi\|_1 = 1\}$ is contained in $\overline{\text{abs conv}}(\mathscr{T})$, where the bar denotes the closure in the strong operator topology of $\mathscr{L}(E_1, E_2)$.

Fix $\phi$ with $\|\phi\|_1 = 1$ and for $k = 1, 2, \ldots$ define $T^{(k)} \in \mathscr{L}(E_1^k, E_2^k)$ by

$$T^{(k)}(x_1, \ldots, x_k) := (T_f^\phi x_1, \ldots, T_f^\phi x_k)$$

and note that this operator is given by the Bochner integral

$$T^{(k)}(x_1, \ldots, x_k) = \int_A \phi(\xi) f^{(k)}(\xi)(x_1, \ldots, x_k) \, d\mu(\xi),$$

where $f^{(k)}(\xi)(x_1, \ldots, x_k) := (f(\xi)x_1, \ldots, f(\xi)x_k)$ is strongly $\mu$-measurable as an $E_2^k$-valued function of the variable $\xi$.

Let us fix $x_1, \ldots, x_k \in E_1$. Let $N \in \mathscr{A}$ be a $\mu$-null set such that (ii) holds in $A \setminus N$. Noting that

$$(|\phi|\mu)(B) := \int_B |\phi| \, d\mu = \int_B 1_{A\setminus N} |\phi| \, d\mu, \quad B \in \mathscr{A},$$

defines a probability measure on $(A, \mathscr{A})$ and writing

$$\phi(\xi) f(\xi) = \text{sgn}(\phi(\xi)) f(\xi) \cdot |\phi(\xi)|,$$

from Proposition 1.17 we deduce that

$$(T_f^\phi x_1, \ldots, T_f^\phi x_k) \in \overline{\text{abs conv}}\{(f(\xi)x_1, \ldots, f(\xi)x_k) : \ \xi \in A \setminus N\}.$$

In particular,

$$(T_f^\phi x_1, \ldots, T_f^\phi x_k) \in \overline{\text{abs conv}}\{(Tx_1, \ldots, Tx_k) : \ T \in \mathscr{T}\}.$$

This means that for every $\varepsilon > 0$ we can find $T \in \text{abs conv}(\mathscr{T})$ such that

$$\|T_f^\phi x_j - Tx_j\| < \varepsilon, \qquad j = 1, \ldots, k.$$

Since the choice of $x_1, \ldots, x_k \in E_1$ and $\varepsilon > 0$ were arbitrary, we have shown that every open set (in the strong operator topology) in $\mathscr{L}(E_1, E_2)$ containing $T_f^\phi$ intersects abs conv$(\mathscr{T})$. This is synonymous to saying that $T_f^\phi \in \overline{\text{abs conv}}(\mathscr{T})$. □

So far we have been concerned with producing new $\varphi$-bounded families from old. We continue with two results which produce $\varphi$-bounded families 'from scratch'. In view of Proposition 9.3 it suffices to prove that such families are $R$-bounded. In both cases, however, the same argument already gives the $\varphi$-boundedness, and we prefer this route for the unity of presentation.

**Theorem 9.8 (Integral means II).** *Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space, $E_1$ and $E_2$ Banach spaces and let $f : A \to \mathscr{L}(E_1, E_2)$ be a function with the property that $\xi \mapsto f(\xi)x$ is strongly $\mu$-measurable for all $x \in E_1$. Suppose that $g : A \to \mathbb{R}$ is a $\mu$-integrable function such that for all $x \in E_1$ we have*

$$\|f(\xi)x\| \leqslant |g(\xi)|\|x\| \quad \mu\text{-almost everywhere.}$$

*For $\phi \in L^\infty(A)$ define $T_f^\phi \in \mathscr{L}(E_1, E_2)$ by*

$$T_f^\phi x := \int_A \phi(\xi) f(\xi) x \, d\mu(\xi), \quad x \in E_1.$$

*The family $\mathscr{T}_f^\phi = \{T_f^\phi : \|\phi\|_\infty \leqslant 1\}$ is $\varphi$-bounded and $\varphi(\mathscr{T}_f^\phi) \leqslant \|g\|_1$.*

*Proof.* For $\phi \in L^\infty(A)$, note that $\xi \mapsto \phi(\xi)f(\xi)x$ is $\mu$-Bochner integrable in $E_2$ for all $x \in E_1$, so the operators $T_f^\phi$ are well-defined and bounded with $\|T_f^\phi\| \leqslant \|\phi\|_\infty \|g\|_1$.

Fix $\phi_1, \dots, \phi_N \in L^\infty(A)$ and $x_1, \dots, x_N \in E_1$. Using the Kahane contraction principle we estimate

$$\begin{aligned}
\Big\| \sum_{n=1}^N \varphi_n T_f^{\phi_n} x_n \Big\|_{L^2(\Omega;E_2)} &= \Big\| \int_A \sum_{n=1}^N \varphi_n \phi_n(\xi) f(\xi) x_n \, d\mu(\xi) \Big\|_{L^2(\Omega;E_2)} \\
&\leqslant \int_A \Big\| \sum_{n=1}^N \varphi_n \phi_n(\xi) f(\xi) x_n \Big\|_{L^2(\Omega;E_2)} d\mu(\xi) \\
&\leqslant \int_A \Big\| \sum_{n=1}^N \varphi_n f(\xi) x_n \Big\|_{L^2(\Omega;E_2)} d\mu(\xi) \\
&\leqslant \int_A |g(\xi)| \Big\| \sum_{n=1}^N \varphi_n x_n \Big\|_{L^2(\Omega;E_1)} d\mu(\xi) \\
&= \|g\|_1 \Big\| \sum_{n=1}^N \varphi_n x_n \Big\|_{L^2(\Omega;E_1)}.
\end{aligned}$$ □

Note that if $E_1$ is separable, we may apply the theorem to the function $g(\xi) := \|f(\xi)\|$, which is then $\mu$-measurable (choose a dense sequence $(x_n)_{n=1}^\infty$ in the unit ball of $E_1$ and note that $\|f(\xi)\| = \sup_{n \geqslant 1} \|f(\xi)x_n\|$). A similar remark applies to the next theorem.

**Theorem 9.9 (Functions with integrable derivative).** *Let $f : (a,b) \to \mathscr{L}(E_1, E_2)$ and $g : (a,b) \to \mathbb{R}$ be such that the functions $t \mapsto f(t)x$ are continuously differentiable, $g$ is integrable, and for all $x \in E_1$ we have*

$$\|f'(t)x\| \leqslant |g(t)|\|x\| \quad \mu\text{-almost everywhere.}$$

*Then $\mathscr{T} := \{f(t) : t \in (a,b)\}$ is $\varphi$-bounded and $\varphi(\mathscr{T}) \leqslant \|f(a+)\| + \|g\|_1$.*

*Proof.* Let us first prove that $f(a+) := \lim_{t\downarrow a} f(t)$ exists in the strong operator topology. For fixed $x \in E_1$, given $\varepsilon > 0$ choose $\delta > 0$ so small that $\int_a^{a+\delta} |g(t)|\,dt < \varepsilon$; then for all $a < a_1 < a_2 < a+\delta$ we have

$$\|f(a_2)x - f(a_1)x\| = \Big\| \int_{a_1}^{a_2} f'(t)x\,dt \Big\| \leqslant \int_{a_1}^{a_2} \|f'(t)x\|\,dt < \varepsilon\|x\|.$$

This gives the claim.

For all $a < t_1 \leqslant \cdots \leqslant t_N < b$ and $x_1, \dots, x_N \in E$ we obtain, using Theorem 9.8,

$$\begin{aligned}
&\Big(\mathbb{E}\Big\| \sum_{n=1}^N \varphi_n f(t_n) x_n \Big\|^2\Big)^{\frac12} \\
&\quad = \Big(\mathbb{E}\Big\| \sum_{n=1}^N \varphi_n \Big[ f(a+)x_n + \int_a^{t_n} f'(t)x_n\,dt\Big]\Big\|^2\Big)^{\frac12} \\
&\quad \leqslant \|f(a+)\| \Big(\mathbb{E}\Big\| \sum_{n=1}^N \varphi_n x_n \Big\|^2\Big)^{\frac12} + \Big(\mathbb{E}\Big\| \sum_{n=1}^N \varphi_n \int_a^b 1_{(a,t_n)}(t) f'(t) x_n\,dt \Big\|^2\Big)^{\frac12} \\
&\quad \leqslant (\|f(a+)\| + \|g\|_1) \Big(\mathbb{E}\Big\| \sum_{n=1}^N \varphi_n x_n \Big\|^2\Big)^{\frac12}.
\end{aligned}$$

□

## 9.2 Examples

We proceed with some important examples of $\varphi$-bounded families, where as before $(\varphi_n)_{n=1}^\infty$ is a sequence of independent symmetric real-valued random variables satisfying $\mathbb{E}\varphi_n^2 = 1$, $n \geqslant 1$. One example has already been recorded: a family of Hilbert space operators is $\varphi$-bounded if and only if it is uniformly bounded.

*Example 9.10 (The contraction principle and $\varphi$-boundedness).* Let $E$ be a Banach space. Every real number $a$ defines a bounded operator $T_a$ on $E$ by scalar multiplication: $T_a x = ax$. The Kahane contraction principle can be reformulated as saying that for every bounded set $A \subseteq \mathbb{R}$, the set $\mathscr{T}_A := \{T_a :\ a \in A\}$ is $\varphi$-bounded in $\mathscr{L}(E)$, with $\varphi(\mathscr{T}_A) = \sup\{|a| :\ a \in A\}$.

*Example 9.11 ($\varphi$-Boundedness in $L^p$).* Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space and let $1 \leqslant p < \infty$ be fixed. If $S$ is a positive bounded operator on $E := L^p(A)$, i.e., $Sf \geqslant 0$ whenever $f \geqslant 0$ (we write $f_1 \geqslant f_2$ to mean that $f_1(\xi) \geqslant f_2(\xi)$ for $\mu$-almost all $\xi \in A$), the set

$$\mathscr{T} := \{T \in \mathscr{L}(E) :\ |Tf| \leqslant S|f| \text{ for all } f \in E\}$$

is $\varphi$-bounded and we have $\varphi(\mathscr{T}) \leqslant K_p\|S\|$, where $K_p$ is a universal constant depending only on $p$.

By Proposition 9.3 it suffices to prove this for Rademacher variables $(r_n)_{n=1}^\infty$. Using Fubini's theorem and the scalar Kahane-Khintchine inequality, we see that for all $T_1, \dots, T_N \in \mathscr{T}$ and $f_1, \dots, f_N \in E$,

$$\begin{aligned}
\mathbb{E}\Big\|\sum_{n=1}^N r_n T_n f_n\Big\|_E^p &= \int_A \mathbb{E}\Big|\sum_{n=1}^N r_n T_n f_n\Big|^p \, d\mu \lesssim_p \int_A \Big(\mathbb{E}\Big|\sum_{n=1}^N r_n T_n f_n\Big|^2\Big)^{\frac{p}{2}} \, d\mu \\
&= \int_A \Big(\sum_{n=1}^N |T_n f_n|^2\Big)^{\frac{p}{2}} \, d\mu \leqslant \int_A \Big(\sum_{n=1}^N (S|f_n|)^2\Big)^{\frac{p}{2}} \, d\mu \\
&= \int_A \Big(\mathbb{E}\Big|\sum_{n=1}^N r_n S|f_n|\Big|^2\Big)^{\frac{p}{2}} \, d\mu \lesssim_p \int_A \mathbb{E}\Big|\sum_{n=1}^N r_n S|f_n|\Big|^p \, d\mu \\
&= \mathbb{E}\Big\|\sum_{n=1}^N r_n S|f_n|\Big\|_E^p \leqslant \|S\|^p \mathbb{E}\Big\|\sum_{n=1}^N r_n |f_n|\Big\|_E^p \\
&= \|S\|^p \int_A \mathbb{E}\Big|\sum_{n=1}^N r_n |f_n|\Big|^p \, d\mu = \|S\|^p \int_A \mathbb{E}\Big|\sum_{n=1}^N r_n f_n\Big|^p \, d\mu \\
&= \|S\|^p \mathbb{E}\Big\|\sum_{n=1}^N r_n f_n\Big\|_E^p.
\end{aligned}$$

In this computation we used that $\mathbb{E}\big|\sum_{n=1}^N r_n a_n\big|^p = \mathbb{E}\big|\sum_{n=1}^N r_n |a_n|\big|^p$ for $a_1, \dots, a_N \in \mathbb{R}$; to see this, just replace $r_n$ by $-r_n$ if $a_n < 0$. The result now follows from the Kahane-Khintchine inequality which permits us to replace the $L^p$-moments by $L^2$-moments.

## 9.3 A multiplier result

Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space and let $\Phi : A \to \gamma(H, E)$ be uniformly bounded and strongly $\mu$-measurable. For $f \in L^2(A; H)$ the integrals

$$R_\Phi f = \int_A \Phi(\xi) f(\xi) \, d\mu(\xi)$$

exist as Bochner integrals in $E$, and the resulting linear operator $R_\Phi : L^2(A; H) \to E$ is bounded.

Turning to the situation where $\Phi : A \to \gamma(H, E_1)$ is uniformly bounded and strongly $\mu$-measurable, suppose next that $E_2$ is another Banach space and $M : A \to \mathscr{L}(E_1, E_2)$ is a uniformly bounded function with the property that $\xi \mapsto M(\xi)x$ is strongly $\mu$-measurable for all $x \in E_1$ (in this situation, with a slight abuse of terminology we call $M$ *strongly $\mu$-measurable*). We put

$$(M\Phi)(\xi) := M(\xi)\Phi(\xi).$$

Let us check that the function $M\Phi$ is strongly $\mu$-measurable. By strong $\mu$-measurability, the range of $\Phi$ is $\mu$-separably-valued in $\gamma(H, E_1)$. Therefore by Proposition 5.10 we may assume $H$ is separable. Choose an orthonormal basis $(h_n)_{n=1}^{\infty}$ for $H$ and let $P_n$ denote the orthogonal projection onto the span of $\{h_1, \dots, h_n\}$. Then $\xi \mapsto (M\Phi P_n)(\xi) := M(\xi)\Phi(\xi)P_n$ is strongly $\mu$-measurable, and the claim follows by noting that $\lim_{n\to\infty} M\Phi P_n = M\Phi$ pointwise in the norm of $\gamma(H, E_1)$ by Proposition 5.12.

As a result, the integral operator $R_{M\Phi}$ is well-defined as a bounded operator from $L^2(A; H)$ to $E_2$. Thus $M$ induces a mapping

$$\widetilde{M} : R_\Phi \mapsto R_{M\Phi}.$$

We shall be interested in finding conditions that guarantee the boundedness of this mapping as an operator from $\gamma(L^2(A; H), E_1)$ to $\gamma(L^2(A; H), E_2)$.

First we check that the operators $R_\Phi$, with $\Phi$ a finite rank simple function, are dense in $\gamma(L^2(A; H), E_1)$.

**Lemma 9.12.** *Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space. The operators $R_\Phi$, with $\Phi : A \to \gamma(H, E)$ a finite rank simple function, form a dense subspace in $\gamma(L^2(A; H), E)$.*

*Proof.* Let $R \in \gamma(L^2(A; H), E)$ be given. By the same argument as in Step 4 of the proof of Theorem 6.17 we may assume that $R \in \gamma(L^2(A; \widetilde{H}), E)$, where $\widetilde{H}$ is a separable closed subspace of $H$. Then, by Proposition 5.12, we even may assume that $\widetilde{H}$ is finite-dimensional.

Let $\mathscr{A}_k$ be a sequence of finite sub-$\sigma$-algebras generating $\mathscr{A}$. Let $I_k$ denote the associated averaging operator on $L^2(A; H)$. Then $\lim_{k\to\infty} R \circ I_k = R$ in $\gamma(L^2(0, T; \widetilde{H}), E)$ by Proposition 5.12. Every $R \circ I_k$ is of the form $R_{\Phi_k}$ for a simple function $\Phi_k : A \to \gamma(\widetilde{H}, E)$, and as function with values in $\gamma(H, E)$, the $\Phi_k$ are finite rank simple functions. $\square$

The following multiplier theorem, due to Kalton and Weis in a slightly simpler setting, connects the notions of $\gamma$-boundedness and $\gamma$-radonification. It states that functions with $\gamma$-bounded range act as multipliers on spaces of $\gamma$-radonifying operators.

**Theorem 9.13 ($\gamma$-Bounded functions as $\gamma$-multipliers).** *Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space. Suppose that $M : A \to \mathscr{L}(E_1, E_2)$ is strongly measurable and has $\gamma$-bounded range $\mathscr{M} := \{M(t) : t \in A\}$. Then for every finite rank simple function $\phi : A \to \gamma(H, E_1)$ the operator $T_{M\phi}$ belongs to $\gamma^\infty(L^2(A; H), E_2)$ and*

$$\|T_{M\phi}\|_{\gamma_p^\infty(L^2(A;H),E_2)} \leqslant \gamma_p(\mathscr{M})\, \|T_\phi\|_{\gamma_p(L^2(A;H),E_1)}, \quad 1 \leqslant p < \infty.$$

*As a result, the map $\widetilde{M} : T_\phi \mapsto T_{M\phi}$ has a unique extension to a bounded operator*

$$\widetilde{M} : \gamma_p(L^2(A;H),E_1) \to \gamma_p^\infty(L^2(A;H),E_2)$$

*of norm* $\|\widetilde{M}\| \leqslant \gamma_p(\mathscr{M})$.

*Remark 9.14.* If $\mathscr{H}$ is any Hilbert space and the Banach space $E$ does not contain a closed subspace isomorphic to $c_0$, then $\gamma^\infty(\mathscr{H},E) = \gamma(\mathscr{H},E)$ by the theorem of Hoffmann-Jorgensen and Kwapień (Theorem 5.9). Thus, in the context of Theorem 9.13, if $E_2$ does not contain a closed subspace isomorphic to $c_0$, then $\widetilde{M}$ extends to a bounded operator acting from $\gamma(L^2(A;H),E_1)$ into $\gamma(L^2(A;H),E_2)$.

*Proof.* The uniqueness part follows from Lemma 9.12.

To prove the boundedness of $\widetilde{M}$ we let $\phi : A \to H \otimes E_1$ be a finite rank simple function which is kept fixed throughout the proof. Since we are fixing $\phi$ there is no loss of generality if we assume $H$ to be finite-dimensional, say with orthonormal basis $(h_n)_{n=1}^N$. Also, by virtue of the strong measurability of $M$, we may assume that the $\sigma$-algebra $\mathscr{A}$ is countably generated. This implies that $L^2(A)$ is separable, say with orthonormal basis $(g_m)_{m\geqslant 1}$.

We split the proof into two steps. In the second step we shall make use of conditional expectations (Lecture 11). For reasons of self-containedness we note that we only condition with respect to finite $\sigma$-algebras and that the conditional expectation operators could easily be replaced by averaging operators (at the expence of less transparent notations).

*Step 1* – In this step we consider the special case of the theorem where $M$ is a simple function. By passing to a common refinement we may suppose that

$$\phi = \sum_{j=1}^k 1_{B_j} U_j, \qquad M = \sum_{j=1}^k 1_{B_j} M_j,$$

with disjoint sets $B_j \in \mathscr{A}$ of finite positive measure; the operators $U_j \in H \otimes E_1$ are of finite rank and the operators $M_j$ belong to $\mathscr{M}$. Then,

$$M\phi = \sum_{j=1}^k 1_{B_j} M_j U_j.$$

This is a simple function with values in $H \otimes E_2$ which defines an operator $T_{M\phi} \in \gamma(L^2(A;H),E_2)$, and

$$\begin{aligned}
\|T_{M\phi}\|^p_{\gamma_p(L^2(A;H),E_2)} &= \mathbb{E}\Big\| \sum_{j=1}^k \sum_{n=1}^N \gamma_{jn} \sqrt{\mu(B_j)} M_j \Phi_j h_n \Big\|^p \\
&\leqslant (\gamma_p(\mathscr{M}))^p \mathbb{E}\Big\| \sum_{j=1}^k \sum_{n=1}^N \gamma_{jn} \sqrt{\mu(B_j)} \Phi_j h_n \Big\|^p \\
&= (\gamma_p(\mathscr{M}))^p \|T_\phi\|^p_{\gamma_p(L^2(A;H),E_1)}.
\end{aligned}$$

*Step 2* – Let $(A_j)_{j\geqslant 1}$ be a generating collection of sets in $\mathscr{A}$ and let, for all $k \geqslant 1$, $\mathscr{A}_k := \sigma(A_1, \dots, A_k)$. Define the functions $M_k : A \to \mathscr{L}(E_1, E_2)$ by

$$M_k x := \mathbb{E}(Mx|\mathscr{A}_k).$$

Since $\mathscr{A}_k$ is a finite $\sigma$-algebra, $M_k$ is a simple function. It is easily checked that for all $f \in L^2(A; H)$ we have $T_{M_k\phi} f = T_{M\phi}\mathbb{E}(f|\mathscr{A}_k)$, and therefore

$$\lim_{k\to\infty} T_{M_k\phi} f = T_{M\phi} f$$

strongly in $E_2$. By the $\gamma$-Fatou lemma (Proposition 5.3) it follows that $T_{M\phi} \in \gamma^\infty(L^2(A; H), E_1)$ and

$$\|T_{M\phi}\|_{\gamma_p^\infty(L^2(A;H),E_2)} \leqslant \liminf_{k\to\infty} \|T_{M\phi}\|_{\gamma_p(L^2(A;H),E_2)} \leqslant \gamma_p(\mathscr{M})\|T_\phi\|_{\gamma_p(L^2(A;H),E_1)}.$$

□

## 9.4 Exercises

1. Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space and let $1 \leqslant p < \infty$. For a function $\phi \in L^\infty(A)$ define the multiplier $M_\phi \in \mathscr{L}(L^p(A; E))$ by
$$(M_\phi f)(\xi) := \phi(\xi) f(\xi), \quad \xi \in A.$$
Show that the set $M = \{M_\phi : \ \|\phi\|_\infty \leqslant 1\}$ is $R$-bounded and give an estimate for $R(M)$.
2. On $l^p$ with $1 \leqslant p \leqslant \infty$, consider the left shift $S : (a_n)_{n\geqslant 1} \mapsto (a_{n+1})_{n\geqslant 1}$. For which values of $p$ is the family $\{S^k : \ k \geqslant 1\}$ $R$-bounded in $\mathscr{L}(l^p)$?
3. In this exercise we prove that analyticity implies $\gamma$-boundedness on compact sets. Let $D \subseteq \mathbb{C}$ be open and let $E$ be a Banach space. A function $f : D \to E$ is said to be *analytic* if the limit
$$f'(z_0) = \lim_{z\to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$
exists in $E$ for all $z_0 \in D$.
   a) Show that analytic functions are continuous.
   b) Use the Hahn-Banach theorem to show that Cauchy's formulas
   $$f(z_0) = \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z - z_0}\, dz, \quad f'(z_0) = \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{(z - z_0)^2}\, dz$$
   hold for an analytic function $f$, where $\Gamma$ is a simple contour in $D$ around $z_0$ (by (a), the integrals make sense as Bochner integrals).

c) Let $f : D \to \mathscr{L}(E_1, E_2)$ be a function such that $z \mapsto f(z)x$ is analytic for all $x \in E_1$. Show that for every compact set $K \subseteq D$ the family

$$\mathscr{T}_K := \{f(z) : \; z \in K\}$$

is $R$-bounded.
*Hint:* Use Theorem 9.9 and (c) to see that $f$ is $\gamma$-bounded on every circle contained in $D$. Then use the first formula in (b) together with Theorem 9.7.

4. (!) Let $\Phi : (0, T) \to \mathscr{L}(H, E_1)$ be stochastically integrable with respect to an $H$-cylindrical Brownian motion $W_H$ and suppose that $M : (0, T) \to \mathscr{L}(E_1, E_2)$ is strongly measurable and has $\gamma$-bounded range $\mathscr{M}$. Prove that $M\Phi : (0, T) \to \mathscr{L}(H, E_2)$ is stochastically integrable with respect to $W_H$ and

$$\mathbb{E}\Big\| \int_0^T M\Phi \, dW_H \Big\|^p \leqslant (\gamma_p(\mathscr{M}))^p \mathbb{E}\Big\| \int_0^T \Phi \, dW_H \Big\|^p,$$

where $\gamma_p(\mathscr{M})$ is the $\gamma$-bound of $\mathscr{M}$ relative to the $L^p$-norm (see the discussion following Definition 9.2).
*Hint:* Use the norm of $\gamma_p(L^2(0, T; H), E_1)$.

5. Let $E_1$ and $E_2$ be Banach spaces. Prove that the following assertions are equivalent:
(1) $E_1$ has cotype 2 and $E_2$ has type 2;
(2) every uniformly bounded subset of $\mathscr{L}(E_1, E_2)$ is $R$-bounded;
(3) every uniformly bounded subset of $\mathscr{L}(E_1, E_2)$ is $\gamma$-bounded.

*Hint:* For the proofs that (2) and (3) imply (1), consider suitable uniformly bounded families of rank one operators from $E_1$ to $E_2$. Recall that the notions of (co)type and Gaussian (co)type are equivalent (see Exercise 3.5).

*Remark:* Via Kwapień's theorem (see the Notes of Lecture 5), from this exercise we infer that for a Banach space $E$ the following assertions are equivalent:
(1) $E$ is isomorphic to a Hilbert space;
(2) every uniformly bounded subset of $\mathscr{L}(E)$ is $R$-bounded;
(3) every uniformly bounded subset of $\mathscr{L}(E)$ is $\gamma$-bounded.

**Notes.** The notion of $R$-boundedness has its origin in the work of Bourgain on vector-valued multiplier theorems and has since then been studied by many authors. The results presented here are taken from the fundamental papers by Clément, de Pagter, Sukochev, Witvliet [24] and Weis [108]. We refer to Denk, Hieber, Prüss [32] and Kunstmann and Weis [61] for more on the history of this notion and bibliographical references. It is well established by now that a large class of operators associated with analytic semigroups are $R$-bounded in $L^p$, a fact which explains the importance of $R$-boundedness for the theory of parabolic PDEs. Profound $R$-boundedness results are also available in harmonic analysis (e.g., in connection with Fourier multipliers) and probability theory (in connection with conditional expectation operators). The examples presented in this lecture only give a glimpse of the rich body of results nowadays available.

The result of Exercise 3 is due to Weis [108]. The result of Exercise 5 is due to Le Merdy and Pisier; see Arendt and Bu [4].

# 10

# Linear equations with additive noise II

In this lecture we pick up the thread of Lecture 8 and continue our investigation of the stochastic abstract Cauchy problem with additive noise,

$$\begin{cases} dU(t) = AU(t)\,dt + B\,dW_H(t), & t \in [0,T], \\ U(0) = x. \end{cases}$$

The goal is to prove optimal Hölder regularity results for the solutions in the parabolic case, that is, for operators $A$ generating an analytic $C_0$-semigroup. Since the problem is solved by

$$U(t) = S(t)x + \int_0^t S(t-s)B\,dW_H(s)$$

it suffices to concentrate on the case $x = 0$. Assuming that $x = 0$ and $B \in \gamma(H,E)$, we shall prove that $U$ has a Hölder continuous version for any exponent $\alpha < \frac{1}{2}$. The main technical tool is the $\gamma$-boundedness of the family $\{t^\beta(-A)^\alpha S(t) :\ t \in (0,T)\}$ for $0 < \alpha < \beta < \frac{1}{2}$ (Lemma 10.17). Thus by the $\gamma$-multiplier theorem (Theorem 9.13) this family acts as a multiplier in $\gamma(L^2(0,T;H),E)$. This provides a powerful tool for estimating the above stochastic integral.

## 10.1 Analytic semigroups

We begin with a discussion of analytic semigroups. In this section, all Banach spaces are complex. In later sections we shall return to the setting of real Banach spaces and apply the results to their complexifications.

We begin with a definition (cf. Exercise 9.3).

**Definition 10.1.** *Let $D \subseteq \mathbb{C}$ be open. A function $f : D \to E$ is* analytic *if*

$$\lim_{z\to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

*exists in* $E$ *for all* $z_0 \in D$.

Clearly, if $f$ is analytic, then $\langle f, x^*\rangle$ is analytic for all $x^* \in E^*$. In combination with the Hahn-Banach theorem, this fact may be used to show that many results on scalar-valued analytic functions extend to the vector-valued setting.

For $\eta \in (0, \pi]$ define the open sector

$$\Sigma_\eta = \{z \in \mathbb{C} \setminus \{0\} : \ |\arg(z)| < \eta\},$$

where the argument is taken in $(-\pi, \pi]$.

**Definition 10.2.** *A* $C_0$*-semigroup* $S$ *on* $E$ *is called* analytic on $\Sigma_\eta$ *if for all* $x \in E$ *the function* $t \mapsto S(t)x$ *extends analytically to* $\Sigma_\eta$ *and satisfies*

$$\lim_{z\in\Sigma_\eta,\, z\to 0} S(z)x = x.$$

*We call* $S$ *an* analytic $C_0$-semigroup *if* $S$ *is analytic on* $\Sigma_\eta$ *for some* $\eta \in (0, \pi]$.

The supremum of all $\eta \in (0, \pi]$ such that $S$ analytic on $\Sigma_\eta$ is called the *angle of analyticity* of $S$.

If $S$ is analytic on $\Sigma_\eta$, then for all $z_1, z_2 \in \Sigma_\eta$ we have

$$S(z_1)S(z_2) = S(z_1 + z_2).$$

Indeed, for each $x \in E$ the functions $z_1 \mapsto S(z_1)S(t)x$, $S(t)S(z_1)x$, and $S(z_1+t)x$ are analytic extensions of $s \mapsto S(s+t)x$ and are therefore equal. Repeating this argument, the functions $z_2 \mapsto S(z_1)S(z_2)x$, $S(z_2)S(z_1)x$, and $S(z_1+z_2)x$ are analytic extensions of $t \mapsto S(z_1 + t)x$ and are therefore equal.

As in the proof of Proposition 7.3, the uniform boundedness theorem implies that if $S$ is analytic on $\Sigma_\eta$, then $S$ is uniformly bounded on $\Sigma_{\eta'} \cap \{z \in \mathbb{C} : \ |z| \leqslant r\}$ for all $0 < \eta' < \eta$ and $r \geqslant 0$. Thus it makes sense to call $S$ a *uniformly bounded analytic* $C_0$*-semigroup* if $S$ is uniformly bounded on $\Sigma_\eta$ for some $\eta \in (0, \pi]$. Clearly, if $A$ generates an analytic $C_0$-semigroup on $\Sigma_\eta$, then for any $0 < \eta' < \eta$ the operator $A - \mu$ generates a uniformly bounded analytic $C_0$-semigroup on $\Sigma_{\eta'}$ if $\mu$ (depending on $\eta'$) is large enough.

**Theorem 10.3.** *For a closed and densely defined operator* $A$ *the following assertions are equivalent:*

(1) *there exists* $\eta \in (0, \frac{1}{2}\pi]$ *such that* $A$ *generates a uniformly bounded analytic* $C_0$*-semigroup on* $\Sigma_\eta$*;*
(2) *there exists* $\theta \in (\frac{1}{2}\pi, \pi]$ *such that* $\Sigma_\theta \subseteq \varrho(A)$ *and* $\sup_{\lambda\in\Sigma_\theta} \|\lambda R(\lambda, A)\| < \infty$*;*
(3) $S(t)x \in \mathscr{D}(A)$ *for all* $x \in E$ *and* $t > 0$*, and* $\sup_{t>0} t\|AS(t)\| < \infty$.

*In this situation, the suprema* $\widetilde{\eta}$ *and* $\widetilde{\theta}$ *for which* (1) *and* (2) *hold are related by* $\frac{1}{2}\pi + \widetilde{\eta} = \widetilde{\theta}$. *Furthermore we have the representation*

$$S(t)x = \frac{1}{2\pi i}\int_{\Gamma} e^{zt}R(z,A)x\,dz, \quad t > 0,\ x \in E,$$

*where* $\Gamma$ *is the upwards oriented boundary of* $\Sigma_{\theta'} \setminus B$ *for some* $\theta' \in (\frac{1}{2}\pi, \theta)$ *and* $B$ *is a closed ball centred at the origin.*

*Proof.* (1)⇒(2): By Proposition 7.8, if $G$ is the generator of a uniformly bounded $C_0$-semigroup, then $\{\mathrm{Re}\lambda > 0\} \subseteq \varrho(G)$ and $\lambda R(\lambda, G)$ is uniformly bounded on every proper sub-sector $\Sigma_\rho$, $0 < \rho < \frac{1}{2}\pi$.

Let $S$ be the $C_0$-semigroup generated by $A$ and let it be uniformly bounded on the sector $\Sigma_\eta$. The implication (1)⇒(2) follows by applying the above observation to the uniformly bounded $C_0$-semigroups $(S(e^{i\eta'}t))_{t\geqslant 0}$ with $0 < \eta' < \eta$, whose generators are $e^{i\eta'}A$. This gives the uniform boundedness of $\lambda R(\lambda, A)$ on the union of all sectors $e^{i\eta'}\Sigma_\rho$ for $0 < \eta' < \eta$ and $0 < \rho < \frac{1}{2}\pi$, which equals $\Sigma_{\frac{1}{2}\pi+\eta'}$. This argument also proves the inequality $\widetilde{\theta} \geqslant \frac{1}{2}\pi + \widetilde{\eta}$.

(2)⇒(3): First we prove that the conditions of (2) imply the integral representation for $S(t)x$.

The integral converges absolutely for all $t > 0$ and $x \in E$, and as a function of $t$ it extends to a bounded analytic function on the sector $\Sigma_{\eta'}$ for any $\eta' < \theta' - \frac{1}{2}\pi$. This proves the inequality $\widetilde{\theta} \leqslant \frac{1}{2}\pi + \widetilde{\eta}$.

Fix $t > 0$ and $x \in E$. For $\mu > 0$ such that $\mu \notin B$ define

$$v_\mu(t)x = \frac{1}{2\pi i}\int_{\Gamma} e^{zt}(\mu - z)^{-1}R(z,A)x\,dz.$$

Our aim is to show that $v_\mu(t)x = S(t)R(\mu, A)x$. Then,

$$S(t)x = \lim_{\mu\to\infty} S(t)\mu R(\mu, A)x = \lim_{\mu\to\infty} \mu v_\mu(t)x = \frac{1}{2\pi i}\int_{\Gamma} e^{zt}R(z,A)x\,dz,$$

where the first equality follows from Lemma 7.9 and the last is obtained by splitting $\Gamma = \Gamma_{r,1} \cup \Gamma_{r,2}$ with $\Gamma_{r,1} = \{z \in \Gamma : \|z\| \leqslant r\}$ and $\Gamma_{r,2} = \{z \in \Gamma : \|z\| \geqslant r\}$: for large fixed $r$, the integral over $\Gamma_{r,2}$ is less than $\varepsilon$, uniformly with respect to $\mu \geqslant 2r$, while the integral over $\Gamma_{r,1}$ tends to $\frac{1}{2\pi i}\int_{\Gamma_{r,1}} e^{zt}R(z,A)x\,dz$ by dominated convergence. Now pass to the limit $r \to \infty$.

The strategy is to prove that $t \mapsto v_\mu(t)x$ is a weak solution of the Cauchy problem

$$\begin{cases} u'(t) = Au(t), & t \in [0,T], \\ u(0) = R(\mu, A)x. \end{cases}$$

Then $t \mapsto v_\mu(t)x$ is a strong solution by Proposition 7.16 and by the uniqueness part of Theorem 7.17 it follows that $v_\mu(t)x = S(t)R(\mu, A)x$.

It is easily checked that $t \mapsto v_\mu(t)$ is integrable on $[0,T]$ (even continuous), and for all $x^* \in \mathscr{D}(A^*)$ we obtain

$$\begin{aligned}
\int_0^t \langle v_\mu(s), A^*x^*\rangle\,ds &= \int_0^t \frac{1}{2\pi i}\int_\Gamma e^{zs}(\mu-z)^{-1}\langle R(z,A)x, A^*x^*\rangle\,dz\,ds\\
&= \int_0^t \frac{1}{2\pi i}\int_\Gamma e^{zs}(\mu-z)^{-1}\langle zR(z,A)x - x, x^*\rangle\,dz\,ds\\
&\overset{(*)}{=} \int_0^t \frac{1}{2\pi i}\int_\Gamma e^{zs}(\mu-z)^{-1}\langle zR(z,A)x, x^*\rangle\,dz\,ds\\
&= \frac{1}{2\pi i}\int_\Gamma (e^{zt}-1)(\mu-z)^{-1}\langle R(z,A)x, x^*\rangle\,dz\\
&\overset{(**)}{=} \frac{1}{2\pi i}\int_\Gamma e^{zt}(\mu-z)^{-1}\langle R(z,A)x, x^*\rangle\,dz - \langle R(\mu,A)x, x^*\rangle.
\end{aligned}$$

Here the equality $(*)$ follows from the observation that by Cauchy's theorem we have

$$\frac{1}{2\pi i}\int_\Gamma (\mu-z)^{-1}e^{zs}\,dz = 0,$$

since $\mu \notin B$ is on the right of $\Gamma$. The equality $(**)$ follows from

$$\frac{1}{2\pi i}\int_\Gamma (\mu-z)^{-1}\langle R(z,A)x, x^*\rangle\,dz = \langle R(\mu,A)x, x^*\rangle$$

by the analyticity of the resolvent (Exercise 7.2) and Cauchy's theorem.

Now we are ready for the proof that (2) implies (3). Fix $t>0$ and $x \in E$. Since

$$M := \sup_{z\in\Gamma}\|AR(z,A)\| = \sup_{z\in\Gamma}\|zR(z,A) - I\|$$

is finite, the integral $\frac{1}{2\pi i}\int_\Gamma e^{tz}R(z,A)Ax\,dz$ converges absolutely. From Hille's theorem we deduce that $S(t)x \in \mathscr{D}(A)$ and

$$AS(t)x = \frac{1}{2\pi i}\int_\Gamma e^{tz}R(z,A)Ax\,dz.$$

By estimating this integral and letting the radius of the ball $B$ in the definition of $\Gamma$ tend to 0, it follows moreover that

$$\|AS(t)x\| \leqslant \frac{M}{\pi}\|x\|\int_0^\infty e^{\rho t\cos\theta'}\,d\rho = t^{-1}\frac{M}{\pi|\cos\theta'|}\|x\|.$$

(3)$\Rightarrow$(1): For all $x \in \mathscr{D}(A^n)$, $t \mapsto S(t)x$ is $n$ times continuously differentiable and $S^{(n)}(t)x = A^nS(t)x = (AS(t/n))^n x$. Since $\mathscr{D}(A^n)$ is dense, the boundedness of $AS(t/n)$ and closedness of the $n$th derivative in $C([0,T];E)$ together imply that the same conclusion holds for $x \in E$. Moreover, $\|S^{(n)}(t)x\| \leqslant C^n n^n/t^n\|x\|$, where $C$ is the supremum in (3). From $n! \geqslant n^n/e^n$ we obtain that for each $t>0$ the series

$$S(z)x := \sum_{n=0}^{\infty} \frac{1}{n!}(z-t)^n S^{(n)}(t)x$$

converges absolutely on the ball $B(t, rt/eC)$ for all $0 < r < 1$ and defines an analytic function there. The union of these balls is the sector $\Sigma_\eta$ with $\sin\eta = 1/eC$. We shall complete the proof by showing that $S(z)$ is uniformly bounded and satisfies $\lim_{z\to 0} S(z)x = x$ in $\Sigma_{\eta'}$ for each $0 < \eta' < \eta$. To this end we fix $0 < r < 1$. For $z \in B(t, rt/eC)$ we have

$$\|S(z)x\| \leqslant \sum_{n=0}^{\infty} \frac{1}{n!} r^n (t/eC)^n C^n n^n / t^n \|x\| \leqslant \sum_{n=0}^{\infty} r^n \|x\|.$$

This proves uniform boundedness on the sectors $\Sigma_{\eta'}$. To prove strong continuity it then suffices to consider $x \in \mathscr{D}(A)$, for which it follows from estimating the identity

$$S(z)x - x = e^{i\theta} \int_0^r S(se^{i\theta})Ax\,ds$$

where $z = re^{i\theta}$. □

*Remark 10.4.* We will use analyticity only through condition (3), which gives a 'real' characterisation of analyticity. In the context of semigroups on real Banach spaces this condition could be taken as the definition for analyticity, which has the advantage of avoiding the digressions through complexified spaces. In concrete examples, however, it is often easier to check analyticity using Definition 10.2 or condition (2) of Theorem 10.3.

By a rescaling argument we obtain:

**Corollary 10.5.** *If $A$ generates an analytic $C_0$-semigroup $S$ on $E$, then*

$$\limsup_{t\downarrow 0} t\|AS(t)\| < \infty.$$

From the fact that $S(t)x \in \mathscr{D}(A)$ for all $t > 0$ and $x \in E$ we deduce:

**Corollary 10.6.** *If $A$ generates an analytic $C_0$-semigroup $S$ on $E$, then for all initial values $x \in E$ the problem* (ACP) *has a unique classical solution, which is given by $u(t) = S(t)x$.*

## 10.2 Fractional powers

Throughout this section we assume that $A$ is the generator of a $C_0$-semigroup $S$ on $E$ which is *uniformly exponentially stable* in the sense that there exist constants $M \geqslant 1$ and $\mu > 0$ such that $\|S(t)\| \leqslant Me^{-\mu t}$ for all $t \geqslant 0$.

The next definition is motivated by the trivial identity

$$c^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-ct}\, dt, \quad c > 0,$$

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

**Definition 10.7.** *For* $0 < \alpha < 1$ *we define the* fractional power $(-A)^{-\alpha}$ *of* $-A$ *by the formula*

$$(-A)^{-\alpha} x := \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} S(t) x\, dt, \quad x \in E.$$

Note that $(-A)^{-\alpha}$ is well-defined and bounded on $E$ and commutes with $S(t)$ for all $t \geqslant 0$. Sometimes it is useful to extend the definition to the limiting values $\alpha \in \{0, 1\}$ by putting $(-A)^0 = I$ and $(-A)^{-1} = -A^{-1}$.

**Lemma 10.8.** *For all* $0 < \alpha, \beta < 1$ *satisfying* $0 < \alpha + \beta < 1$ *we have*

$$(-A)^{-\alpha}(-A)^{-\beta} = (-A)^{-\beta}(-A)^{-\alpha} = (-A)^{-\alpha-\beta}.$$

*Proof.* It suffices to prove that $(-A)^{-\alpha}(-A)^{-\beta} = (-A)^{-\alpha-\beta}$; the other identity follows upon interchanging $\alpha$ and $\beta$.

For all $x \in E$ we have

$$\begin{aligned}
(-A)^{-\alpha}(-A)^{-\beta} x &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty \int_0^\infty t^{\alpha-1} s^{\beta-1} S(s+t) x\, ds\, dt \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty \int_t^\infty t^{\alpha-1} (s-t)^{\beta-1} S(s) x\, ds\, dt \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty \Big( \int_0^s t^{\alpha-1} (s-t)^{\beta-1}\, dt \Big) S(s) x\, ds \\
&\overset{(*)}{=} \frac{1}{\Gamma(\alpha+\beta)} \int_0^\infty s^{\alpha+\beta-1} S(s) x\, ds = (-A)^{-\alpha-\beta} x,
\end{aligned}$$

where the identity $(*)$ follows from

$$\frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^s t^{\alpha-1}(s-t)^{\beta-1} dt = \frac{s^{\alpha+\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \tau^{\alpha-1}(1-\tau)^{\beta-1} d\tau = \frac{s^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)}.$$

Indeed, computing as above,

$$\begin{aligned}
\Gamma(\alpha+\beta) \int_0^1 \tau^{\alpha-1}(1-\tau)^{\beta-1}\, d\tau &= \int_0^\infty \int_0^1 s^{\alpha+\beta-1} \tau^{\alpha-1} (1-\tau)^{\beta-1} e^{-s}\, d\tau\, ds \\
&= \int_0^\infty \int_0^s t^{\alpha-1} (s-t)^{\beta-1} e^{-s}\, dt\, ds \\
&= \int_0^\infty \int_t^\infty t^{\alpha-1} (s-t)^{\beta-1} e^{-s}\, ds\, dt \\
&= \int_0^\infty \int_0^\infty t^{\alpha-1} s^{\beta-1} e^{-s-t}\, ds\, dt \\
&= \Gamma(\alpha)\Gamma(\beta).
\end{aligned}$$

$\square$

**Lemma 10.9.** *We have* $\sup_{0<\alpha<1} \|(-A)^{-\alpha}\| < \infty$.

*Proof.* We estimate $\|(-A)^{-\alpha}x\|$ by

$$\frac{1}{\Gamma(\alpha)}\Big\| \int_0^1 t^{\alpha-1} S(t)x\,dt\Big\| + \frac{1}{\Gamma(\alpha)}\Big\| \int_1^\infty t^{\alpha-1} S(t)x\,dt\Big\| =: (\mathrm{I}) + (\mathrm{II}).$$

Now,

$$(\mathrm{I}) \leqslant \frac{M\|x\|}{\alpha\Gamma(\alpha)} = \frac{M\|x\|}{\Gamma(\alpha+1)}, \quad (\mathrm{II}) \leqslant \frac{M\|x\|}{\Gamma(\alpha)} \int_1^\infty t^{\alpha-1}e^{-\mu t}\,dt \leqslant \frac{M\|x\|}{\mu^\alpha},$$

and both right hand sides are uniformly bounded for $0 < \alpha < 1$. $\square$

**Lemma 10.10.** *For all* $x \in E$, $\alpha \mapsto (-A)^{-\alpha}x$ *is continuous on* $[0,1]$.

*Proof.* First let $x \in \mathscr{D}(A)$ and put $Ax = y$. An integration by parts gives

$$\begin{aligned}(-A)^{-\alpha}x - x &= \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} S(t)x\,dt - x \\ &= -\frac{1}{\alpha\Gamma(\alpha)} \int_0^\infty t^\alpha S(t)y\,dt - x \\ &= -\int_0^\infty \Big(\frac{t^\alpha}{\Gamma(\alpha+1)} - 1\Big) S(t)y\,dt,\end{aligned}$$

where we used that $x = A^{-1}y = -\int_0^\infty S(t)y\,dt$. Hence, for any $r \geqslant 1$,

$$\|(-A)^{-\alpha}x - x\| \leqslant M\|y\| \int_0^r \Big|\frac{t^\alpha}{\Gamma(\alpha+1)} - 1\Big|\,dt + CM\|y\| \int_r^\infty t^\alpha e^{-\mu t}\,dt,$$

where $C = \sup\big\{\big|\frac{1}{\Gamma(\alpha+1)} - \frac{1}{t^\alpha}\big| :\ 0 < \alpha < 1,\ t \geqslant 1\big\}$. Choosing $r \geqslant 1$ so large that the second term is less than $\varepsilon$ and then passing to the limit $\alpha \downarrow 0$ in the first, we obtain the continuity of $\alpha \mapsto (-A)^{-\alpha}x$ at $\alpha = 0$ for $x \in \mathscr{D}(A)$. In view of the previous lemma, continuity at $\alpha = 0$ for all $x \in E$ follows from this.

The continuity of $\alpha \mapsto (-A)^{-\alpha}x$ at $\alpha = 1$ is proved in the same way, this time noting that for all $x \in \mathscr{D}(A)$ we have

$$(-A)^{-\alpha}x - (-A)^{-1}x = \int_0^\infty \Big(\frac{t^{\alpha-1}}{\Gamma(\alpha)} - 1\Big) S(t)x\,ds.$$

Finally the continuity for $\alpha \in (0,1)$ follows from the continuity at $\alpha = 0$ and the 'semigroup' property of Lemma 10.8. $\square$

**Lemma 10.11.** *For* $0 < \alpha < 1$ *the operator* $(-A)^{-\alpha}$ *is injective.*

*Proof.* Suppose $(-A)^{-\alpha}x = 0$. Then Lemma 10.8 implies $(-A)^{-\beta}x = 0$ for all $\alpha < \beta < 1$, and Lemma 10.10 gives $A^{-1}x = 0$. Hence $x = 0$. $\square$

This lemma suggests the following definition.

**Definition 10.12.** *For* $0 < \alpha < 1$ *define* $(-A)^{\alpha} := ((-A)^{-\alpha})^{-1}$.

As an unbounded operator with the range of $(-A)^{-\alpha}$ as its natural domain, $(-A)^{\alpha}$ is a closed and injective operator in $E$. With respect to the norm

$$\|x\|_{\mathscr{D}((-A)^{\alpha})} := \|(-A)^{\alpha}x\|, \tag{10.1}$$

$\mathscr{D}((-A)^{\alpha})$ is a Banach space and $(-A)^{\alpha} : \mathscr{D}((-A)^{\alpha}) \to E$ is an isometric isomorphism. For later reference we note that $\mathscr{D}(A)$ is dense in $\mathscr{D}((-A)^{\alpha})$. Indeed, for any $x \in \mathscr{D}((-A)^{\alpha})$ we have $\lim_{\lambda\to\infty} \lambda R(\lambda, A)(-A)^{\alpha}x = (-A)^{\alpha}x$, and since $R(\lambda, A)$ and $(-A)^{\alpha}$ commute this means that $\lim_{\lambda\to\infty} \lambda R(\lambda, A)x = x$ in the norm of $\mathscr{D}((-A)^{\alpha})$.

**Lemma 10.13.** *For* $0 < \alpha < 1$ *we have*

$$(-A)^{\alpha-1}(-A)^{-\alpha} = (-A)^{-\alpha}(-A)^{\alpha-1} = (-A)^{-1}.$$

*Proof.* This follows from Lemmas 10.8 and 10.10:

$$(-A)^{-1}x = \lim_{\beta\uparrow 1}(-A)^{-\beta}x = \lim_{\beta\uparrow 1}(-A)^{-\alpha}(-A)^{\alpha-\beta}x = (-A)^{-\alpha}(-A)^{\alpha-1}x. \quad \square$$

In the next two lemmas we assume that the $C_0$-semigroup $S$, in addition to being uniformly exponentially stable, is analytic.

**Lemma 10.14.** *For all* $0 < \alpha < 1$ *and* $t > 0$ *the operator* $(-A)^{\alpha}S(t)$ *is bounded and we have*

$$\sup_{t>0} t^{\alpha}\|(-A)^{\alpha}S(t)\| < \infty.$$

*Proof.* Since $S$ is analytic, $S(t)$ maps $E$ into $\mathscr{D}(A)$ and $\sup_{t>0} t\|AS(t)\| < \infty$. The boundedness of $(-A)^{\alpha}S(t)$ follows from the boundedness of $AS(t)$ by the identity $(-A)^{\alpha}S(t) = -(-A)^{\alpha-1}AS(t)$.

To prove the estimate, note that for all $x \in E$ we have

$$(-A)^{\alpha}S(t)x = \frac{-1}{\Gamma(1-\alpha)}\int_0^{\infty} s^{-\alpha}AS(t+s)x\,ds,$$

so, for $t > 0$,

$$\begin{aligned}\|(-A)^{\alpha}S(t)x\| &\leqslant \frac{C}{\Gamma(1-\alpha)}\int_0^{\infty} s^{-\alpha}(t+s)^{-1}\|x\|\,ds \\ &= \frac{Ct^{-\alpha}}{\Gamma(1-\alpha)}\int_0^{\infty} \tau^{-\alpha}(1+\tau)^{-1}\|x\|\,d\tau. \qquad \square\end{aligned}$$

**Lemma 10.15.** *For all* $0 < \alpha < 1$ *we have*

$$\sup_{t>0} t^{-\alpha}\|S(t)(-A)^{-\alpha} - (-A)^{-\alpha}\| < \infty.$$

*Proof.* From the identity $(-A)^{-\alpha}(-A)x = (-A)^{1-\alpha}x$ for $x \in \mathscr{D}(A)$ and Lemma 10.14 we obtain, for $t > 0$,

$$\|S(t)(-A)^{-\alpha}x - (-A)^{-\alpha}x\| \leqslant \int_0^t \|(-A)^{-\alpha}AS(s)x\|\,ds \leqslant \frac{Ct^\alpha}{\alpha}\|x\|,$$

where $C = \sup_{t>0} t^{1-\alpha}\|(-A)^{1-\alpha}S(t)\|$. □

In the next section we shall consider generators $A$ of analytic $C_0$-semigroups which are not necessarily uniformly exponentially stable. In this situation, fractional powers still can be defined for the shifted operators $A - \lambda$ with $\lambda$ large enough. The next lemma states that the resulting domains $\mathscr{D}((\lambda - A)^\alpha)$ are independent of $\lambda$.

To make things more precise, let $A$ be the generator of an arbitrary $C_0$-semigroup on $E$ and suppose $M \geqslant 1$ and $\mu \in \mathbb{R}$ are such that $\|S(t)\| \leqslant Me^{\mu t}$ for all $t \geqslant 0$.

**Lemma 10.16.** *For all* $0 < \alpha < 1$ *and* $\lambda_1, \lambda_2 > \mu$ *we have*

$$\mathscr{D}((\lambda_1 - A)^\alpha) = \mathscr{D}((\lambda_2 - A)^\alpha)$$

*with equivalent norms.*

*Proof.* The linear operator $(\lambda_2 - A)^{-\alpha}(\lambda_1 - A)^\alpha$ is a bounded and injective mapping from $\mathscr{D}((\lambda_1 - A)^\alpha)$ onto $\mathscr{D}((\lambda_2 - A)^\alpha)$ with inverse $(\lambda_1 - A)^{-\alpha}(\lambda_2 - A)^\alpha$. Thus $\mathscr{D}((\lambda_1 - A)^\alpha)$ and $\mathscr{D}((\lambda_2 - A)^\alpha)$ are isomorphic as Banach spaces. It remains to prove that $\mathscr{D}((\lambda_1 - A)^\alpha) = \mathscr{D}((\lambda_2 - A)^\alpha)$ as linear subspaces of $E$. But this follows from the fact that these spaces are the completions of $\mathscr{D}(A)$ with respect to the equivalent norms $\|\cdot\|_{\mathscr{D}((\lambda_1 - A)^\alpha)}$ and $\|\cdot\|_{\mathscr{D}((\lambda_2 - A)^\alpha)}$. □

This proposition justifies the notation $E_\alpha := \mathscr{D}((\lambda - A)^\alpha))$; this defines $E_\alpha$ as a Banach space up to an equivalent norm.

## 10.3 Hölder regularity

We now turn to the stochastic abstract Cauchy problem

$$\begin{cases} dU(t) = AU(t)\,dt + B\,dW_H(t), & t \in [0, T], \\ U(0) = 0. \end{cases} \tag{SACP$_0$}$$

We shall assume throughout this section that $A$ generates an analytic $C_0$-semigroup $S$ on $E$ satisfying $\|S(t)\| \leqslant Me^{\mu t}$ for certain $M \geqslant 1$, $\mu \in \mathbb{R}$, and all $t \geqslant 0$.

The key lemma for proving Hölder regularity of the solutions of (SACP$_0$) reads as follows.

**Lemma 10.17.** *For all real numbers $\alpha, \beta, \eta \geqslant 0$ satisfying $0 \leqslant \alpha + \eta < \beta < 1$ and $\lambda > \mu$, the set*

$$\{t^\beta (\lambda - A)^\eta S(t) : \ t \in (0,T)\}$$

*is R-bounded (and hence $\gamma$-bounded) in $\mathscr{L}(E, E_\alpha)$, with $\gamma$-bound $O(T^{\beta-\alpha-\eta})$.*

*Proof.* For all $x \in E$ the $\mathscr{L}(E, E_\alpha)$-valued function $\Psi(t)x := t^\beta(\lambda - A)^\eta S(t)x$ is continuously differentiable on $(0,T)$ with derivative

$$\Psi'(t)x = \beta t^{\beta-1}(\lambda - A)^\eta S(t)x + t^\beta (\lambda - A)^\eta AS(t)x,$$

where the second expression on the right hand side is well-defined since we may write $AS(t) = S(t/2)AS(t/2)$. By Lemma 10.14,

$$\|\Psi'(t)\|_{\mathscr{L}(E,E_\alpha)} \leqslant Ct^{\beta-\alpha-\eta-1}, \quad t \in (0,T),$$

where $C$ is a constant depending on $T$. Here we estimated the second term as

$$\begin{aligned} \|(\lambda - A)^\eta AS(t)\|_{\mathscr{L}(E,E_\alpha)} &= \|(\lambda - A)^\eta S(t/2)AS(t/2)\|_{\mathscr{L}(E,E_\alpha)} \\ &\leqslant Ct^{-\alpha-\eta}\|AS(t/2)\| \leqslant C't^{-\alpha-\eta-1}. \end{aligned}$$

Since $t^{\beta-\alpha-\eta-1}$ is integrable, the lemma follows from Theorem 9.9. $\square$

After these preparations we are ready to state and prove the main results of this lecture. The first is an existence result.

**Theorem 10.18.** *If $B \in \gamma(H,E)$, then the $\mathscr{L}(H,E)$-valued function $t \mapsto S(t)B$ is stochastically integrable on $(0,T)$ with respect to $W_H$. As a consequence, the stochastic Cauchy problem* (SACP$_0$) *associated with $A$ and $B$ admits a unique strong solution.*

*Proof.* By Theorems 8.6 and 8.10 it suffices to check that the function $\Phi(t) = S(t)B$ is stochastically integrable with respect to $W_H$, or equivalently, that the operator

$$R_\Phi f := \int_0^T \Phi(t)f(t)\,dt, \quad f \in L^2(0,T;H),$$

is $\gamma$-radonifying from $L^2(0,T;H)$ to $E$.

Pick a number $\beta \in (0, \frac{1}{2})$ and write

$$\Phi(t) = t^\beta S(t)[t^{-\beta}B] := t^\beta S(t)\Psi(t),$$

where $\Psi(t) := t^{-\beta}B$. By Lemma 10.17 and the $\gamma$-multiplier theorem (Theorem 9.13), the operator $R_\Phi$ belongs to $\gamma(L^2(0,T;H),E)$ once we know that $R_\Psi \in \gamma(L^2(0,T;H),E)$. But this is immediate from the result of Exercise 5.3, since $t \mapsto t^{-\beta}$ belongs to $L^2(0,T)$ and $B$ belongs to $\gamma(H,E)$. $\square$

Under the assumptions of the theorem we define the $E$-valued process $(U(t))_{t\in[0,T]}$ by

$$U(t) := \int_0^t S(t-s)B\, dW_H(s).$$

In order to formulate the second main result, for a Banach space $F$ and $0 \leqslant \beta < 1$ we define $C^\beta([0,T];F)$ as the space of all continuous functions $u : [0,T] \to F$ for which

$$\sup_{0\leqslant s<t\leqslant T} \frac{\|u(t)-u(s)\|}{(t-s)^\beta} < \infty.$$

The elements of $C^\beta([0,T];F)$ are said to be *Hölder continuous* of exponent $\beta$.

**Theorem 10.19 (Hölder regularity).** *Under the assumptions of the previous theorem, for all $\alpha \geqslant 0$ and $\beta \geqslant 0$ satisfying $\alpha+\beta < \frac{1}{2}$ and $1 \leqslant p < \infty$ the solution $U$ belongs to $L^p(\Omega;E_\alpha)$ and there exists a constant $C \geqslant 0$ such that for all $0 \leqslant s,t \leqslant T$,*

$$\big(\mathbb{E}\|U(t)-U(s)\|_{E_\alpha}^p\big)^{\frac{1}{p}} \leqslant C|t-s|^\beta.$$

*As a consequence, for all $\alpha \geqslant 0$ and $\beta \geqslant 0$ satisfying $\alpha + \beta < \frac{1}{2}$ the process $(U(t))_{t\in[0,T]}$ has a version with trajectories in $C^\beta([0,T];E_\alpha)$.*

*Proof.* By the Kahane-Khintchine inequality it suffices to prove the $L^p$-estimate for $p = 2$.

Fix $\alpha \geqslant 0$ and $\beta \geqslant 0$ such that $\alpha + \beta < \frac{1}{2}$. Let us first prove that for all $t \in [0,T]$ the random $U(t)$ takes its values in $E_\alpha$ almost surely. We do so by showing that $S(\cdot)B$ is stochastically integrable as an $\mathscr{L}(H,E_\alpha)$-valued function. Fix $\alpha < \theta < \frac{1}{2}$. Then $\{t^\theta S(t) : t \in (0,T)\}$ is $\gamma$-bounded in $\mathscr{L}(E,E_\alpha)$ by Lemma 10.17. As we have seen, the function $t \mapsto t^{-\theta}B$ defines an operator in $\gamma(L^2(0,T;H),E)$ of norm $\|t^{-\theta}\|_{L^2(0,T)}\|B\|_{\gamma(H,E)}$. Now Theorem 9.13 and the identity $S(t)B = t^\theta S(t)t^{-\theta}B$ imply that

$$\|R_{S(\cdot)B}\|_{\gamma(L^2(0,T;H),E_\alpha)} \leqslant C\|B\|_{\gamma(H,E)}.$$

Fix $0 \leqslant s \leqslant t \leqslant T$. By the triangle inequality in $L^2(\Omega;E_\alpha)$,

$$\begin{aligned}\big(\mathbb{E}\|U(t)-U(s)\|_{E_\alpha}^2\big)^{\frac{1}{2}} \leqslant{}& \Big(\mathbb{E}\Big\|\int_0^s [S(t-r)-S(s-r)]B\, dW(r)\Big\|_{E_\alpha}^2\Big)^{\frac{1}{2}} \\ &+ \Big(\mathbb{E}\Big\|\int_s^t S(t-r)B\, dW(r)\Big\|_{E_\alpha}^2\Big)^{\frac{1}{2}}.\end{aligned}$$

Choose $\lambda \in \mathbb{R}$ sufficiently large in order that the fractional powers of $\lambda - A$ exist. For the first term we have, for any choice of $\varepsilon > 0$ such that $\alpha+\beta+\varepsilon < \frac{1}{2}$, and using Lemmas 10.8 and 10.15,

$$
\begin{aligned}
&\mathbb{E}\Big\| \int_0^s [S(t-r) - S(s-r)]B\,dW(r)\Big\|_{E_\alpha}^2 \\
&\simeq \mathbb{E}\Big\| \int_0^s (s-r)^{\alpha+\beta+\varepsilon}(\lambda - A)^{\alpha+\beta} S(s-r) \\
&\qquad\qquad\qquad \times (s-r)^{-\alpha-\beta-\varepsilon}[S(t-s) - I](\lambda - A)^{-\beta} B\,dW(r)\Big\|^2 \\
&\leqslant C^2 \mathbb{E}\Big\| \int_0^s (s-r)^{-\alpha-\beta-\varepsilon}[S(t-s) - I](\lambda - A)^{-\beta} B\,dW(r)\Big\|^2 \\
&= C^2 \|[S(t-s) - I](\lambda - A)^{-\beta} B\|_{\gamma(H,E)}^2 \int_0^s (s-r)^{-2(\alpha+\beta+\varepsilon)}\,dr \\
&\leqslant C^2 \|[S(t-s) - I](\lambda - A)^{-\beta}\|^2 \|B\|_{\gamma(H,E)}^2 s^{1-2(\alpha+\beta+\varepsilon)} \\
&\leqslant C_T^2 (t-s)^{2\beta} \|B\|_{\gamma(H,E)}^2.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\mathbb{E}\Big\| \int_s^t S(t-r)B\,dW(r)\Big\|_{E_\alpha}^2 \\
&\qquad \eqsim \mathbb{E}\Big\| \int_s^t (t-r)^{\frac12-\beta}(\lambda - A)^{\alpha} S(t-r)(t-r)^{-\frac12+\beta} B\,dW(r)\Big\|^2 \\
&\qquad \leqslant C^2 \mathbb{E}\Big\| \int_s^t (t-r)^{-\frac12+\beta} B\,dW(r)\Big\|^2 \\
&\qquad = C^2 \|B\|_{\gamma(H,E)}^2 \int_s^t (t-r)^{-1+2\beta}\,dr \\
&\qquad \leqslant C_T^2 \|B\|_{\gamma(H,E)}^2 (t-s)^{2\beta}.
\end{aligned}
$$

The first part of the theorem follows by combining these estimates.

For the second part, pick $\beta < \beta' < \frac12 - \alpha$. Given $p \geqslant 1$, by the above we find a constant $C$ such that for $0 \leqslant s, t \leqslant T$,

$$\mathbb{E}\|U(t) - U(s)\|_{E_\alpha}^p \leqslant C^p |t-s|^{\beta' p}.$$

For $p$ large enough the existence of a version with $\beta$-Hölder continuous trajectories now follows from Kolmogorov's theorem (Theorem 6.9). □

*Example 10.20.* For the stochastic heat equation in $L^2(0,1)$ with Dirichlet boundary conditions, Theorem 10.19 implies the existence of a solution $U$ with trajectories in $\in C^\eta([0,T]; C^\theta[0,1])$ for all $\eta, \theta \geqslant 0$ satisfying $2\eta + \theta < \frac12$. This will be shown as a special case of a more general space-time regularity result in the last lecture.

## 10.4 Exercises

1. a) Show that the heat semigroup $S$ on $L^p(\mathbb{R}^d)$ with $1 \leqslant p < \infty$ (Example 7.20) is analytic on the sector $\Sigma_{\frac12\pi}$, and uniformly bounded on every proper subsector of $\Sigma_{\frac12\pi}$.

*Hint:* Put $S(z)f = K_z * f$, where $K_z$ is the analytic extension of the heat kernel. Use Young's inequality together with the estimate

$$\|K_z\|_1 = \Big(\frac{|z|}{\mathrm{Re}(z)}\Big)^{\frac{1}{2}d}.$$

b) Show that in $L^2(\mathbb{R}^d)$, $S$ is contractive on $\Sigma_{\frac{1}{2}\pi}$.

*Remark:* Using interpolation theory, the above results imply the estimate

$$\|S(t)\|_p \leqslant \Big(\frac{|z|}{\mathrm{Re}(z)}\Big)^{|\frac{1}{2}-\frac{1}{p}|d}, \quad \mathrm{Re}(z) > 0.$$

2. This exercise gives a two-dimensional example of a bounded analytic $C_0$-semigroup which is uniformly exponentially stable, contractive on $\mathbb{R}_+$, and fails to be contractive on any open sector containing $\mathbb{R}_+$.

   On $\mathbb{R}^2$ consider the norm $\|\cdot\|_Q$ induced by the inner product $[x,y]_Q := [Qx,y]$, where $[\cdot,\cdot]$ represents the standard inner product of $\mathbb{R}^2$ and

$$Q = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}.$$

   a) Show that the symmetric matrix $Q$ is positive and conclude that $[\cdot,\cdot]_Q$ does indeed define an inner product on $\mathbb{R}^2$.

   On $(\mathbb{R}^2, \|\cdot\|_Q)$ we consider the $C_0$-semigroup $S$,

$$S(t) = e^{-t/2}\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}.$$

   b) Show that $\|S(t)\|_Q^2 = \frac{1}{2}e^{-t}(t^2 + 2 + t\sqrt{t^2+4})$ and conclude that $S$ is contractive on $\mathbb{R}_+$.

   *Hint:* Use the fact that $\|S(t)\|_Q^2$ equals the largest eigenvalue of $S(t)S^*(t)$ (the adjoint refers to the inner product $[\cdot,\cdot]_Q$).

   On the complexification $\mathbb{C}^2$ of $\mathbb{R}^2$ we consider the inner product

$$\langle x, y\rangle_Q := \langle Qx, y\rangle,$$

   where this time $\langle\cdot,\cdot\rangle$ represents the standard inner product of $\mathbb{C}^2$. Prove that the complexified semigroup $S$ has the following properties:

   c) $S$ extends to an entire $C_0$-semigroup which is uniformly bounded on the sector $\Sigma_\eta$ for all $0 < \eta < \frac{1}{2}\pi$.

   d) $S$ fails to be contractive on any open sector $\Sigma_\eta$.

3. This exercise gives necessary and sufficient conditions for a closed densely defined operator $A$ in $E$ to generate an analytic $C_0$-semigroup which is contractive on a sector $\Sigma_\eta$. For $x \in E$ we define

$$\partial(x) = \{x^* \in E^* : \ \|x\| = \|x^*\|, \ \langle x, x^*\rangle = \|x\|\|x^*\|\}.$$

By the Hahn-Banach theorem, for all $x \in E$ we have $\partial(x) \neq \varnothing$. Let $A$ be a closed densely defined operator in $E$ and assume that $\varrho(A) \cap (0,\infty) \neq \varnothing$. Prove that the following assertions are equivalent:

(1) $A$ generates an analytic $C_0$-semigroup on $E$ which is contractive on an open sector $\Sigma_\eta$;

(2) There exists a constant $C \geqslant 0$ such that for all non-zero $x \in \mathscr{D}(A)$ and all $x^* \in \partial(x)$ we have
$$|\mathrm{Im}\langle Ax, x^*\rangle| \leqslant -C\,\mathrm{Re}\langle Ax, x^*\rangle;$$

(3) There exists a constant $C \geqslant 0$ such that for all non-zero $x \in \mathscr{D}(A)$ there exists $x^* \in \partial(x)$ such that
$$|\mathrm{Im}\langle Ax, x^*\rangle| \leqslant -C\,\mathrm{Re}\langle Ax, x^*\rangle.$$

*Hint:* For (1)⇒(2) differentiate the function $\mathrm{Re}\langle S(te^{i\eta'})x, x^*\rangle$ for $|\eta'| < \eta$ and $x^* \in \partial(x)$. For (3)⇒(1) observe that if $\cot\eta = C$, then for $x$ and $x^*$ as indicated and $\lambda = re^{i\eta'}$ with $|\eta'| < \eta$ we have $\|(\lambda - A)x\| \geqslant r\|x\| = |\lambda|\,\|x\|$.

4. Suppose that $A$ is a closed linear operator with $(0,\infty) \subseteq \varrho(A)$ and $\sup_{\lambda>0}(\lambda+1)\|R(\lambda,A)\| < \infty$.

 a) Show that $\Sigma_\eta \cup B \subseteq \varrho(A)$ and $\sup_{\lambda\in\Sigma_\eta\cup B}\|\lambda R(\lambda,A)\| < \infty$ for some $\eta > 0$ and some ball $B$ centred at the origin.

 Define
$$(-A)^{-\alpha} := \frac{1}{2\pi i}\int_\Gamma (-z)^{-\alpha} R(z,A)\,dz,$$
 where $\Gamma$ is the upwards oriented boundary of $\Sigma_\eta \cup B$, where $B$ is a closed ball centred at the origin.

 b) Show, by using Cauchy's formula, that
$$(-A)^{-\alpha} = \frac{\sin\alpha\pi}{\pi}\int_0^\infty \lambda^{-\alpha} R(\lambda,A)\,d\lambda.$$

 c) Now assume that $A$ generates a uniformly exponentially stable $C_0$-semigroup $S$ and prove that the definition in b) agrees with Definition 10.7.

5. In this exercise we sketch an alternative approach to Theorem 10.19, whose notations and assumptions we use.

 Let $U(t) = \int_0^t S(t-s)B\,dW_H(t)$. Being the weak solution of the problem $dU(t) = AU(t)\,dt + B\,dW_H$ with initial value $U(0) = 0$, the process $U$ has a version with integrable trajectories. Using this version we define
$$V(t) := \int_0^t U(s)\,ds.$$
 Let $W_B$ be the Brownian motion canonically associated with $B$ and $W_H$ (see Proposition 8.8). Fixing $0 \leqslant \alpha < \frac{1}{2}$, we note that $W_B$ has a version with trajectories in $C^\alpha([0,T];E)$ by Kolmogorov's theorem.

a) Show that almost surely, the following identity holds for all $t \in [0,T]$:

$$V(t) = \int_0^t S(t-s)W_B(s)\,ds.$$

b) Show that almost surely, $U$ has trajectories in $C^\alpha([0,T];E)$.
*Hint:* Show that almost surely, the trajectories of $V$ belong to $C^1([0,T];E)$ and have derivatives in $C^\alpha([0,T];E)$.

c) Refine this argument to obtain the result of Theorem 10.19.

**Notes.** The results of Sections 10.1 and 10.2 are standard.

Theorem 10.3 can be found in most textbooks on semigroups (see the Notes of Lecture 7). We have tried to shorten the proof as much as possible. Of course, much more is to be said about the representation of the operators $S(t)$ in terms of the resolvent $R(\lambda, A)$. Indeed, this formula is a special case of the complex inversion formula for the Laplace transform, and suitable generalisations can be given to arbitrary $C_0$-semigroups. We refer to Arendt, Batty, Hieber, Neubrander [3] for a thorough discussion of this topic. A systematic treatment of analytic semigroups and their applications to parabolic evolution equations is given in the monograph of Lunardi [72]. Exercise 2 is taken from [42].

Fractional powers of unbounded operators are discussed in Arendt, Batty, Hieber, Neubrander [3], Haase [45], Lunardi [72], and Pazy [89]. We followed the presentation of [89]. Our approach is rather *ad hoc* and was designed to keep the technicalities at a minimum. A more systematic approach starts from the Dunford integral along the lines of Exercise 4.

The results of Section 10.3 are taken from [34]. The proof of Theorem 10.19 presented here contains a simplification due to Veraar. Our results generalise the Hilbert space case which is due to Da Prato, Kwapień, Zabczyk [26]. The approach of [26] is based on a factorisation trick which is based on the identity

$$\frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)}\int_r^t (t-s)^{\alpha-1}(s-r)^{-\alpha}\,ds = 1,$$

valid for $0 < \alpha < 1$ and $t > r \geqslant 0$. This identity allows one to write the solution process as a repeated integral. Hölder regularity is then obtained by applying the stochastic Fubini theorem and exploiting the regularising properties of fractional integrals. This method was extended to Banach spaces by Millet and Smoleński [77]. The idea of Exercise 5 is taken from Da Prato, Kwapień, Zabczyk [26].

# 11

# Conditional expectations and martingales

Having finished our discussion of the stochastic Cauchy problem with additive noise, we now turn to the more difficult case of equations with multiplicative noise, where the fixed operator $B \in \mathscr{L}(H,E)$ is replaced by an operator-valued function $B : E \to \mathscr{L}(H,E)$:

$$\begin{cases} dU(t) = AU(t)\,dt + B(U(t))\,dt, \quad t \in [0,T], \\ U(0) = u_0. \end{cases}$$

The solutions are then no longer given in closed form by the explicit formula (8.1). Instead, they arise as fixed points of the stochastic integral equation

$$U(t) = S(t)x + \int_0^t S(t-s)B(U(s))\,dW_H(s).$$

The new difficulty here is that integrand is an $\mathscr{L}(H,E)$-valued process depending on $U$. This requires an extension of the stochastic integration theory of Lecture 6 to this more general situation. As it turns out, in the setting of UMD Banach spaces this can be achieved by a decoupling technique which reduces the construction of the stochastic integral to the one already covered.

In this lecture we introduce the notion of $E$-valued martingales. They will be used to define UMD Banach spaces as the class of Banach spaces $E$ such that certain a priori estimates hold for $E$-valued martingales. This may sound rather technical, but the important fact is that Hilbert spaces, $L^p$-spaces ($1 < p < \infty$), and spaces constructed from these, are UMD spaces. From the point of view of applications, the UMD spaces therefore constitute an important class of spaces.

## 11.1 Conditional expectations

Throughout this section we fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a sub-$\sigma$-algebra $\mathscr{G}$ of $\mathscr{F}$. For $1 \leqslant p \leqslant \infty$ we denote by $L^p(\Omega, \mathscr{G})$ the subspace of all $\xi \in L^p(\Omega)$ having a $\mathscr{G}$-measurable representative. With this notation, $L^p(\Omega) = L^p(\Omega, \mathscr{F})$.

**Lemma 11.1.** *$L^p(\Omega, \mathscr{G})$ is a closed subspace of $L^p(\Omega)$.*

*Proof.* Suppose that $(\xi_n)_{n=1}^\infty$ is a sequence in $L^p(\Omega, \mathscr{G})$ such that $\lim_{n\to\infty} \xi_n = \xi$ in $L^p(\Omega)$. We may assume that the $\xi_n$ are pointwise defined and $\mathscr{G}$-measurable. After passing to a subsequence (when $1 \leqslant p < \infty$) we may furthermore assume that $\lim_{n\to\infty} \xi_n = \xi$ almost surely. The set $C$ of all $\omega \in \Omega$ where the sequence $(\xi_n(\omega))_{n=1}^\infty$ converges is $\mathscr{G}$-measurable. Put $\widetilde{\xi} := \lim_{n\to\infty} 1_C \xi_n$, where the limit exists pointwise. The random variable $\widetilde{\xi}$ is $\mathscr{G}$-measurable and agrees almost surely with $\xi$. This shows that $\xi$ defines an element of $L^p(\Omega, \mathscr{G})$. $\square$

Our aim is to show that $L^p(\Omega, \mathscr{G})$ is the range of a contractive projection in $L^p(\Omega)$. For $p = 2$ this is clear: we have the orthogonal decomposition

$$L^2(\Omega) = L^2(\Omega, \mathscr{G}) \oplus L^2(\Omega, \mathscr{G})^\perp$$

and the projection we have in mind is the orthogonal projection, denoted by $P_{\mathscr{G}}$, onto $L^2(\Omega, \mathscr{G})$ along this decomposition. Following common usage we write

$$\mathbb{E}(\xi|\mathscr{G}) := P_{\mathscr{G}}\xi, \quad \xi \in L^2(\Omega),$$

and call $\mathbb{E}(\xi|\mathscr{G})$ the *conditional expectation* of $\xi$ with respect to $\mathscr{G}$. Let us emphasise that $\mathbb{E}(\xi|\mathscr{G})$ is defined as an element of $L^2(\Omega, \mathscr{G})$, that is, as an equivalence class of random variables.

**Lemma 11.2.** *For all $\xi \in L^2(\Omega)$ and $G \in \mathscr{G}$ we have*

$$\int_G \mathbb{E}(\xi|\mathscr{G})\, d\mathbb{P} = \int_G \xi\, d\mathbb{P}.$$

*As a consequence, if $\xi \geqslant 0$ almost surely, then $\mathbb{E}(\xi|\mathscr{G}) \geqslant 0$ almost surely.*

*Proof.* By definition we have $\xi - \mathbb{E}(\xi|\mathscr{G}) \perp L^2(\Omega, \mathscr{G})$. If $G \in \mathscr{G}$, then $1_G \in L^2(\Omega, \mathscr{G})$ and therefore

$$\int_\Omega 1_G(\xi - \mathbb{E}(\xi|\mathscr{G}))\, d\mathbb{P} = 0.$$

This gives the desired identity. For the second assertion, choose a $\mathscr{G}$-measurable representative of $g := \mathbb{E}(\xi|\mathscr{G})$ and apply the identity to the $\mathscr{G}$-measurable set $\{g < 0\}$. $\square$

Taking $G = \Omega$ we obtain the identity $\mathbb{E}(\mathbb{E}(\xi|\mathscr{G})) = \mathbb{E}\xi$. This will be used in the lemma, which asserts that the mapping $\xi \mapsto \mathbb{E}(\xi|\mathscr{G})$ is $L^1$-bounded.

**Lemma 11.3.** *For all $\xi \in L^2(\Omega)$ we have $\mathbb{E}|\mathbb{E}(\xi|\mathscr{G})| \leqslant \mathbb{E}|\xi|$.*

*Proof.* It suffices to check that $|\mathbb{E}(\xi|\mathscr{G})| \leqslant \mathbb{E}(|\xi|\,|\mathscr{G})$, since then the lemma follows from $\mathbb{E}|\mathbb{E}(\xi|\mathscr{G})| \leqslant \mathbb{E}\mathbb{E}(|\xi|\,|\mathscr{G}) = \mathbb{E}|\xi|$. Splitting $\xi$ into positive and negative parts, almost surely we have

$$\begin{aligned}|\mathbb{E}(\xi|\mathscr{G})| &= |\mathbb{E}(\xi^+|\mathscr{G}) - \mathbb{E}(\xi^-|\mathscr{G})| \\ &\leqslant |\mathbb{E}(\xi^+|\mathscr{G})| + |\mathbb{E}(\xi^-|\mathscr{G})| = \mathbb{E}(\xi^+|\mathscr{G}) + \mathbb{E}(\xi^-|\mathscr{G}) = \mathbb{E}(|\xi|\,|\mathscr{G}). \quad \square\end{aligned}$$

Since $L^2(\Omega)$ is dense in $L^1(\Omega)$ this lemma shows that the conditional expectation operator has a unique extension to a contractive projection on $L^1(\Omega)$, which we also denote by $\mathbb{E}(\cdot|\mathscr{G})$. This projection is again positive in the sense that it maps positive random variables to positive random variables; this follows from Lemma 11.2 by approximation.

**Lemma 11.4 (Conditional Jensen inequality).** *If $\phi : \mathbb{R} \to \mathbb{R}$ is convex, then for all $\xi \in L^1(\Omega)$ such that $\phi \circ \xi \in L^1(\Omega)$ we have, almost surely,*

$$\phi \circ \mathbb{E}(\xi|\mathscr{G}) \leqslant \mathbb{E}(\phi \circ \xi|\mathscr{G}).$$

*Proof.* If $a, b \in \mathbb{R}$ are such that $at + b \leqslant \phi(t)$ for all $t \in \mathbb{R}$, then the positivity of the conditional expectation operator gives

$$a\mathbb{E}(\xi|\mathscr{G}) + b = \mathbb{E}(a\xi + b|\mathscr{G}) \leqslant \mathbb{E}(\phi \circ \xi|\mathscr{G})$$

almost surely. Since $\phi$ is convex we can find real sequences $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ such that $\phi(t) = \sup_{n\geqslant 1}(a_n t + b_n)$ for all $t \in \mathbb{R}$; we leave the proof of this fact as an exercise. Hence almost surely,

$$\phi \circ \mathbb{E}(\xi|\mathscr{G}) = \sup_{n\geqslant 1} a_n\mathbb{E}(\xi|\mathscr{G}) + b_n \leqslant \mathbb{E}(\phi \circ \xi|\mathscr{G}). \quad \square$$

**Theorem 11.5 ($L^p$-contractivity).** *For all $1 \leqslant p \leqslant \infty$ the conditional expectation operator extends to a contractive positive projection on $L^p(\Omega)$ with range $L^p(\Omega, \mathscr{G})$. For $\xi \in L^p(\Omega)$, the random variable $\mathbb{E}(\xi|\mathscr{G})$ is the unique element of $L^p(\Omega, \mathscr{G})$ with the property that for all $G \in \mathscr{G}$,*

$$\int_G \mathbb{E}(\xi|\mathscr{G})\,d\mathbb{P} = \int_G \xi\,d\mathbb{P}. \tag{11.1}$$

*Proof.* For $1 \leqslant p < \infty$ the $L^p$-contractivity follows from Lemma 11.4 applied to the convex function $\phi(t) = |t|^p$. For $p = \infty$ we argue as follows. If $\xi \in L^\infty(\Omega)$, then $0 \leqslant |\xi| \leqslant \|\xi\|_\infty 1_\Omega$ and therefore $0 \leqslant \mathbb{E}(|\xi|\,|\mathscr{G}) \leqslant \|\xi\|_\infty 1_\Omega$ almost surely. Hence, $\mathbb{E}(|\xi|\,|\mathscr{G}) \in L^\infty(\Omega)$ and $\|\mathbb{E}(|\xi|\,|\mathscr{G})\|_\infty \leqslant \|\xi\|_\infty$.

For $2 \leqslant p \leqslant \infty$, (11.1) follows from Lemma 11.2. For $\xi \in L^p(\Omega)$ with $1 \leqslant p < 2$ we choose a sequence $(\xi_n)_{n=1}^\infty$ in $L^2(\Omega)$ such that $\lim_{n\to\infty} \xi_n = \xi$ in $L^p(\Omega)$. Then $\lim_{n\to\infty} \mathbb{E}(\xi_n|\mathscr{G}) = \mathbb{E}(\xi|\mathscr{G})$ in $L^p(\Omega)$ and therefore, for any $G \in \mathscr{G}$,

$$\int_G \mathbb{E}(\xi|\mathscr{G})\,d\mathbb{P} = \lim_{n\to\infty}\int_G \mathbb{E}(\xi_n|\mathscr{G})\,d\mathbb{P} = \lim_{n\to\infty}\int_G \xi_n\,d\mathbb{P} = \int_G \xi\,d\mathbb{P}.$$

If $\eta \in L^p(\Omega,\mathscr{G})$ satisfies $\int_G \eta\,d\mathbb{P} = \int_G \xi\,d\mathbb{P}$ for all $G \in \mathscr{G}$, then $\int_G \eta\,d\mathbb{P} = \int_G \mathbb{E}(\xi|\mathscr{G})\,d\mathbb{P}$ for all $G \in \mathscr{G}$. Since both $\eta$ and $\mathbb{E}(\xi|\mathscr{G})$ are $\mathscr{G}$-measurable, as in the proof of the second part of Lemma 11.2 this implies that $\eta = \mathbb{E}(\xi|\mathscr{G})$ almost surely.

In particular, $\mathbb{E}(\mathbb{E}(\xi|\mathscr{G})|\mathscr{G}) = \mathbb{E}(\xi|\mathscr{G})$ for all $\xi \in L^p(\Omega)$ and $\mathbb{E}(\xi|\mathscr{G}) = \xi$ for all $\xi \in L^p(\Omega,\mathscr{G})$. This shows that $\mathbb{E}(\cdot|\mathscr{G})$ is a projection onto $L^p(\Omega,\mathscr{G})$. □

The next two results develop some properties of conditional expectations.

**Proposition 11.6.**

(1) *If $\xi \in L^1(\Omega)$ and $\mathscr{H}$ is a sub-$\sigma$-algebra of $\mathscr{G}$, then almost surely*

$$\mathbb{E}(\mathbb{E}(\xi|\mathscr{G})|\mathscr{H}) = \mathbb{E}(\xi|\mathscr{H}).$$

(2) *If $\xi \in L^1(\Omega)$ is independent of $\mathscr{G}$ (that is, $\xi$ is independent of $1_G$ for all $G \in \mathscr{G}$), then almost surely*

$$\mathbb{E}(\xi|\mathscr{G}) = \mathbb{E}\xi.$$

(3) *If $\xi \in L^p(\Omega)$ and $\eta \in L^q(\Omega,\mathscr{G})$ with $1 \leqslant p,q \leqslant \infty$, $\frac{1}{p}+\frac{1}{q} = 1$, then almost surely*

$$\mathbb{E}(\eta\xi|\mathscr{G}) = \eta\mathbb{E}(\xi|\mathscr{G}).$$

*Proof.* (1): For all $H \in \mathscr{H}$ we have $\int_H \mathbb{E}(\mathbb{E}(\xi|\mathscr{G})|\mathscr{H})\,d\mathbb{P} = \int_H \mathbb{E}(\xi|\mathscr{G})\,d\mathbb{P} = \int_H \xi\,d\mathbb{P}$ by Theorem 11.5, first applied to $\mathscr{H}$ and then to $\mathscr{G}$ (observe that $H \in \mathscr{G}$). Now the result follows from the uniqueness part of the theorem.

(2): For all $G \in \mathscr{G}$ we have $\int_G \mathbb{E}\xi\,d\mathbb{P} = \mathbb{E}1_G\mathbb{E}\xi = \mathbb{E}1_G\xi = \int_G \xi\,d\mathbb{P}$, and the result follows from the uniqueness part of Theorem 11.5.

(3): For all $G, G' \in \mathscr{G}$ we have $\int_G 1_{G'}\mathbb{E}(\xi|\mathscr{G})\,d\mathbb{P} = \int_{G\cap G'} \mathbb{E}(\xi|\mathscr{G})\,d\mathbb{P} = \int_{G\cap G'} \xi\,d\mathbb{P} = \int_G \xi 1_{G'}\,d\mathbb{P}$. Hence $\mathbb{E}(\xi 1_{G'}|\mathscr{G}) = 1_{G'}\mathbb{E}(\xi|\mathscr{G})$ by the uniqueness part of Theorem 11.5. By linearity, this gives the result for simple functions $\eta$, and the general case follows by approximation. □

If $\mathscr{C}$ is a collection of subsets of $\Omega$, then $\sigma(\mathscr{C})$ denotes the $\sigma$-algebra generated by $\mathscr{C}$, that is, the smallest $\sigma$-algebra in $\Omega$ which contains all sets of $\mathscr{C}$. In this context we shall use self-explanatory notations such as $\mathbb{E}(\xi|\mathscr{C}) := \mathbb{E}(\xi|\sigma(\mathscr{C}))$ and $\mathbb{E}(\xi|\mathscr{C}_1,\mathscr{C}_2) := \mathbb{E}(\xi|\sigma(\mathscr{C}_1 \cup \mathscr{C}_2))$.

If $\eta : \Omega \to E$ is a random variable, then $\sigma(\eta)$ denotes the $\sigma$-algebra $\{\eta^{-1}(B) : B \in \mathscr{B}(E)\}$. This is the smallest $\sigma$-algebra in $\Omega$ with respect to which $\eta$ is Borel measurable. Again, notations such as $\mathbb{E}(\xi|\eta) := \mathbb{E}(\xi|\sigma(\eta))$ and $\mathbb{E}(\xi|\eta_1,\eta_2) := \mathbb{E}(\xi|\sigma(\eta_1,\eta_2))$ are self-explanatory.

**Proposition 11.7.** *Let $\mathscr{G}$ and $\mathscr{H}$ be sub-$\sigma$-algebras of $\mathscr{F}$, let $\xi \in L^1(\Omega)$, and suppose that $\mathscr{H}$ is independent of $\sigma(\xi,\mathscr{G})$. Then, almost surely,*

$$\mathbb{E}(\xi|\mathscr{G},\mathscr{H}) = \mathbb{E}(\xi|\mathscr{G}).$$

*Proof.* First we claim that $\sigma(\mathscr{G},\mathscr{H})$ is generated by the collection $\mathscr{C}$ of all sets of the form $G\cap H$ with $G\in\mathscr{G}$ and $H\in\mathscr{H}$. Indeed, from $G=G\cap\Omega\in\mathscr{C}$ and $H=\Omega\cap H\in\mathscr{C}$ we see that $\mathscr{C}$ contains both $\mathscr{G}$ and $\mathscr{H}$.

Next, from $(G_1\cap H_1)\cap(G_2\cap H_2)=(G_1\cap G_2)\cap(H_1\cap H_2)$ it follows that $\mathscr{C}$ is closed under taking finite intersections. This being said, the strategy is to apply Dynkin's lemma. By considering positive and negative parts separately we may assume that $\xi\geqslant 0$ almost surely. Also we may assume that $\mathbb{E}\xi>0$, since otherwise there is nothing to prove.

For $G\cap H\in\mathscr{C}$ we have

$$\begin{aligned}\int_{G\cap H}\mathbb{E}(\xi|\mathscr{G},\mathscr{H})\,d\mathbb{P} &= \int_{G\cap H}\xi\,d\mathbb{P}=\mathbb{E}(1_G 1_H\xi)\\ &\overset{\text{(i)}}{=} \mathbb{E}1_H\mathbb{E}(1_G\xi)=\mathbb{E}1_H\mathbb{E}(1_G\mathbb{E}(\xi|\mathscr{G}))\\ &\overset{\text{(ii)}}{=} \mathbb{E}(1_G1_H\mathbb{E}(\xi|\mathscr{G}))=\int_{G\cap H}\mathbb{E}(\xi|\mathscr{G})\,d\mathbb{P}.\end{aligned}$$

In (i) and(ii) we used the independence of $\mathscr{H}$ and $\sigma(\xi,\mathscr{G})$. By Dynkin's lemma, applied to the probability measures $\mu(C):=\frac{1}{\mathbb{E}\xi}\int_C\mathbb{E}(\xi|\mathscr{G},\mathscr{H})\,d\mathbb{P}$ and $\nu(C):=\frac{1}{\mathbb{E}\xi}\int_C\mathbb{E}(\xi|\mathscr{G})\,d\mathbb{P}$ it follows that $\mu=\nu$ on $\sigma(\mathscr{C})=\sigma(\mathscr{G},\mathscr{H})$. This means that

$$\int_C\mathbb{E}(\xi|\mathscr{G},\mathscr{H})\,d\mathbb{P}=\int_C\mathbb{E}(\xi|\mathscr{G})\,d\mathbb{P}\quad\forall C\in\sigma(\mathscr{G},\mathscr{H})$$

and the result follows. □

## 11.2 Vector-valued conditional expectations

Our next aim is to extend the conditional expectation operators from $L^p(\Omega)$ to $L^p(\Omega;E)$, where $E$ is an arbitrary Banach space.

Let us fix $1\leqslant p<\infty$ for the moment and let $(A,\mathscr{A},\mu)$ be an arbitrary $\sigma$-finite measure space. For a Banach space $E$ we denote by $L^p(A)\otimes E$ the linear span of all functions of the form $f\otimes x$ with $f\in L^p(A)$ and $x\in E$.

**Lemma 11.8.** *$L^p(A)\otimes E$ is dense in $L^p(A;E)$.*

*Proof.* It has been observed in Lecture 1 that the $\mu$-simple functions are dense in $L^p(A;E)$. Clearly these belong to $L^p(A)\otimes E$. □

Suppose next that a bounded linear operator $T$ on $L^p(A)$ is given. We may define a linear operator $T\otimes I$ on $L^p(A)\otimes E$ by the formula

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

We leave it to the reader to check that the resulting linear operator on $L^p(A)\otimes E$ is well-defined. In view of Lemma 11.8 one may now ask whether $T\otimes I$ extends to a bounded operator on $L^p(A;E)$. Unfortunately, without additional assumptions this is generally not the case. For *positive* operators $T$ on $L^p(A)$ we have the following result.

**Proposition 11.9.** *If $T$ is a positive operator on $L^p(A)$, then $T \otimes I$ extends uniquely to a bounded operator on $L^p(A;E)$ and we have*

$$\|T \otimes I\| = \|T\|.$$

*Proof.* Let $g \in L^p(A) \otimes E$ be a $\mu$-simple function, say $g = \sum_{n=1}^N 1_{A_n} \otimes x_n$ with the sets $A_n \in \mathscr{A}$ mutually disjoint. Then from the positivity of $T$ we have $|T1_A| = T1_{A_n}$ and we obtain the estimates

$$\begin{aligned}
\Big\|(T \otimes I) \sum_{n=1}^N 1_{A_n} \otimes x_n\Big\|^p_{L^p(A;E)} &= \int_A \Big\| \sum_{n=1}^N T1_{A_n} \otimes x_n \Big\|^p d\mu \\
&\leqslant \int_A \Big| \sum_{n=1}^N |T1_{A_n}| \|x_n\| \Big|^p d\mu \\
&= \int_A \Big| T \sum_{n=1}^N 1_{A_n} \|x_n\| \Big|^p d\mu \\
&\leqslant \|T\|^p \int_A \Big| \sum_{n=1}^N 1_{A_n} \|x_n\| \Big|^p d\mu \\
&= \|T\|^p \Big\| \sum_{n=1}^N 1_{A_n} \otimes x_n \Big\|^p_{L^p(A;E)}.
\end{aligned}$$

Since the $\mu$-simple functions are dense in $L^p(A;E)$, this proves that $T \otimes I$ has a unique extension to a bounded operator on $L^p(A;E)$ of norm $\|T \otimes I\| \leqslant \|T\|$. Equality $\|T \otimes I\| = \|T\|$ is obtained by considering functions of the form $f \otimes x$ with $f \in L^p(A)$ and $x \in E$ of norm one. $\square$

Returning to conditional expectations we obtain the following result:

**Theorem 11.10.** *For $1 \leqslant p \leqslant \infty$ the operator $\mathbb{E}(\cdot|\mathscr{G}) \otimes I$ extends uniquely to a contractive projection on $L^p(\Omega;E)$ with range $L^p(\Omega, \mathscr{G}; E)$. For all $X \in L^p(\Omega;E)$, the random variable*

$$\mathbb{E}(X|\mathscr{G}) := (\mathbb{E}(\cdot|\mathscr{G}) \otimes I)X$$

*is the unique element of $L^p(\Omega, \mathscr{G}; E)$ with the property that for all $G \in \mathscr{G}$,*

$$\int_G \mathbb{E}(X|\mathscr{G})\, d\mathbb{P} = \int_G X\, d\mathbb{P}.$$

*Proof.* For $1 \leqslant p < \infty$ the $L^p$-contractivity follows from Proposition 11.9.

Before continuing with the case $p = \infty$, let us note that for a simple random variable of the form $X = \sum_{n=1}^N 1_{A_n} \otimes x_n$ with disjoint sets $A_n \in \mathscr{F}$ we have

$$\|\mathbb{E}(X|\mathscr{G})\| = \Big\| \sum_{n=1}^{N} \mathbb{E}(1_{A_n}|\mathscr{G}) \otimes x_n \Big\| \leqslant \sum_{n=1}^{N} \mathbb{E}(1_{A_n}|\mathscr{G})\|x_n\| = \mathbb{E}(\|X\|\,|\mathscr{G}).$$

By a density argument, this inequality extends to arbitrary random variables $X \in L^1(\Omega; E)$. By inclusion, this implies the corresponding inequality for random variables $X \in L^p(\Omega; E)$, $1 \leqslant p \leqslant \infty$.

Next let $X \in L^\infty(\Omega; E)$. Then $\|X\| \in L^\infty(\Omega)$, and by the inequality which has just been proved together with the contractivity of the conditional expectation in $L^\infty(\Omega)$ we obtain

$$\|\mathbb{E}(X|\mathscr{G})\|_{L^\infty(\Omega;E)} \leqslant \|\mathbb{E}(\|X\|\,|\mathscr{G})\|_{L^\infty(\Omega)} \leqslant \big\|\|X\|\big\|_{L^\infty(\Omega)} = \|X\|_{L^\infty(\Omega;E)}.$$

This proves that the conditional expectation is a contraction in $L^\infty(\Omega; E)$.

For simple random variables $X$, the identity $\int_G \mathbb{E}(X|\mathscr{G})\,d\mathbb{P} = \int_G X\,d\mathbb{P}$ follows from the corresponding assertion in the scalar case. By density, the identity extends to random variables $X \in L^1(\Omega; E)$, and hence for $X \in L^p(\Omega; E)$, $1 \leqslant p \leqslant \infty$. The uniqueness assertion follows from the scalar case via Corollary 1.14. $\square$

We leave it to the reader to check that Propostions 11.6 and 11.7 extend to the vector-valued setting and finish this section with two important examples. We have already encountered the first example in the proof of Theorem 6.17.

*Example 11.11 (Averaging).* Consider a decomposition $(0, 1) = \bigcup_{n=1}^{N} I_n$, where the $I_n$ are disjoint intervals with Lebesgue measure $|I_n| > 0$. Let $\mathscr{F}$ be the Borel $\sigma$-algebra of $(0, 1)$ and let $\mathscr{G}$ be the $\sigma$-algebra generated by the intervals $I_n$. Let $E$ be a Banach space. Then for all $f \in L^1(0, 1; E)$ we have

$$\mathbb{E}(f|\mathscr{G}) = \sum_{n=1}^{N} c_n 1_{I_n} \quad \text{with} \quad c_n = \frac{1}{|I_n|} \int_{I_n} f(t)\,dt.$$

This is verified by checking the condition of Theorem 11.5.

*Example 11.12 (Sums of independent random variables).* Let $(\xi_n)_{n=1}^\infty$ be a sequence of independent integrable $E$-valued random variables. For each $n \geqslant 1$ let $\mathscr{F}_n := \sigma(\xi_1, \ldots, \xi_n)$ and put $S_n = \xi_1 + \cdots + \xi_n$. Then for all $N \geqslant n \geqslant 1$ we have

$$\mathbb{E}(S_N|\mathscr{F}_n) = S_n + \mathbb{E}(\xi_{n+1}) + \cdots + \mathbb{E}(\xi_N).$$

This follows from Proposition 11.6 (1), (2). In particular, if the $\xi_n$ are centred,

$$\mathbb{E}(S_N|\mathscr{F}_n) = S_n.$$

## 11.3 Martingales

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and $(I, \leqslant)$ a partially ordered set, that is, a set $I$ with a relation $\leqslant$ which satisfies the following properties:

(1) $i \leqslant i$ for all $i \in I$;
(2) $i \leqslant j$ and $j \leqslant i$ imply $i = j$;
(3) $i \leqslant j$ and $j \leqslant k$ imply $i \leqslant k$.

**Definition 11.13.** *Let $I$ be a partially ordered set. A* filtration *with index set $I$ is a family $(\mathscr{F}_i)_{i\in I}$ of sub-$\sigma$-algebras of $\mathscr{F}$ such that $\mathscr{F}_i \subseteq \mathscr{F}_j$ whenever $i \leqslant j$ in $I$. A family $(X_i)_{i\in I}$ of $E$-valued random variables is* adapted *to the filtration $(\mathscr{F}_i)_{i\in I}$ if each $X_i$ is strongly $\mathscr{F}_i$-measurable.*

In this definition the random variables are defined pointwise. The definitions carry over to equivalence classes modulo null sets, provided one replaces 'is strongly $\mathscr{F}_i$-measurable' with 'has a strongly $\mathscr{F}_i$-measurable representative' in the definition of adaptedness.

Every family $X = (X_i)_{i\in I}$ is adapted to the filtration $(\mathscr{F}_i^X)_{i\in I}$, where $\mathscr{F}_i^X := \sigma(X_j : \ j \leqslant i)$. We call this filtration the *filtration generated by $X$*.

**Definition 11.14.** *A family $(M_i)_{i\in I}$ of integrable $E$-valued random variables is an $E$-valued* martingale *with respect to a filtration $(\mathscr{F}_i)_{i\in I}$ if it is adapted with respect to $(\mathscr{F}_i)_{i\in I}$ and*

$$\mathbb{E}(M_j|\mathscr{F}_i) = M_i$$

*almost surely whenever $i \leqslant j$ in $I$. If in addition $\mathbb{E}\|M_i\|^p < \infty$ for all $i \in I$, then we call $(M_i)_{i\in I}$ an $E$-valued $L^p$-*martingale*.*

*Example 11.15.* Let $X \in L^1(\Omega; E)$ be given. For any filtration $(\mathscr{F}_i)_{i\in I}$, the family $(X_i)_{i\in I}$ defined by

$$X_i := \mathbb{E}(X|\mathscr{F}_i)$$

is a martingale with respect to $(\mathscr{F}_i)_{i\in I}$; this follows from Proposition 11.6 (1).

In most examples, $I$ is a finite or infinite interval in $\mathbb{Z}$ or $\mathbb{R}$. In these cases one speaks of *discrete time martingales* and *continuous time martingales*. Here are two examples.

*Example 11.16 (Sums of independent random variables).* If $X = (X_n)_{n=1}^\infty$ is a sequence of independent integrable $E$-valued random variables satisfying $\mathbb{E}X_n = 0$ for all $n \geqslant 1$, then the partial sum sequence $(S_n)_{n=1}^\infty$ is a martingale with respect to the filtration $(\mathscr{F}_n^X)_{n=1}^\infty$, where

$$\mathscr{F}_n^X := \sigma(X_1, \dots, X_n).$$

This is immediate from Example 11.12.

*Example 11.17 (Brownian motion).* Every Brownian motion $(W(t))_{t\in[0,T]}$ is a martingale with respect to the filtration $(\mathscr{F}_t^W)_{t\in[0,T]}$ defined by

$$\mathscr{F}_t^W := \sigma(W(s) : \ 0 \leqslant s \leqslant t).$$

Adaptedness and integrability being clear, it remains to show that

$$\mathbb{E}(W(t)|\mathscr{F}_s^W) = W(s)$$

almost surely for all $0 \leqslant s \leqslant t \leqslant T$. Writing $W(t) = W(s) + (W(t) - W(s))$, and noting that $W(s)$ is $\mathscr{F}_s^W$-measurable and $W(t) - W(s)$ is independent of $\mathscr{F}_s^W$ (this will be proved in a moment), with Proposition 11.6 (2), we obtain

$$\mathbb{E}(W(t)|\mathscr{F}_s^W) = W(s) + \mathbb{E}(W(t) - W(s)) = W(s)$$

almost surely.

That $W(t) - W(s)$ is independent of $\mathscr{F}_s^W$ is a consequence of the following general observation:

**Lemma 11.18.** *A random variable $X$ is independent of the family $(Y_j)_{j\in J}$ if and only if $X$ is independent of the $\sigma$-algebra generated by $(Y_j)_{j\in J}$.*

*Proof.* Let $X$ take its values in $E$ and each $Y_j$ in $E_j$.

We begin with the 'only if' part. Suppose that $X$ is independent of $(Y_j)_{j\in J}$. By definition this means that $X$ is independent of $(Y_{j_1}, \dots, Y_{j_N})$ for all $j_1, \dots, j_N \in J$. In particular,

$$\mathbb{P}(\{X \in B\} \cap C) = \mathbb{P}\{X \in B\}\mathbb{P}(C) \tag{11.2}$$

for all sets $C = \{Y_{j_1} \in B_1, \dots, Y_{j_N} \in B_N\}$. The collection of all such sets $C$, which we shall denote by $\mathscr{C}$, is closed under taking finite intersections and generates $\sigma((Y_j)_{j\in J})$. We must show that (11.2) holds for all $C \in \sigma((Y_j)_{j\in J})$. Fix $B \in \mathscr{B}(E)$ and assume without loss of generality that $\mathbb{P}\{X \in B\} > 0$. Consider the probability measure $\mathbb{P}_B$ on $(\Omega, \mathscr{F})$ defined by

$$\mathbb{P}_B(F) := \frac{\mathbb{P}(\{X \in B\} \cap F)}{\mathbb{P}\{X \in B\}}.$$

The measures $\mathbb{P}_B$ and $\mathbb{P}$ coincide on $\mathscr{C}$, and therefore they coincide on $\sigma(\mathscr{C}) = \sigma((Y_j)_{j\in J})$ by Dynkin's lemma.

To prove the 'if' part it suffices to observe that the sets $\{(Y_{j_1}, \dots, Y_{j_N}) \in B\})$ with $B \in \mathscr{B}(E_{j_1} \times \cdots \times E_{j_N})$ belong to $\sigma((Y_j)_{j\in J})$. $\square$

More generally, this argument can be used to show that two families $(X_i)_{i\in I}$ and $(Y_j)_{j\in J}$ are independent of each other if and only if $\sigma((X_i)_{i\in I})$ and $\sigma((Y_j)_{j\in J})$ are independent.

Our final example will be used in the next lecture.

*Example 11.19 (Martingale transforms).* A real-valued sequence $v = (v_n)_{n=1}^N$ is said to be *predictable* with respect to a filtration $(\mathscr{F}_n)_{n=1}^N$ if $v_n$ is $\mathscr{F}_{n-1}$-measurable for $n = 1, \dots, N$ (with the understanding that $\mathscr{F}_0 = \{\varnothing, \Omega\}$, so $v_1$ is constant almost surely). If $M = (M_n)_{n=1}^N$ is an $E$-valued martingale with respect to $(\mathscr{F}_n)_{n=1}^N$, the sequence $v * M = ((v * M)_n)_{n=1}^N$ defined by

$$(v * M)_n := \sum_{j=1}^{n} v_j (M_j - M_{j-1}), \quad n = 1, \dots, N$$

(with the understanding that $M_0 = 0$) is called the *martingale transform* of $M$ by $v$.

The intuitive meaning is as follows. Suppose the increment $M_j - M_{j-1}$ represents the outcome of the $j$-th gambling game. The assumption that $M$ is a martingale means that the game is fair. Let $v_j$ be the stake a player puts on this game. The requirement that $v_j$ be $\mathscr{F}_{j-1}$-measurable means that the stake has to be decided knowing the outcomes of the first $j-1$ games only. The random variable $(v * M)_n$ then represents the total winnings after game $n$. An obvious question is whether the player can devise a favourable strategy. Under a mild additional assumption the answer is 'no': if the $v_n$ are bounded, then $v * M$ is a martingale with respect to $(\mathscr{F}_n)_{n=1}^N$. Let us prove this. Clearly, $v * M$ is adapted with respect to $(\mathscr{F}_n)_{n=1}^N$ and the random variables $v_n(M_n - M_{n-1})$ are integrable. By Proposition 11.6 (3) and the $\mathscr{F}_{n-1}$-measurability of $(v * M)_{n-1}$ and $v_n$,

$$\mathbb{E}((v * M)_n)|\mathscr{F}_{n-1}) = (v * M)_{n-1} + v_n \mathbb{E}(M_n - M_{n-1}|\mathscr{F}_{n-1}) = (v * M)_{n-1}.$$

## 11.4 $L^p$-martingales

An important inequality for $L^p$-martingales, due to Doob, states that for $1 < p < \infty$ the maximum of an $L^p$-martingale is in $L^p$ again.

Let $M = (M_n)_{n=1}^N$ be an $E$-valued martingale with respect to $\mathbb{F} = (\mathscr{F}_n)_{n=1}^N$ and define $M_N^* : \Omega \to \mathbb{R}_+$ by

$$M_N^* := \max_{1 \leqslant n \leqslant N} \|M_n\|.$$

**Theorem 11.20 (Doob).** *For all $r > 0$ we have*

$$\mathbb{P}\{M_N^* > r\} \leqslant \frac{1}{r} \mathbb{E}\|M_N\|.$$

*If $1 < p < \infty$ and $M_N \in L^p(\Omega; E)$, then $M_N^* \in L^p(\Omega)$ and*

$$\|M_N^*\|_p \leqslant \frac{p}{p-1} \|M_N\|_p.$$

*Proof.* The proof proceeds in two steps.

*Step 1* – We claim that for all $r > 0$,

$$r\mathbb{P}\{M_N^* > r\} \leqslant \mathbb{E}(1_{\{M_N^* > r\}} \|M_N\|). \tag{11.3}$$

This implies the first inequality.

Let us fix $r > 0$ and define $\tau : \Omega \to \{1, \ldots, N+1\}$ by $\tau := \min\{1 \leqslant n \leqslant N : \ \|M_n\| > r\}$ with the convention that $\min \varnothing := N+1$. Then $\{M_N^* > r\} = \{\tau \leqslant N\}$. On the set $\{\tau = n\}$ we have $\|M_n\| > r$ and therefore

$$
\begin{aligned}
r\mathbb{P}\{M_N^* > r\} = r\sum_{n=1}^{N} \mathbb{P}\{\tau = n\} &\leqslant \sum_{n=1}^{N} \mathbb{E}(1_{\{\tau=n\}}\|M_n\|) \\
&\overset{(*)}{\leqslant} \sum_{n=1}^{N} \mathbb{E}(1_{\{\tau=n\}}\|M_N\|) = \mathbb{E}(1_{\{\tau\leqslant N\}}\|M_N\|) \\
&= \mathbb{E}(1_{\{M_N^*>r\}}\|M_N\|)
\end{aligned}
$$

which gives (11.3). The inequality $(*)$ follows from the martingale property, since almost surely we have

$$\|M_n\| = \|\mathbb{E}(M_N|\mathscr{F}_n)\| \leqslant \mathbb{E}(\|M_N\||\mathscr{F}_n).$$

*Step 2* – Next let $1 < p < \infty$ and assume that $\|M_N\|_p < \infty$. We may assume that $\|M_N^*\|_p > 0$, since otherwise there is nothing to prove. Integrating by parts and using (11.3) and Hölder's inequality,

$$
\begin{aligned}
\|M_N^*\|_p^p = \int_0^\infty pr^{p-1}\mathbb{P}\{M_N^* > r\}\, dr &\leqslant \int_0^\infty pr^{p-2}\mathbb{E}(1_{\{M_N^*>r\}}\|M_N\|)\, dr \\
&= \mathbb{E}\Big(\|M_N\| \int_0^{M_N^*} pr^{p-2}\, dr\Big) \\
&= \frac{p}{p-1}\mathbb{E}(\|M_N\|(M_N^*)^{p-1}) \\
&\leqslant \frac{p}{p-1}\|M_N\|_p\|M_N^*\|_p^{p-1}.
\end{aligned}
$$

The result follows upon dividing both sides by $\|M_N^*\|_p^{p-1}$. □

We shall apply the first part of Doob's inequality to prove the following result on convergence of certain $L^p$-martingales.

Suppose a filtration $(\mathscr{F}_n)_{n=1}^\infty$ is given on $(\Omega, \mathscr{F}, \mathbb{P})$. We denote by $\mathscr{F}_\infty$ the $\sigma$-algebra generated by $(\mathscr{F}_n)_{n=1}^\infty$, that is, $\mathscr{F}_\infty$ is the smallest $\sigma$-algebra containing each of the $\mathscr{F}_n$.

**Theorem 11.21.** *Let $1 \leqslant p < \infty$ and assume that $X \in L^p(\Omega; E)$. Then,*

$$\lim_{n\to\infty} \mathbb{E}(X|\mathscr{F}_n) = \mathbb{E}(X|\mathscr{F}_\infty)$$

*both in $L^p(\Omega; E)$ and almost surely.*

*Proof.* We claim that $\bigcup_{n=1}^\infty L^p(\Omega, \mathscr{F}_n; E)$ is dense in $L^p(\Omega, \mathscr{F}_\infty; E)$. Assuming this for the moment we first show how the $L^p$-convergence is obtained from this.

For all $Y \in L^p(\Omega, \mathscr{F}_m; E)$ and $n \geqslant m$ we have $\mathbb{E}(Y|\mathscr{F}_n) = \mathbb{E}(Y|\mathscr{F}_\infty) = Y$, and therefore we trivially have $\lim_{n\to\infty} \mathbb{E}(Y|\mathscr{F}_n) = Y$ in $L^p(\Omega, \mathscr{F}_\infty; E)$. Since the conditional operators are contractive, it follows that $\lim_{n\to\infty} \mathbb{E}(Y|\mathscr{F}_n) = Y$ in $L^p(\Omega, \mathscr{F}_\infty; E)$ for all $Y \in L^p(\Omega, \mathscr{F}_\infty; E)$. In particular this is true for $Y = \mathbb{E}(X|\mathscr{F}_\infty)$.

Let us next prove that $\bigcup_{n=1}^\infty L^p(\Omega, \mathscr{F}_n; E)$ is dense in $L^p(\Omega, \mathscr{F}_\infty; E)$. Let $\mathscr{G}$ be the collection of all sets $G \in \mathscr{F}_\infty$ with the property that for all $\varepsilon > 0$ there exists an $n \geqslant 1$ and a set $F \in \mathscr{F}_n$ such that $\mathbb{P}(F \Delta G) < \varepsilon$. Here, $F \Delta G = (F \setminus G) \cup (G \setminus F)$ is the symmetric difference of $F$ and $G$. It is easily checked that the collection of all approximable sets is a sub-$\sigma$-algebra of $\mathscr{F}_\infty$. Clearly, this $\sigma$-algebra contains each $\mathscr{F}_n$, and therefore it contains $\mathscr{F}_\infty$.

By what we have shown so far, $G \in \mathscr{F}_\infty$ implies that $1_G = \lim_{k\to\infty} 1_{G_k}$ in $L^p(\Omega; E)$, where $G_k \in \mathscr{F}_{n_k}$ for some $n_k \geqslant 1$. It follows that every simple function of $L^p(\Omega, \mathscr{F}_\infty; E)$ is contained in the closure of $\bigcup_{n=1}^\infty L^p(\Omega, \mathscr{F}_n; E)$ in $L^p(\Omega, \mathscr{F}_\infty; E)$. As a consequence, all of $L^p(\Omega, \mathscr{F}_\infty; E)$ is contained in the closure of $\bigcup_{n=1}^\infty L^p(\Omega, \mathscr{F}_n; E)$.

So far we have proved the $L^p$-convergence. To prove the almost sure convergence, note that by the first part of Theorem 11.20 and monotone convergence we have

$$\mathbb{P}\Big\{ \sup_{n\geqslant 1} \|M_n\| > r \Big\} \leqslant \frac{1}{r} \sup_{n\geqslant 1} \mathbb{E}\|M_n\|,$$

where we put $M_n := \mathbb{E}(X|\mathscr{F}_n)$ for brevity. Applying this with $X$ replaced by $X - M_N$, for all $n \geqslant N$ we obtain

$$\mathbb{P}\Big\{ \sup_{n\geqslant N} \|M_n - M_N\| > r \Big\} \leqslant \frac{1}{r} \sup_{n\geqslant N} \mathbb{E}\|M_n - M_N\|.$$

By what we have proved already we find indices $N_1 < N_2 < \dots$ such that

$$\sup_{n\geqslant N_k} \mathbb{E}\|M_n - M_{N_k}\| < \frac{1}{2^{2k}}.$$

With $r = 1/2^k$ this gives

$$\mathbb{P}\Big\{ \sup_{n\geqslant N_k} \|M_n - M_{N_k}\| > \frac{1}{2^k} \Big\} \leqslant \frac{1}{2^k}.$$

The Borel-Cantelli lemma now implies that $\lim_{n\to\infty} M_n = M_\infty = \mathbb{E}(X|\mathscr{F}_\infty)$ almost surely. □

## 11.5 Exercises

1. Let $f, g$ be random variables on $\Omega$. Prove that if $f$ is $\sigma(g)$-measurable, then $f = \phi \circ g$ for some Borel function $\phi$.
   *Hint:* First suppose that $f = 1_A$ with $A \in \sigma(g)$.

2. a) Let $X_1, \dots, X_N$ be independent and identically distributed integrable $E$-valued random variables and put $S_N = X_1 + \dots + X_N$. Show that $\mathbb{E}(X_1|S_N) = \dots = \mathbb{E}(X_N|S_N)$ and deduce that $\mathbb{E}(X_n|S_N) = S_N/N$ for all $n = 1, \dots, N$.
   b) (!) Let $X$ and $Y$ be independent and identically distributed integrable $E$-valued random variables on $\Omega$. Prove that $\mathbb{E}(X - Y|X + Y) = 0$.
3. Let $(M_i)_{i\in I}$ be a martingale with respect to the filtration $(\mathscr{F}_i)_{i\in I}$, and let $(\mathscr{G}_i)_{i\in I}$ be a filtration which is independent of $(\mathscr{F}_i)_{i\in I}$. Define the filtration $(\mathscr{H}_i)_{i\in I}$ by $\mathscr{H}_i := \sigma(\mathscr{F}_i, \mathscr{G}_i)$. Show that $(M_i)_{i\in I}$ is a martingale with respect to $(\mathscr{H}_i)_{i\in I}$.
4. Let $W_H$ be an $H$-cylindrical Brownian motion. The filtration $(\mathscr{F}_t^{W_H})_{t\in[0,T]}$ generated by $W_H$ is defined by $\mathscr{F}_t^{W_H} := \sigma(W_H(s)h : \ s \in [0, t], \ h \in H)$.
   a) Show that for all $h \in H$ and $0 \leqslant s \leqslant t \leqslant T$ the increment $W_H(t)h - W_H(s)h$ is independent of $\mathscr{F}_s^{W_H}$.
   b) Show that for all $h \in H$ the Brownian motion $(W(t)h)_{t\in[0,T]}$ is a martingale with respect to $(\mathscr{F}_t^{W_H})_{t\in[0,T]}$.
5. This exercise is a continuation of Exercise 6.3 on averaging operators. Using the notations introduced there, show that for all $f \in L^p(0, T; E)$ we have $\lim_{n\to\infty} A_n f = f$ almost everywhere.

**Notes.** An elementary introduction to the theory of martingales is the book by Williams [109]; for more comprehensive treatments we refer to Kallenberg [55] and Rogers and Williams [95]. A systematic account of the vector-valued theory can be found in Diestel and Uhl [36].

The results of Sections 11.1 and 11.2 are standard. The approach taken in Section 11.1 by first defining conditional expectations in $L^2(\Omega)$ by an orthogonal projection is the most elementary one and, in our opinion, the most intuitive. A shorter, but less elementary approach is to *define* conditional expectations in $L^1(\Omega)$ by the identity (11.1) and then to use the Radon-Nikodým theorem to prove their existence and uniqueness.

For further results on vector-valued extensions of positive operators we refer to the nice paper by Haase [46].

The proof of Doob's inequality (Theorem 11.20) is standard and can be found in many textbooks. It only requires the fact that $(\|M_n\|)_{n=1}^N$ is a non-negative submartingale, that is, it satisfies $\|M_n\| \leqslant \mathbb{E}(\|M_n\||\mathscr{F}_m)$ almost surely for all $m \leqslant n$.

The proof of the martingale convergence theorem (Theorem 11.21) is taken from [36]. In the scalar theory it is true that any $L^1$-bounded martingale converges almost surely, with convergence in $L^1$ if the martingale is uniformly integrable (which is the case, e.g., if the martingale is $L^p$-bounded for some $1 < p < \infty$). For Banach space-valued martingales $E$, the same result holds if $E$ has the so-called Radon-Nikodým property. Examples of spaces with this property are reflexive spaces and separable dual spaces. We refer to [36] for the full story.

It is worth mentioning the following result of Davis, Ghoussoub, Johnson, Kwapień, Maurey [30], which generalises the Itô-Nisio theorem to $E$-valued martingales:

**Theorem 11.22.** *Let $E$ be an arbitrary Banach space and suppose that $(M_n)_{n=1}^{\infty}$ is an $L^1$-bounded $E$-valued martingale. For an $E$-valued random variable $M$ the following assertions are equivalent:*

(1) *For all $x^* \in E^*$ we have $\lim_{n\to\infty}\langle M_n, x^*\rangle = \langle M, x^*\rangle$ almost surely;*
(2) *For all $x^* \in E^*$ we have $\lim_{n\to\infty}\langle M_n, x^*\rangle = \langle M, x^*\rangle$ in probability;*
(3) $\lim_{n\to\infty} M_n = M$ *almost surely;*
(4) $\lim_{n\to\infty} M_n = M$ *in probability.*

*If $M \in L^p(\Omega; E)$ for some $1 \leqslant p < \infty$, then $M_n \in L^p(\Omega; E)$ for all $n \geqslant 1$ and we have $\lim_{n\to\infty} M_n = M$ in $L^p(\Omega; E)$.*

Note that the Itô-Nisio theorem holds without any integrability conditions. It is clear that in the above theorem we need to impose integrability of the random variables $M_n$ in order to define the their conditional expectations. In [30] a simple example is given which shows that even the $L^1$-boundedness condition on the $M_n$ cannot be omitted.

# 12

# UMD-spaces

This lecture is devoted to the study of a class of Banach spaces, the so-called UMD-spaces, which share many of the good properties of Hilbert spaces and is sufficiently broad to include $L^p$-spaces for $1 < p < \infty$.

Experience has shown that the class of UMD-spaces is precisely the 'right' one for pursuing vector-valued stochastic analysis as well as vector-valued harmonic analysis. Indeed, many classical Hilbert space-valued results from both areas can be extended to the UMD-valued case, and often this fact characterises the UMD-property.

The relevant fact for our purposes is that the UMD-spaces are those Banach spaces $E$ in which the Wiener integral of Lecture 6 can be extended from $\mathscr{L}(H,E)$-valued functions to $\mathscr{L}(H,E)$-valued stochastic processes. This is the subject matter of the next lecture. In the present lecture, we define UMD-spaces in terms of $L^p$-bounds for signed $E$-valued martingale difference sequences and study some of their elementary properties. At first sight, the definition of the UMD-property depends on the parameter $1 < p < \infty$. It is a deep result of Maurey and Pisier that the UMD-property is independent of $1 < p < \infty$. This theorem, which is proved in detail, enables us to prove that $L^p$-spaces are UMD-spaces for $1 < p < \infty$.

## 12.1 UMD$_p$-spaces

We begin with a definition.

**Definition 12.1.** *Let $(M_n)_{n=1}^N$ be an $E$-valued martingale. The sequence $(d_n)_{n=1}^N$ defined by $d_n := M_n - M_{n-1}$ (with the understanding that $M_0 = 0$) is called the* martingale difference sequence *associated with $(M_n)_{n=1}^N$.*

We call $(d_n)_{n=1}^N$ an *$L^p$-martingale difference sequence* if it is the difference sequence of an $L^p$-martingale.

If $(M_n)_{n=1}^N$ is a martingale with respect to the filtration $(\mathscr{F}_n)_{n=1}^N$, then $(d_n)_{n=1}^N$ is adapted to $(\mathscr{F}_n)_{n=1}^N$ and $\mathbb{E}(d_n|\mathscr{F}_m) = 0$ for $1 \leqslant m < n \leqslant N$.

It is easy to see that these two properties characterise martingale difference sequences.

**Proposition 12.2.** *Every $L^2$-martingale difference sequence with values in a Hilbert space $H$ is orthogonal in $L^2(\Omega; H)$.*

*Proof.* We use the notations introduced above. For $1 \leqslant m < n \leqslant N$, from $\mathbb{E}d_n = \mathbb{E}(\mathbb{E}(d_n|\mathscr{F}_{n-1})) = 0$ we deduce that

$$\mathbb{E}[d_m, d_n] = \mathbb{E}(\mathbb{E}([d_m, d_n]|\mathscr{F}_{n-1})) = \mathbb{E}([d_m, \mathbb{E}(d_n|\mathscr{F}_{n-1})]) = 0.$$

The second identity follows from Proposition 11.6 (3) if $d_m$ is replaced by a random variable in $g \in L^2(\Omega, \mathscr{F}_m) \otimes H$, and the general case follows from this since $L^2(\Omega, \mathscr{F}_m) \otimes H$ is dense in $L^2(\Omega, \mathscr{F}_m; H)$. $\square$

This suggests that in the context of stochastic analysis in Banach spaces, martingale difference sequences provide a substitute for orthogonal sequences. To formalise this idea we note that in the situation of Proposition 12.2, for any choice of signs $\varepsilon_n = \pm 1$ we have

$$\mathbb{E}\Big\|\sum_{n=1}^{N} \varepsilon_n d_n\Big\|^2 = \mathbb{E}\Big\|\sum_{n=1}^{N} d_n\Big\|^2. \tag{12.1}$$

It is this property that is generalised in the next definition. The exponent 2 has no special significance in the context of Banach spaces, and therefore we replace it by an exponent $1 < p < \infty$.

**Definition 12.3.** *Let $1 < p < \infty$. A Banach space $E$ is said to be a UMD$_p$-space if there exists a constant $\beta$ such that for all $E$-valued $L^p$-martingale difference sequences $(d_n)_{n=1}^N$ we have*

$$\mathbb{E}\Big\|\sum_{n=1}^{N} \varepsilon_n d_n\Big\|^p \leqslant \beta^p \mathbb{E}\Big\|\sum_{n=1}^{N} d_n\Big\|^p.$$

If $(d_n)_{n=1}^N$ is an $E$-valued martingale difference sequence, then the same is true for $(\varepsilon_n d_n)_{n=1}^N$. This gives the reverse inequality

$$\mathbb{E}\Big\|\sum_{n=1}^{N} d_n\Big\|^p \leqslant \beta^p \mathbb{E}\Big\|\sum_{n=1}^{N} \varepsilon_n d_n\Big\|^p.$$

The term ‘UMD’ is an abbreviation for ‘unconditional martingale differences’. The least possible constant $\beta$ in the above inequalities is called the *UMD$_p$-constant* of $E$, notation $\beta_p(E)$.

Every Hilbert space $H$ is a UMD$_2$-space, with $\beta_2(H) = 1$; this is the content of (12.1). It is a trivial consequence of the definition that every closed subspace $F$ of a UMD$_p$-space is a UMD$_p$-space, with $\beta_p(F) \leqslant \beta_p(E)$.

As we shall see in the next section, if a Banach space is UMD$_p$ for some $1 < p < \infty$, then it is UMD$_p$ for all $1 < p < \infty$. In particular, Hilbert spaces are UMD$_p$ for all $1 < p < \infty$. Taking this for granted for the moment, the next result implies that for $1 < p < \infty$ the spaces $L^p(A)$, and more generally $L^p(A;H)$ for Hilbert spaces $H$, are UMD$_p$-spaces.

**Theorem 12.4.** *Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space and let $1 < p < \infty$. If $E$ is a UMD$_p$-space, then $L^p(A;E)$ is a UMD$_p$-space, with $\beta_p(L^p(A;E)) = \beta_p(E)$.*

*Proof.* Let $(d_n)_{n=1}^N$ be an $L^p$-martingale difference sequence with values in $L^p(A;E)$. With Fubini's theorem, for all choices of signs $\varepsilon_n = \pm 1$ we obtain

$$
\begin{aligned}
\mathbb{E}\Big\|\sum_{n=1}^N \varepsilon_n d_n\Big\|_{L^p(A)}^p &= \int_A \mathbb{E}\Big\|\sum_{n=1}^N \varepsilon_n d_n\Big\|^p \,d\mu \\
&\leqslant \beta_p(E)^p \int_A \mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|^p \,d\mu = \beta_p(E)^p \mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|_{L^p(A)}^p.
\end{aligned}
$$

In this computation we used that for $\mu$-almost all $\xi \in A$ the sequences $(d_n(\xi))_{n=1}^N$ is an $E$-valued martingale difference sequence; this follows from the observation that under the identification $L^p(\Omega; L^p(A;E)) \simeq L^p(A; L^p(\Omega;E))$ we have $\mathbb{E}_{L^p(A;E)}(\cdot|\mathscr{F}_n) = I \otimes \mathbb{E}_{L^p(A)}(\cdot|\mathscr{F}_n)$. This proves that $L^p(A;E)$ is a UMD$_p$-space, with $\beta_p(L^p(A;E)) \leqslant \beta_p(E)$.

If $f \in L^p(A)$ has norm 1, then $x \mapsto f \otimes x$ defines an isometric embedding of $E$ into $L^p(A;E)$; this gives the opposite inequality $\beta_p(E) \leqslant \beta_p(L^p(A;E))$. $\square$

Duality provides another way to produce new UMD-spaces from old:

**Proposition 12.5.** *Let $1 < p, q < \infty$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Then $E$ is a UMD$_p$-space if and only if $E^*$ is a UMD$_q$-space, and in this situation we have $\beta_p(E) = \beta_q(E^*)$.*

*Proof.* Suppose $E$ is a UMD$_p$-space and let $(d_n^*)_{n=1}^N$ be an $E^*$-valued $L^q$-martingale difference sequence with respect to $(\mathscr{F}_n)_{n=1}^N$. Fix an arbitrary $Y \in L^p(\Omega, \mathscr{F}_N; E)$ of norm 1, and define the $E$-valued $L^p$-martingale $(M_n)_{n=1}^N$ by $M_n := \mathbb{E}(Y|\mathscr{F}_n)$. Let $(d_n)_{n=1}^N$ be its difference sequence. Then $Y = \sum_{m=1}^N d_m$. If $1 \leqslant m < n \leqslant N$, then

$$
\mathbb{E}\langle d_m, d_n^* \rangle = \mathbb{E}\mathbb{E}(\langle d_m, d_n^* \rangle | \mathscr{F}_{n-1}) = \mathbb{E}\langle d_m, \mathbb{E}(d_n^*|\mathscr{F}_{n-1}) \rangle = 0.
$$

The second identity is justified as in the proof of Proposition 12.2. A similar computation shows that $\mathbb{E}\langle d_m, d_n^* \rangle = 0$ if $1 \leqslant n < m \leqslant N$. Hence,

$$
\begin{aligned}
\Big|\mathbb{E}\Big\langle Y, \sum_{n=1}^N \varepsilon_n d_n^*\Big\rangle\Big| &= \Big|\mathbb{E}\Big\langle \sum_{m=1}^N d_m, \sum_{n=1}^N \varepsilon_n d_n^*\Big\rangle\Big| \\
&= \Big|\mathbb{E}\Big\langle \sum_{m=1}^N \varepsilon_m d_m, \sum_{n=1}^N d_n^*\Big\rangle\Big| \\
&\leqslant \beta_p(E)\Big(\mathbb{E}\Big\| \sum_{m=1}^N d_m\Big\|^p\Big)^{\frac{1}{p}} \Big(\mathbb{E}\Big\| \sum_{n=1}^N d_n^*\Big\|^q\Big)^{\frac{1}{q}} \\
&= \beta_p(E)\Big(\mathbb{E}\Big\| \sum_{n=1}^N d_n^*\Big\|^q\Big)^{\frac{1}{q}}.
\end{aligned}
$$

In the last identity we used the assumption that $\|Y\|_p = 1$. Since $L^p(\Omega, \mathscr{F}_N; E)$ is norming for $L^q(\Omega, \mathscr{F}_N; E^*)$ (see Exercise 1.5), by taking the supremum over all $Y$ of norm 1 we obtain the estimate

$$
\Big(\mathbb{E}\Big\| \sum_{n=1}^N \varepsilon_n d_n^*\Big\|^q\Big)^{\frac{1}{q}} \leqslant \beta_p(E)\Big(\mathbb{E}\Big\| \sum_{n=1}^N d_n^*\Big\|^q\Big)^{\frac{1}{q}}.
$$

This proves that $E^*$ is a $\mathrm{UMD}_q$-space, with $\beta_q(E^*) \leqslant \beta_p(E)$.

If $E^*$ is a $\mathrm{UMD}_q$-space, the result just proved implies that $E^{**}$ is a $\mathrm{UMD}_p$-space, with $\beta_p(E^{**}) \leqslant \beta_q(E^*)$. Hence also $E$, being isometrically contained as a closed subspace in $E^{**}$ (by the Hahn-Banach theorem each $x \in E$ defines a functional $\phi_x$ in $E^{**}$ of norm $\|\phi_x\| = \|x\|$ by the formula $\langle x^*, \phi_x\rangle := \langle x, x^*\rangle$), is a $\mathrm{UMD}_p$-space, with $\beta_p(E) \leqslant \beta_p(E^{**}) \leqslant \beta_q(E^*)$.

Combining both parts, we obtain the equality $\beta_p(E) = \beta_q(E^*)$. □

*Remark 12.6.* It can be shown that every $\mathrm{UMD}_p$-space $E$ is *reflexive*, that is, the canonical mapping $x \mapsto \phi_x$ from $E$ to $E^{**}$ is surjective. This fact will not be needed in what follows.

## 12.2 $p$-Independence of the $\mathrm{UMD}_p$-property

This section is devoted to the proof of the highly non-trivial fact, already mentioned above, that the $\mathrm{UMD}_p$-property is independent of the parameter $1 < p < \infty$. The work consists of two parts: a reduction of the problem to difference sequences of so-called Haar martingales, and then proving the $p$-independence for this class of martingales.

### 12.2.1 Reduction to Haar martingales

A probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is said to be *divisible* if for all $F \in \mathscr{F}$ and $0 < r < 1$ we have $F = F_1 \cup F_2$ with $F_1, F_2 \in \mathscr{F}$ and

$$
\mathbb{P}(F_1) = r\mathbb{P}(F), \quad \mathbb{P}(F_2) = (1-r)\mathbb{P}(F).
$$

For $1 < p < \infty$, let us say that $E$ has the *UMD$_p^{\mathrm{div}}$-property* if there exists a constant $\beta_p^{\mathrm{div}}(E)$ such that

$$\mathbb{E}\Big\| \sum_{n=1}^{N} \varepsilon_n d_n \Big\|^p \leqslant (\beta_p^{\mathrm{div}}(E))^p \mathbb{E}\Big\| \sum_{n=1}^{N} d_n \Big\|^p$$

for all $E$-valued $L^p$-martingale difference sequences $(d_n)_{n=1}^N$ defined on a divisible probability space. Trivially, if $E$ has the UMD$_p$-property, then it has the UMD$_p^{\mathrm{div}}$-property and $\beta_p^{\mathrm{div}}(E) \leqslant \beta_p(E)$. The next lemma establishes the converse.

**Lemma 12.7.** *Let $1 < p < \infty$. If $E$ has the UMD$_p^{\mathrm{div}}$-property, then it has the UMD$_p$-property and $\beta_p(E) = \beta_p^{\mathrm{div}}(E)$.*

*Proof.* Suppose that $(d_n)_{n=1}^N$ is an $E$-valued $L^p$-martingale difference sequence with respect to a filtration $(\mathscr{F}_n)_{n=1}^N$ on an arbitrary probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The idea is to enlarge the probability space in such a way that it becomes divisible, without affecting the $L^p$-estimates for the martingale differences.

Consider $\widetilde{\Omega} := \Omega \times [0,1]$, $\widetilde{\mathscr{F}} := \mathscr{F} \times \mathscr{B}([0,1])$, and $\widetilde{P} := \mathbb{P} \times m$, where $m$ is the Lebesgue measure on the Borel $\sigma$-algebra $\mathscr{B}([0,1])$. The probability space $(\widetilde{\Omega}, \widetilde{\mathscr{F}}, \widetilde{\mathbb{P}})$ is divisible: this follows from the intermediate value theorem applied to the continuous function $t \mapsto \widetilde{\mathbb{P}}(\widetilde{F} \cap (\Omega \times [0,t]))$, where $\widetilde{F} \in \widetilde{\mathscr{F}}$.

Let $(M_n)_{n=1}^N$ be the martingale associated with $(d_n)_{n=1}^N$. Define $\widetilde{M}_n(\omega, t) := M_n(\omega)$ and $\widetilde{\mathscr{F}}_n := \mathscr{F}_n \times \mathscr{B}([0,1])$. It is easily checked that $(\widetilde{M}_n)_{n=1}^N$ is a martingale with respect to $(\widetilde{\mathscr{F}}_n)_{n=1}^N$ and, for every sequence of signs $(\varepsilon_n)_{n=1}^N$, its difference sequence $(\widetilde{d}_n)_{n=1}^N$ satisfies

$$\begin{aligned}\mathbb{E}\Big\| \sum_{n=1}^{N} \varepsilon_n d_n \Big\|^p &= \widetilde{\mathbb{E}}\Big\| \sum_{n=1}^{N} \varepsilon_n \widetilde{d}_n \Big\|^p \\ &\leqslant (\beta_p^{\mathrm{div}}(E))^p \widetilde{\mathbb{E}}\Big\| \sum_{n=1}^{N} \widetilde{d}_n \Big\|^p = (\beta_p^{\mathrm{div}}(E))^p \mathbb{E}\Big\| \sum_{n=1}^{N} d_n \Big\|^p\end{aligned}$$

using the UMD$_p^{\mathrm{div}}$-property of $E$. □

In the next step we restrict the class of probability spaces even further. If $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space, we call a sub-$\sigma$-algebra $\mathscr{G}$ of $\mathscr{F}$ *dyadic* if it is generated by $2^m$ sets of measure $2^{-m}$ for some integer $m \geqslant 0$. We call a filtration in $(\Omega, \mathscr{F}, \mathbb{P})$ *dyadic* if each of its constituting $\sigma$-algebras is dyadic. For $1 < p < \infty$, let us say that $E$ has the *UMD$_p^{\mathrm{dyad}}$-property* if there exists a constant $\beta_p^{\mathrm{dyad}}(E)$ such that

$$\mathbb{E}\Big\| \sum_{n=1}^{N} \varepsilon_n d_n \Big\|^p \leqslant (\beta_p^{\mathrm{dyad}}(E))^p \mathbb{E}\Big\| \sum_{n=1}^{N} d_n \Big\|^p$$

holds for all $E$-valued $L^p$-martingale difference sequences $(d_n)_{n=1}^N$ with respect to a dyadic filtration $(\mathscr{F}_n)_{n=1}^N$ on a divisible probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Trivially, if $E$ has the UMD$_p$-property, then it has the UMD$_p^{\rm dyad}$-property and $\beta_p^{\rm dyad}(E) \leqslant \beta_p(E)$. In order to establish the converse we need a simple approximation result. The proof appears somewhat technical, but by drawing a picture one sees that it is nearly trivial.

**Lemma 12.8.** *Let $1 \leqslant p < \infty$ and $\varepsilon > 0$ be given. If $f$ is a simple random variable on a divisible probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and $\mathscr{G}$ is a dyadic sub-$\sigma$-algebra of $\mathscr{F}$, there exists a dyadic sub-$\sigma$-algebra $\mathscr{G} \subseteq \mathscr{H} \subseteq \mathscr{F}$ and an $\mathscr{H}$-measurable simple random variable $h$ such that $\|f - h\|_p < \varepsilon$.*

*Proof.* Suppose $\mathscr{G}$ is generated by $2^m$ sets of measure $2^{-m}$.

It suffices to prove the lemma for indicator functions $f = 1_F$ with $F \in \mathscr{F}$. Considering $F \in \mathscr{F}$ as fixed, we write $1_F = \sum_G 1_{F \cap G}$ where the sum extends over the $2^m$ generating sets $G$ of $\mathscr{G}$.

Take one such $G$ and let $(b_j^G)_{j=1}^\infty$ denote the digits in the binary expansion of the real number $\mathbb{P}(F \cap G)$. Informally, we use the digits to write $F \cap G$ inductively as a union, up to a null set, of disjoint 'dyadic' subsets of maximal measure.

To be more precise, inductively define sets $A_j^G$ and $B_j^G$ by $A_0^G = F \cap G$ and $B_0^G = \varnothing$, and requiring, for $j \geqslant 1$, that $B_j^G \subseteq A_{j-1}^G$ satisfies $B_j^G \in \mathscr{F}$ and $\mathbb{P}(B_j^G) = b_j^G 2^{-j}$ (we may take $B_j^G := \varnothing$ if $b_j^G = 0$). Then put $A_j^G := A_{j-1}^G \setminus B_j^G$ and continue.

The sets $B_j^G \in \mathscr{F}$ thus constructed are disjoint, contained in $G$, and satisfy $\mathbb{P}\big((F \cap G) \setminus \bigcup_{j=1}^\infty B_j^G\big) = 0$. Let $n \geqslant 1$ be the first integer such that

$$\mathbb{P}\Big((F \cap G) \setminus \bigcup_{j=1}^n B_j^G\Big) < \frac{\varepsilon^p}{2^m}.$$

For each $1 \leqslant j \leqslant n$ such that $b_j^G = 1$ we have $\mathbb{P}(B_j^G) = 2^{-j}$. If follows that we can split $G$ into disjoint subsets of measure $2^{-n}$ in such a way that each $B_j^G$, $1 \leqslant j \leqslant n$, is a finite union of these subsets.

We subdivide each of the $2^m$ generating sets $G$ in this way. The number $n$ varies over $G$, but by considering further subdivisions we may assume it to be independent of $G$. Let $\mathscr{H}$ be the $\sigma$-algebra generated by the $2^n$ sets of measure $2^{-n}$ thus obtained. This $\sigma$-algebra is dyadic, it contains $\mathscr{G}$, and the simple function

$$h := \sum_G \sum_{\substack{1 \leqslant j \leqslant n \\ b_j^G = 1}} 1_{B_j^G}$$

is $\mathscr{H}$-measurable and satisfies $\|f - h\|_p < \varepsilon$. $\square$

**Lemma 12.9.** *Let $1 < p < \infty$. If $E$ has the UMD$_p^{\rm dyad}$-property, then it has the UMD$_p$-property and $\beta_p(E) = \beta_p^{\rm dyad}(E)$.*

*Proof.* Suppose that $(d_n)_{n=1}^N$ is an $E$-valued $L^p$-martingale difference sequence with respect to a filtration $(\mathscr{F}_n)_{n=1}^N$ on a divisible probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The idea is approximate the $d_n$ with simple functions as in the previous lemma.

Fixing $\varepsilon > 0$, we can find $\mathscr{F}_n$-measurable simple functions $s_n : \Omega \to E$ such that $\|d_n - s_n\|_p < \frac{\varepsilon}{N}$. By repeated application of Lemma 12.8 we find a sequence of dyadic $\sigma$-algebras $(\widetilde{\mathscr{F}_n})_{n=1}^N$ such that $\widetilde{\mathscr{F}}_{n-1} \subseteq \widetilde{\mathscr{F}_n} \subseteq \mathscr{F}_n$ (with the understanding that $\widetilde{\mathscr{F}_0} = \{\varnothing, \Omega\}$) and a sequence of step functions $(\widetilde{s}_n)_{n=1}^N$ such that each $\widetilde{s}_n$ is $\widetilde{\mathscr{F}_n}$-measurable and satisfies $\|s_n - \widetilde{s}_n\|_p < \frac{\varepsilon}{N}$.

Consider the sequence $(\widetilde{d}_n)_{n=1}^N$ defined by $\widetilde{d}_n := \mathbb{E}(d_n|\widetilde{\mathscr{F}_n})$. To see that this is a martingale difference sequence with respect to the filtration $(\widetilde{\mathscr{F}_n})_{n=1}^N$, note that for $1 < n \leqslant N$,

$$\begin{aligned}\mathbb{E}(\widetilde{d}_n|\widetilde{\mathscr{F}}_{n-1}) &= \mathbb{E}(\mathbb{E}(d_n|\widetilde{\mathscr{F}_n})|\widetilde{\mathscr{F}}_{n-1}) \\ &= \mathbb{E}(d_n|\widetilde{\mathscr{F}}_{n-1}) = \mathbb{E}(\mathbb{E}(d_n|\mathscr{F}_{n-1})|\widetilde{\mathscr{F}}_{n-1}) = 0.\end{aligned}$$

Then, by the $L^p$-contractivity of conditional expectations,

$$\begin{aligned}\|d_n - \widetilde{d}_n\|_p &\leqslant \frac{2\varepsilon}{N} + \|\widetilde{s}_n - \widetilde{d}_n\|_p \\ &= \frac{2\varepsilon}{N} + \|\mathbb{E}(\widetilde{s}_n - \widetilde{d}_n|\widetilde{\mathscr{F}_n})\|_p \\ &= \frac{2\varepsilon}{N} + \|\mathbb{E}(\widetilde{s}_n - d_n|\widetilde{\mathscr{F}_n})\|_p \leqslant \frac{2\varepsilon}{N} + \|\widetilde{s}_n - d_n\|_p = \frac{4\varepsilon}{N}.\end{aligned}$$

Hence,

$$\begin{aligned}\Big\|\sum_{n=1}^N \varepsilon_n d_n\Big\|_p &\leqslant 4\varepsilon + \Big\|\sum_{n=1}^N \varepsilon_n \widetilde{d}_n\Big\|_p \\ &\leqslant 4\varepsilon + \beta_p^{\mathrm{dyad}}(E)\Big\|\sum_{n=1}^N \widetilde{d}_n\Big\|_p \leqslant 4\varepsilon(1 + \beta_p^{\mathrm{dyad}}(E))\Big\|\sum_{n=1}^N d_n\Big\|_p.\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, this shows that $E$ has the UMD$_p^{\mathrm{div}}$-property with $\beta_p^{\mathrm{div}}(E) \leqslant \beta_p^{\mathrm{dyad}}(E)$. Together with Lemma 12.7 this proves the result. $\square$

The final reduction consists of shrinking the class of difference sequences to *Haar martingale difference sequences*, which are defined as difference sequences of martingales with respect to a *Haar filtration*. This is a filtration $(\mathscr{F}_n)_{n=1}^N$, where $\mathscr{F}_1 = \{\varnothing, \Omega\}$ and each $\mathscr{F}_n$ (with $n \geqslant 1$) is obtained from $\mathscr{F}_{n-1}$ by dividing precisely one atom of $\mathscr{F}_{n-1}$ of maximal measure into two sets of equal measure (an *atom* of a $\sigma$-algebra $\mathscr{G}$ is a set $G \in \mathscr{G}$ such that $H \subseteq G$ with $H \in \mathscr{G}$ implies $H \in \{\varnothing, G\}$). By construction, each $\mathscr{F}_n$ is generated by $n$ atoms, whose measures equal $2^{-k-1}$ or $2^{-k}$, where $k$ is the unique integer such that $2^{k-1} < n \leqslant 2^k$.

For $1 < p < \infty$, let us say that $E$ has the *UMD$_p^{\mathrm{Haar}}$-property* if there exists a constant $\beta_p^{\mathrm{Haar}}(E)$ such that

$$\mathbb{E}\Big\| \sum_{n=1}^{N} \varepsilon_n d_n \Big\|^p \leqslant (\beta_p^{\mathrm{Haar}}(E))^p \mathbb{E}\Big\| \sum_{n=1}^{N} d_n \Big\|^p$$

holds for all $E$-valued Haar martingale difference sequences $(d_n)_{n=1}^N$ defined on a divisible probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Trivially, if $E$ has the UMD$_p$-property, then it has the UMD$_p^{\mathrm{Haar}}$-property and $\beta_p^{\mathrm{Haar}}(E) \leqslant \beta_p(E)$.

**Lemma 12.10.** *Let $1 < p < \infty$. If $E$ has the UMD$_p^{\mathrm{Haar}}$-property, then it has the UMD$_p$-property and $\beta_p^{\mathrm{Haar}}(E) = \beta_p(E)$.*

*Proof.* Suppose that $(d_n)_{n=1}^N$ is an $E$-valued $L^p$-martingale difference sequence with respect to a dyadic filtration $(\mathscr{F}_n)_{n=1}^N$ on a divisible probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The idea is to 'embed' $(d_n)_{n=1}^N$ into a Haar martingale difference sequence. To be more precise, we shall construct an $L^p$-martingale difference sequence $(\widetilde{d}_k)_{k=1}^K$ with respect to a Haar filtration $(\widetilde{\mathscr{F}}_k)_{k=1}^K$ such that $M_n = \widetilde{M}_{k_n}$ and $\mathscr{F}_n = \widetilde{\mathscr{F}}_{k_n}$ for some subsequence $k_1 < \cdots < k_N$. Once this has been done, we note that $d_n = \sum_{j=k_{n-1}+1}^{k_n} \widetilde{d}_j$ and

$$\begin{aligned} \Big\| \sum_{n=1}^{N} \varepsilon_n d_n \Big\|_p &= \Big\| \sum_{n=1}^{N} \sum_{j=k_{n-1}+1}^{k_n} \varepsilon_n \widetilde{d}_j \Big\|_p = \Big\| \sum_{k=1}^{K} \widetilde{\varepsilon}_k \widetilde{d}_k \Big\|_p \\ &\leqslant \beta_p^{\mathrm{Haar}}(E) \Big\| \sum_{k=1}^{K} \widetilde{d}_k \Big\|_p = \beta_p^{\mathrm{Haar}}(E) \Big\| \sum_{n=1}^{N} d_n \Big\|_p, \end{aligned}$$

where $\widetilde{\varepsilon}_k = \varepsilon_{k_n}$ for $k = k_{n-1}+1, \dots, k_n$.

Each $\mathscr{F}_n$ is dyadic and therefore it is generated by $k_n := 2^{l_n}$ atoms of measure $2^{-l_n}$. Since each atom of $\mathscr{F}_{n-1}$ is a finite union of atoms in $\mathscr{F}_n$ we have $k_1 < \cdots < k_N$. The $\sigma$-algebras $\widetilde{\mathscr{F}}_k$, with $k_{n-1} < k < k_n$ can now be constructed by splitting the atoms of $\widetilde{\mathscr{F}}_{k_{n-1}}$ one by one into two disjoint subsets of equal measure, so as to arrive at the atoms of $\widetilde{\mathscr{F}}_{k_n}$ by repeating this procedure $k_n - k_{n-1}$ times.

Now take $\widetilde{M}_{k_n} := M_n$ and $\widetilde{M}_k := E(\widetilde{M}_{k_n} | \widetilde{\mathscr{F}}_k)$ if $k_{n-1} < k \leqslant k_n$. ☐

### 12.2.2 *p*-Independence for Haar martingales

By the reductions of the previous subsection, in order to prove the $p$-independence of the UMD$_p$-property it suffices to consider Haar martingale difference sequences. Such sequences have a special property which is captured in the next lemma.

**Lemma 12.11.** *If $(d_n)_{n=1}^N$ is an $E$-valued Haar martingale difference sequence, then $\|d_{n+1}\|$ is $\mathscr{F}_n$-measurable for all $n = 1, \dots, N-1$.*

*Proof.* Suppose that $\mathscr{F}_{n+1}$ is obtained by splitting one of the $n+1$ generating atoms of $\mathscr{F}_n$, say $A$, into subsets $A_1$ and $A_2$ of equal measure. Then $M_{n+1}$ and $M_n$ only differ on $A$, so $d_{n+1} = 0$ on $\complement A$. Also, $d_{n+1}$ is constant on $A_1$ and $A_2$, say with values $x_1$ and $x_2$. Then,

$$\mathbb{P}(A_1)x_1 + \mathbb{P}(A_2)x_2 = \int_A d_{n+1}\,d\mathbb{P} = \int_A \mathbb{E}(d_{n+1}|\mathscr{F}_n)\,d\mathbb{P} = 0,$$

and from $\mathbb{P}(A_1) = \mathbb{P}(A_2)$ we deduce that $x_1 + x_2 = 0$. Hence, $\|d_{n+1}\| = 1_{A_1}\|x_1\| + 1_{A_2}\|x_2\| = 1_A\|x_1\|$ is $\mathscr{F}_n$-measurable. □

In what follows we let $f = (f_n)_{n=1}^N$ be an $E$-valued Haar martingale with difference sequence $(d_n)_{n=1}^N$. By the lemma, the non-negative random variables $\|d_{n+1}\|$ are $\mathscr{F}_n$-measurable for $n = 1, \dots, N-1$.

For a fixed sequence of signs $\varepsilon = (\varepsilon_n)_{n=1}^N$ we denote by $g = (g_n)_{n=1}^N$ the martingale transform $g_n = \sum_{j=1}^n \varepsilon_j d_j$. Further we let

$$f^*(\omega) := \max_{1\leqslant n\leqslant N} \|f_n(\omega)\|, \quad g^*(\omega) := \max_{1\leqslant n\leqslant N} \|g_n(\omega)\|.$$

In the proof of the next lemma we use the following notation: if $(X_n)_{n=1}^N$ is a sequence of $E$-valued random variables and $\tau : \Omega \to \{1, \dots, N\}$ is another random variable, we define the random variable $X_\tau : \Omega \to E$ by

$$X_\tau(\omega) := X_{\tau(\omega)}(\omega).$$

**Lemma 12.12.** *Suppose that $E$ is a UMD$_q$-space for some $1 < q < \infty$. For all $\delta > 0$ and $\beta > 2\delta + 1$ and all $\lambda > 0$ we have*

$$\mathbb{P}\{g^* > \beta\lambda,\ f^* \leqslant \delta\lambda\} \leqslant \alpha^q \mathbb{P}\{g^* > \lambda\},$$

*where $\alpha = 4\delta\beta_q(E)/(\beta - 2\delta - 1)$.*

*Proof.* Since $\mathscr{F}_1 = \{\varnothing, \Omega\}$, the random variable $f_1 = d_1$ is constant almost surely. If the constant value is greater than $\delta\lambda$, then the left hand side in the above inequality vanishes and there is nothing to prove. We may therefore assume that $f_1 \leqslant \delta\lambda$ almost surely.

Let

$$\begin{aligned}
\mu(\omega) &:= \min\{1 \leqslant n \leqslant N :\ \|g_n(\omega)\| > \lambda\},\\
\nu(\omega) &:= \min\{1 \leqslant n \leqslant N :\ \|g_n(\omega)\| > \beta\lambda\},\\
\sigma(\omega) &:= \min\{1 \leqslant n \leqslant N :\ \|f_n(\omega)\| > \delta\lambda \ \text{ or } \ \|d_{n+1}\| > 2\delta\lambda\}
\end{aligned}$$

with the convention that $\min\varnothing := N + 1$. In the third definition we further use the convention that $d_{N+1} := 0$.

Let $v_n$ be the indicator function of the set $\{\mu < n \leqslant \min\{\nu, \sigma\}\}$. Since $d = (d_n)_{n=1}^N$ is a Haar martingale difference sequence, the sequence $v = (v_n)_{n=1}^N$ is predictable by Lemma 12.11 and therefore

$$F_n := \sum_{j=1}^{n} v_j d_j$$

defines a martingale $F = (F_n)_{n=1}^N$ by the result of Example 11.19. On the set $\{\sigma \leqslant \mu\}$ we have $v_j \equiv 0$ for all $j$ and therefore $F_N \equiv 0$ there. In particular this is the case on the set $\{\mu = N+1\} = \{g^* \leqslant \lambda\}$. On the set $\{\sigma > \mu\}$ we have

$$\|F_N\| = \Big\| \sum_{\mu < j \leqslant \min\{\nu,\sigma\}} d_j \Big\| = \|f_{\min\{\nu,\sigma\}} - f_\mu\| \leqslant 4\delta\lambda.$$

To see this, first note that $\mu(\omega) < \sigma(\omega)$ implies $\|f_\mu(\omega)\| \leqslant \delta\lambda$. Also, if $\min\{\nu(\omega), \sigma(\omega)\} = 1$, then by the assumption above $\|f_{\min\{\nu,\sigma\}}(\omega)\| = \|f_1(\omega)\| \leqslant \delta\lambda$; if $\min\{\nu(\omega), \sigma(\omega)\} > 1$, then from $\|f_{\min\{\nu,\sigma\}-1}(\omega)\| \leqslant \delta\lambda$ and $\|d_{\min\{\nu,\sigma\}}(\omega)\| \leqslant 2\delta\lambda$ it follows that $\|f_{\min\{\nu,\sigma\}}(\omega)\| \leqslant \|f_{\min\{\nu,\sigma\}-1}(\omega)\| + \|d_{\min\{\nu,\sigma\}}(\omega)\| \leqslant 3\delta\lambda$. This proves the claim.

We infer that

$$\mathbb{E}\|F_n\|^q \leqslant (4\delta\lambda)^q \mathbb{P}\{g^* > \lambda\}.$$

Now consider the martingale transform $G$ of $F$ by $\varepsilon$,

$$G_n := \sum_{j=1}^{n} \varepsilon_j v_j d_j.$$

On the set $\{\nu \leqslant N,\ \sigma = N+1\}$ we have $\min\{\nu, \sigma\} = \nu$ and

$$\|G_N\| = \Big\| \sum_{\mu < j \leqslant \nu} \varepsilon_j d_j \Big\| = \|g_\nu - g_\mu\| > \beta\lambda - 2\delta\lambda - \lambda,$$

where the last inequality uses that on the set $\{\nu \leqslant N,\ \sigma = N+1\}$ we have $\|g_\nu(\omega)\| > \beta\lambda$ and $\|g_\mu(\omega)\| \leqslant \|g_{\mu-1}(\omega)\| + \|d_\mu(\omega)\| \leqslant \lambda + 2\delta\lambda$.

By Chebyshev's inequality and the UMD$_q$-property,

$$\begin{aligned}
\mathbb{P}\{g^* > \beta\lambda,\ f^* \leqslant \delta\lambda\} &\leqslant \mathbb{P}\{\nu \leqslant N,\ \sigma = N+1\} \\
&\leqslant \mathbb{P}\{\|G_N\| > \beta\lambda - 2\delta\lambda - \lambda\} \\
&\leqslant \frac{1}{(\beta\lambda - 2\delta\lambda - \lambda)^q} \mathbb{E}\|G_N\|^q \\
&\leqslant \frac{(\beta_q(E))^q}{(\beta\lambda - 2\delta\lambda - \lambda)^q} \mathbb{E}\|F_N\|^q \\
&\leqslant \frac{(4\delta)^q (\beta_q(E))^q}{(\beta - 2\delta - 1)^q} \mathbb{P}\{g^* > \lambda\}.
\end{aligned}$$

In the first inequality we used that $f^*(\omega) \leqslant \delta\lambda$ implies that $\|d_j(\omega)\| \leqslant 2\delta\lambda$ for all $j$. This proves the lemma. □

**Theorem 12.13.** *If $E$ is a UMD$_q$-space for some $1 < q < \infty$, then it is a UMD$_p$-space for all $1 < p < \infty$.*

*Proof.* By the results of the previous subsection it suffices to show that $E$ has the $\mathrm{UMD}_p^{\mathrm{Haar}}$-property for all $1 < p < \infty$. Thus we find ourselves in the situation of the previous lemma and need to prove the estimate

$$\mathbb{E}\|g_N\|^p \leqslant b^p \mathbb{E}\|f_N\|^p$$

with a constant $b \geqslant 0$ depending only on $p$, $q$, and $E$, but not on $f$, $g$ and $N$.

Fix an arbitrary number $\beta > 1$. For $\delta > 0$ so small that $\beta > 2\delta + 1$, let $\alpha = \alpha_{\beta,\delta,q,E}$ be as in the lemma. Then, by an integration by parts and Doob's maximal inequality,

$$\begin{aligned}
\mathbb{E}\|g_N\|^p \leqslant \mathbb{E}\|g^*\|^p &= \beta^p \int_0^\infty p\lambda^{p-1}\mathbb{P}\{g^* > \beta\lambda\}\,d\lambda \\
&\leqslant \alpha^q\beta^p \int_0^\infty p\lambda^{p-1}\mathbb{P}\{g^* > \lambda\}\,d\lambda \\
&\qquad + \beta^p \int_0^\infty p\lambda^{p-1}\mathbb{P}\{f^* > \delta\lambda\}\,d\lambda \\
&\leqslant \alpha^q\beta^p\mathbb{E}\|g^*\|^p + \frac{\beta^p}{\delta^p}\mathbb{E}\|f^*\|^p \\
&\leqslant C_p^p\alpha^q\beta^p\mathbb{E}\|g_N\|^p + \frac{C_p^p\beta^p}{\delta^p}\mathbb{E}\|f_N\|^p,
\end{aligned}$$

where $C_p = p/(p-1)$. Since $\lim_{\delta\downarrow 0}\alpha_{\beta,\delta,q,E} = 0$, by taking $\delta > 0$ small enough we may arrange that $C_p^p\alpha^q\beta^p < 1$. Noting that $\mathbb{E}\|g_N\|^p < \infty$ since $g_N$ is simple (recall that $\mathscr{F}_N$ is a finite $\sigma$-algebra) it follows that

$$\mathbb{E}\|g_N\|^p \leqslant \frac{C_p^p\beta^p}{(1 - C_p^p\alpha^q\beta^p)\delta^p}\mathbb{E}\|f_N\|^p.$$

This concludes the proof. □

This theorem justifies the following definition.

**Definition 12.14.** *A Banach space is called a* UMD-space *if it is a $\mathrm{UMD}_p$-space for some (and hence, for all) $1 < p < \infty$.*

By combining Theorem 12.13 with the results of the previous section we see that all Hilbert spaces and all spaces $L^p(A)$ with $1 < p < \infty$ are UMD-spaces.

## 12.3 The vector-valued Stein inequality

In this final section we prove an extension, due to Bourgain, of a beautiful result of Stein which asserts that conditional expectation operators corresponding to the $\sigma$-algebras of a filtration form an $R$-bounded family.

**Theorem 12.15 (Vector-valued Stein inequality).** *Let $E$ be a UMD-space and fix $1 < p < \infty$. If $(\mathscr{F}_t)_{t\in[0,T]}$ is a filtration on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, then the family of conditional expectation operators $\{\mathbb{E}(\cdot|\mathscr{F}_t) : t \in [0,T]\}$ is R-bounded (and hence $\gamma$-bounded) on $L^p(\Omega; E)$.*

*Proof.* Let $(\widetilde{r}_n)_{n=1}^N$ be a Rademacher sequence on a second probability space $(\widetilde{\Omega}, \widetilde{\mathscr{F}}, \widetilde{\mathbb{P}})$ and define $\widetilde{\mathscr{F}}_n = \sigma(\widetilde{r}_1, \dots, \widetilde{r}_n)$, $n = 1, \dots, N$. Fix $t_1 < \dots < t_N$ in $[0,T]$. On the product space $\Omega \times \widetilde{\Omega}$ define the filtration $(\mathscr{G}_m)_{m=1}^{2N}$ by

$$\begin{aligned} \mathscr{G}_{2n-1} &:= \mathscr{F}_{t_n} \times \widetilde{\mathscr{F}}_{n-1}, \quad n = 1, \dots, N, \\ \mathscr{G}_{2n} &:= \mathscr{F}_{t_n} \times \widetilde{\mathscr{F}}_n, \qquad n = 1, \dots, N. \end{aligned}$$

For a random variable $X \in L^p(\Omega \times \widetilde{\Omega}; E)$ define the martingale $(M_m)_{m=1}^{2M}$ by

$$M_m := \mathbb{E}(X|\mathscr{G}_m), \quad m = 1, \dots, 2N.$$

Let $(d_m)_{m=1}^{2M}$ be the associated martingale difference sequence. Then by the UMD$_p$-property of $E$,

$$\Big\| \sum_{n=1}^N d_{2n} \Big\|_{L^p(\Omega;E)} \leqslant \beta_p(E) \Big\| \sum_{m=1}^{2N} d_m \Big\|_{L^p(\Omega;E)}. \tag{12.2}$$

Indeed, the sum on the left hand side equals $\frac{1}{2}\big(\sum_{m=1}^{2N} d_m + \sum_{m=1}^{2N} (-1)^m d_m\big)$.

Now fix $f_1, \dots, f_N \in L^p(\Omega; E)$ and put $X := \sum_{n=1}^N \widetilde{r}_n f_n$. For this choice of $X$ we have

$$\begin{aligned} M_{2n-1} &= \sum_{j=1}^N \mathbb{E}(\widetilde{r}_j f_j | \mathscr{F}_{t_n} \times \widetilde{\mathscr{F}}_{n-1}) = \sum_{j=1}^{n-1} \widetilde{r}_j \mathbb{E}(f_j|\mathscr{F}_{t_n}), \\ M_{2n} &= \sum_{j=1}^N \mathbb{E}(\widetilde{r}_j f_j | \mathscr{F}_{t_n} \times \widetilde{\mathscr{F}}_n) = \sum_{j=1}^{n} \widetilde{r}_j \mathbb{E}(f_j|\mathscr{F}_{t_n}). \end{aligned}$$

Therefore $d_{2n-1} = 0$ and $d_{2n} = \widetilde{r}_n \mathbb{E}(f_j|\mathscr{F}_{t_n})$. It then follows from (12.2) that

$$\begin{aligned} \widetilde{\mathbb{E}} \Big\| \sum_{n=1}^N \widetilde{r}_n \mathbb{E}(f_n|\mathscr{F}_{t_n}) \Big\|_{L^p(\Omega;E)}^p &= \widetilde{\mathbb{E}} \Big\| \sum_{n=1}^N d_{2n} \Big\|_{L^p(\Omega;E)}^p \\ &\leqslant (\beta_p(E))^p \widetilde{\mathbb{E}} \Big\| \sum_{m=1}^{2N} d_m \Big\|_{L^p(\Omega;E)}^p \\ &= (\beta_p(E))^p \widetilde{\mathbb{E}} \Big\| \sum_{n=1}^N \widetilde{r}_n f_n \Big\|_{L^p(\Omega;E)}^p. \qquad \square \end{aligned}$$

## 12.4 Exercises

1. Prove that a Banach space $E$ is a UMD$_p$-space $E$ if and only if for some (and hence, for all) $1 < p < \infty$ there exist constants $\beta_p^{\pm}(E)$ such that for all $E$-valued $L^p$-martingale difference sequences $(d_n)_{n=1}^N$ and all Rademacher sequences $(\widetilde{r}_n)_{n=1}^N$ independent of $(d_n)_{n=1}^N$ we have
$$\frac{1}{(\beta_p^-(E))^p}\mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|^p \leqslant \mathbb{E}\Big\|\sum_{n=1}^N \widetilde{r}_n d_n\Big\|^p \leqslant (\beta_p^+(E))^p\mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|^p.$$

2. Let $1 < p < \infty$. Prove that if $H$ is a Hilbert space and $(d_n)_{n=1}^N$ is an $H$-valued $L^p$-martingale difference sequence, then
$$\frac{1}{c_p^p}\mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|^p \leqslant \mathbb{E}\Big(\sum_{n=1}^N \|d_n\|^2\Big)^{\frac{p}{2}} \leqslant C_p^p\mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|^p,$$
with constant depending only on $p$.
*Hint:* Combine Exercise 1 with the Kahane-Khintchine inequalities.

3. Prove that if $X$ is a UMD-space and $Y$ is a closed subspace, then $X/Y$ is a UMD-space and give an estimate for its UMD constant.

4. A sequence $(x_n)_{n=1}^\infty$ in a Banach space $E$ is called a *Schauder basis* if every $x \in E$ admits a unique representation $x = \sum_{n=1}^\infty a_n x_n$ with convergence in $E$. Using a closed graph argument one can show that the projections
$$D_N \sum_{n=1}^\infty a_n x_n := \sum_{n=1}^N a_n x_n,$$
are bounded. In fact, by the uniform boundedness theorem we even have $\sup_{N\geqslant 1} \|D_N\| < \infty$.
A Schauder basis is called *unconditional* if there exists a constant $0 < C < \infty$ such that for all $N \geqslant 1$, all scalars $a_1, \dots, a_N$, and all signs $\varepsilon_1, \dots, \varepsilon_N \in \{-1, +1\}$ we have
$$\frac{1}{C}\Big\|\sum_{n=1}^N a_n x_n\Big\| \leqslant \Big\|\sum_{n=1}^N \varepsilon_n a_n x_n\Big\| \leqslant C\Big\|\sum_{n=1}^N a_n x_n\Big\|.$$
The least admissible constant $C$ is called the *unconditionality constant* of $(x_n)_{n=1}^\infty$.
Let $(x_n)_{n=1}^\infty$ be an unconditional Schauder basis of $E$ with unconditionality constant $C$.
   a) Show that if $(r_n)_{n=1}^\infty$ is a Rademacher sequence, then for all $N \geqslant 1$ and all scalars $a_1, \dots, a_N$ we have
$$\frac{1}{C^2}\Big\|\sum_{n=1}^N a_n x_n\Big\|^2 \leqslant \mathbb{E}\Big\|\sum_{n=1}^N r_n a_n x_n\Big\|^2 \leqslant C^2\Big\|\sum_{n=1}^N a_n x_n\Big\|^2.$$

b) Show that $\sup_{N\geqslant 1}\|D_N\| \leqslant C$.

Assume next that $E$ is a UMD-space.

c) Show that the sequence $(D_N)_{N=1}^{\infty}$ is $R$-bounded.
*Hint:* Use a) and the vector-valued Stein inequality.

5. In this exercise we prove a vector-valued version of a multiplier theorem due to Marcinkiewicz. Let $(x_n)_{n=1}^{\infty}$ be a Schauder basis of the UMD Banach space $E$ which has an *unconditional blocking*, meaning that there is a sequence $0 = N_0 < N_1 < \dots$ and a constant $0 < C < \infty$ such that the corresponding block projections $\Delta_j := D_{N_j} - D_{N_{j-1}}$ (where $D_0 = 0$) satisfy
$$\frac{1}{C}\Big\|\sum_{j=1}^{k}\Delta_j x\Big\| \leqslant \Big\|\sum_{j=1}^{k}\varepsilon_j\Delta_j x\Big\| \leqslant C\Big\|\sum_{j=1}^{k}\Delta_j x\Big\|$$
for all choices $\varepsilon_n \in \{-1,1\}$. Suppose that $(\lambda_n)_{n=1}^{N}$ is a scalar sequence such that:

(i) $\sup\limits_{n\geqslant 1}|\lambda_n| < \infty$;

(ii) $\sup\limits_{j\geqslant 1}\sum\limits_{n=N_{j-1}+1}^{N_j-1}|\lambda_{n+1}-\lambda_n| < \infty$.

where $\lambda_0 = 0$. Prove that the multiplier
$$M\sum_{n=1}^{\infty}a_n x_n := \sum_{n=1}^{\infty}\lambda_n a_n x_n$$
is bounded.
*Hint:* Write
$$Mx = \sum_{j=1}^{\infty}\lambda_{N_j}\Delta_j x + \sum_{j=1}^{\infty}\sum_{n=N_{j-1}+1}^{N_j-1}(\lambda_n - \lambda_{n+1})D_n\Delta_j x.$$
Now use a randomisation argument, the result of the previous exercise, and Proposition 9.6.

*Remark.* It can be shown that the trigonometric system $(e_n)_{n\in\mathbb{Z}}$, where $e_n(\theta) = e^{in\theta}$, is a Schauder basis in $L^p(\mathbb{T})$ for all $1 < p < \infty$, but this basis is unconditional only for $p = 2$. However, it is a classical result of Littlewood and Paley that the dyadic blocking of $(e_n)_{n\in\mathbb{Z}}$ is unconditional in $L^p(\mathbb{T})$ for all $1 < p < \infty$ (in this blocking, the $j$-th block runs over the indices $2^{j-1} \leqslant |n| < 2^j$). In combination with the exercise, this gives the classical formulation of the Marcinkiewicz multiplier theorem.

**Notes.** The importance of UMD-spaces extends far beyond the domain of stochastic analysis. In fact, the subject was created in an effort to extend classical Fourier multiplier theorems to Banach-space valued functions. On the unit circle $\mathbb{T}$, an important Fourier multiplier is the Riesz projection

$$\sum_{n\in\mathbb{Z}} c_n e^{in\theta} \mapsto \sum_{n=0}^{\infty} c_n e^{in\theta}.$$

This projection, which corresponds to the multiplier $1_{\{n\geqslant 0\}}$, is bounded in $L^p(\mathbb{T})$ for all $1 < p < \infty$. On the real line, the Hilbert transform defined by the principle value integral

$$Hf(x) := \frac{1}{\pi}\mathrm{PV}\int_{-\infty}^{\infty} \frac{f(y)}{x-y}\,dy$$

is bounded on $L^p(\mathbb{R})$ for all $1 < p < \infty$; it can be shown that this operator corresponds to the multiplier $\frac{1}{i}(1_{\mathbb{R}_+} - 1_{\mathbb{R}_-})$. Both results are classical theorems of M. Riesz. In the Banach space-valued situation the validity of these results characterise the UMD-property:

**Theorem 12.16.** *Let $1 < p < \infty$. For a Banach space $E$ the following assertions are equivalent:*

(1) *$E$ is a UMD$_p$-space;*
(2) *The Riesz projection is bounded on $L^p(\mathbb{T}; E)$;*
(3) *The Hilbert transform is bounded on $L^p(\mathbb{R}; E)$.*

The implications (1)⇒(2) and (1)⇒(3) are due to Burkholder [18] and McConnell [74], and their converses to Bourgain [10]. We refer to the review papers [20, 97] for more details. Recently, far-reaching generalisations of Theorem 12.16 to the boundedness of Fourier multipliers and singular integral operators in vector-valued $L^p$-spaces have been proved by several authors. We refer to the excellent lecture notes by Kunstmann and Weis [61] for an overview and references to the literature.

The independence of the UMD$_p$-property of the parameter $1 < p < \infty$ (Theorem 12.13) was first proved by Maurey [73], who gives credit to Pisier. The proof via Lemma 12.12 presented here is adapted from Burkholder [19]. The reductions of Section 12.2.1 are a variation of those proposed in [73] and carried out in detail in the lecture notes of De Pagter [87] and the M.Sc. thesis of Hytönen [50].

Several alternative proofs of the $p$-independence exist; some of them characterise the UMD$_p$-property in terms of some other property not involving the parameter $p$. In order to state two such characterisations, due to Burkholder [17, 20], we need to introduce the following terminology.

A Banach space is called a *weak UMD-space* if there exists a constant $\beta$ such that for all $L^1$-martingale difference sequences $(d_n)_{n=1}^N$, all sequences of signs $(\varepsilon_n)_{n=1}^N$, and all $r > 0$ we have

$$r\,\mathbb{P}\Big\{\Big\|\sum_{n=1}^N \varepsilon_n d_n\Big\| > r\Big\} \leqslant \beta\mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|.$$

A Banach space $E$ is called $\zeta$*-convex* if there exists a function $\zeta$ on $E \times E$, convex in both variables separately, satisfying $\zeta(0,0) > 0$ and $\zeta(x,y) \leqslant \|x+y\|$ if $\|x\| = \|y\| = 1$.

**Theorem 12.17.** *For a Banach space $E$ the following assertions are equivalent:*

(1) *$E$ is a UMD-space;*
(2) *$E$ is a weak UMD-space;*
(3) *$E$ is $\zeta$-convex.*

For Hilbert spaces one may take $\zeta(x,y) := 1 + [x,y]$. For $L^p$-spaces an explicit expression for a function $\zeta$ appears to be unknown.

The scalar version of Theorem 12.15 is due to Stein [100]. Its extension to UMD-spaces is due to Bourgain, who stated the result without proof in [12]. The proof presented here is taken from [24].

The result of Exercise 4 is due to Clément, De Pagter, Sukochev, Witvliet [24] and Berkson and Gillespie [6]. Exercise 5 is an abstract version of Bourgain's version of the Marcinkiewicz multiplier theorem [12]. Other classical multiplier theorems, such as the Mihlin multiplier theorem, can be extended to UMD-spaces as well. As was first shown by Weis [108] it is even possible to consider operator-valued multipliers; typically one has to replace boundedness assumptions by suitable $R$-boundedness assumptions. We refer to Kunstmann and Weis [61] for an overview and further references.

# 13

# Stochastic integration II: the Itô integral

We have seen in Lecture 6 how to integrate functions $\Phi : (0,T) \to \mathscr{L}(H,E)$ with respect to an $H$-cylindrical Brownian motion $W_H$. In this lecture we address the problem of extending the theory of stochastic integration to processes $\Phi : (0,T)\times\Omega \to \mathscr{L}(H,E)$. As it turns out, very satisfactory results can be obtained in the setting of UMD Banach spaces $E$. The reason for this is that in these spaces we can prove a decoupling theorem for certain martingale difference sequence which, in the context of stochastic integrals, enables us to replace $W_H$ by an independent copy $\widetilde{W}_H$. The stochastic integral of $\Phi$ with respect to $\widetilde{W}_H$ can be defined path by path using the results of Lecture 6, and the decoupling inequality allows us to translate integrability criteria for this integral to the integral with respect to $W_H$.

## 13.1 Decoupling

We begin with an abstract decoupling result for a suitable class of martingale difference sequences.

Let $1 < p < \infty$ be fixed and suppose that $(\xi_n)_{n=1}^N$ is a sequence of centred integrable random variables in $L^p(\Omega)$. We assume that a filtration $(\mathscr{F}_n)_{n=1}^N$ is given such that the following conditions are satisfied for $n = 1,\dots,N$:

(1) $\xi_n$ is $\mathscr{F}_n$-measurable for all $1 \leqslant n \leqslant N$;
(2) $\xi_n$ is independent of $\mathscr{F}_m$ for all $1 \leqslant m < n \leqslant N$.

Note that $\mathbb{E}(\xi_n|\mathscr{F}_m) = \mathbb{E}\xi_n = 0$ for $1 \leqslant m < n \leqslant N$, so $(\xi_n)_{n=1}^N$ is a martingale difference sequence with respect to $(\mathscr{F}_n)_{n=1}^N$.

On the product space $(\Omega\times\Omega, \mathscr{F}\times\mathscr{F}, \mathbb{P}\times\mathbb{P})$ we define, with a slight abuse of notation,

$$\xi_n(\omega,\widetilde{\omega}) := \xi_n(\omega), \quad \widetilde{\xi}_n(\omega,\widetilde{\omega}) := \xi_n(\widetilde{\omega}). \tag{13.1}$$

The sequences $(\xi_n)_{n=1}^N$ and $(\widetilde{\xi}_n)_{n=1}^N$ are independent and identically distributed. The point here is that we identify each $\xi_n$ with a random variable on

$\Omega \times \Omega$ which depends only on the first coordinate and introduce an independent copy $\widetilde{\xi}_n$ which depends only on the second coordinate. Clearly, $(\xi_n)_{n=1}^N$ and $(\widetilde{\xi}_n)_{n=1}^N$ are martingale difference sequences on $\Omega \times \Omega$ with respect to the filtrations $(\mathscr{F}_n)_{n=1}^N$ and $(\widetilde{\mathscr{F}}_n)_{n=1}^N$ defined by

$$\mathscr{F}_n := \mathscr{F}_n \times \{\varnothing, \Omega\}, \quad \widetilde{\mathscr{F}}_n := \{\varnothing, \Omega\} \times \mathscr{F}_n, \tag{13.2}$$

where again there is a slight abuse of notation in the first definition.

Let $(v_n)_{n=1}^N$ be a predictable sequence of $E$-valued random variables on $\Omega$. Recall that this means that $v_n$ is $\mathscr{F}_{n-1}$-measurable for $n = 1, \dots, N$, with the understanding that $\mathscr{F}_0 = \{\varnothing, \Omega\}$ (so that $v_1$ is constant almost surely). We identify $(v_n)_{n=1}^N$ with a predictable sequence $(v_n)_{n=1}^N$ on $\Omega \times \Omega$ in the same way as above by putting $v_n(\omega, \widetilde{\omega}) := v_n(\omega)$.

**Theorem 13.1 (Decoupling).** *If, in addition to the above assumptions, $E$ is a UMD-space, then*

$$\mathbb{E}\Big\| \sum_{n=1}^N \xi_n v_n \Big\|^p \eqsim_{p,E} \mathbb{E}\Big\| \sum_{n=1}^N \widetilde{\xi}_n v_n \Big\|^p$$

*with constants depending on $p$ and $E$ only.*

*Proof.* The proof uses a trick similar to that of Theorem 12.15.

For $n = 1, \dots, N$ define

$$d_{2n-1} := \tfrac{1}{2}(\xi_n + \widetilde{\xi}_n)v_n \quad \text{and} \quad d_{2n} := \tfrac{1}{2}(\xi_n - \widetilde{\xi}_n)v_n.$$

We claim that $(d_j)_{j=1}^{2N}$ is a martingale difference sequence with respect to the filtration $(\mathscr{D}_j)_{j=1}^{2N}$, where

$$\mathscr{D}_{2n-1} = \sigma(\mathscr{F}_{n-1}, \widetilde{\mathscr{F}}_{n-1}, \xi_n + \widetilde{\xi}_n), \quad \mathscr{D}_{2n} = \sigma(\mathscr{F}_n, \widetilde{\mathscr{F}}_n).$$

In view of

$$\sum_{n=1}^N \xi_n v_n = \sum_{j=1}^{2N} d_j \quad \text{and} \quad \sum_{n=1}^N \widetilde{\xi}_n v_n = \sum_{j=1}^{2N} (-1)^{j+1} d_j,$$

the result then follows from the definition of the UMD$_p$-property.

It remains to prove the claim. We begin by observing that $(d_n)_{n=1}^{2N}$ is $(\mathscr{D}_n)_{n=1}^{2N}$-adapted. Moreover,

$$\begin{aligned}\mathbb{E}(d_{2n}|\mathscr{D}_{2n-1}) &\overset{\text{(i)}}{=} \tfrac{1}{2} v_n \mathbb{E}(\xi_n - \widetilde{\xi}_n | \mathscr{F}_{n-1}, \widetilde{\mathscr{F}}_{n-1}, \xi_n + \widetilde{\xi}_n) \\ &\overset{\text{(ii)}}{=} \tfrac{1}{2} v_n \mathbb{E}(\xi_n - \widetilde{\xi}_n | \xi_n + \widetilde{\xi}_n) \overset{\text{(iii)}}{=} 0.\end{aligned}$$

Here (i) follows from the $\mathscr{F}_{n-1}$-measurability of $v_n$, (ii) from Proposition 11.7 and the independence of $\sigma(\xi_n, \widetilde{\xi}_n)$ and $\sigma(\mathscr{F}_{n-1}, \widetilde{\mathscr{F}}_{n-1})$ (which follows from the independence of $\xi_n$ and $\mathscr{F}_{n-1}$), and (iii) uses that $\xi_n$ and $\widetilde{\xi}_n$ are independent and identically distributed (Exercise 11.2). Similarly,

$$\mathbb{E}(d_{2n-1}|\mathscr{D}_{2n-2}) = \tfrac{1}{2}v_n\mathbb{E}(\xi_n + \widetilde{\xi}_n|\mathscr{F}_{n-1}, \widetilde{\mathscr{F}}_{n-1}) = \tfrac{1}{2}v_n\mathbb{E}(\xi_n + \widetilde{\xi}_n) = 0$$

since $\xi_n + \widetilde{\xi}_n$ is independent of $\sigma(\mathscr{F}_{n-1}, \widetilde{\mathscr{F}}_{n-1})$ and $\xi_n$, $\widetilde{\xi}_n$ are centred. □

## 13.2 Stochastic integration

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. A function $\Phi : (0,T) \times \Omega \to \mathscr{L}(H,E)$ is said to be a *finite rank adapted step process* with respect to a given filtration $\mathbb{F} = (\mathscr{F}_t)_{t\in[0,T]}$ if it is of the form

$$\Phi(t,\omega) = \sum_{m=1}^{M}\sum_{n=1}^{N} 1_{(t_{n-1},t_n)}(t)1_{A_{mn}}(\omega)\sum_{j=1}^{k} h_j \otimes x_{jmn}, \tag{13.3}$$

where $0 \leqslant t_0 < \cdots < t_N \leqslant T$, for each $n = 1, \ldots, N$ the sets $A_{1n}, \ldots, A_{Mn}$ are disjoint and belong to $\mathscr{F}_{t_{n-1}}$, the vectors $h_1, \ldots, h_k \in H$ are orthonormal, and the vectors $x_{jmn}$ belong to $E$.

In what follows we assume that $W_H$ is an $H$-cylindrical Brownian motion on $(\Omega, \mathscr{F}, \mathbb{P})$, adapted to $\mathbb{F}$ in the sense that the random variables $W_H(t)h$ are $\mathscr{F}_t$-measurable and the increments $W_H(t)h - W_H(s)h$ are independent of $\mathscr{F}_s$ for $t > s$. It follows from Exercise 11.4 that the filtration $\mathbb{F}^{W_H}$ generated by $W_H$ has these properties.

The *stochastic integral* with respect to $W_H$ of a finite rank adapted step process $\Phi$ of the form (13.3) is defined as

$$\int_0^T \Phi(t)\,dW_H(t) := \sum_{m=1}^{M}\sum_{n=1}^{N} 1_{A_{mn}}\sum_{j=1}^{k}(W_H(t_n)h_j - W_H(t_{n-1})h_j)x_{jmn}.$$

We leave it to the reader to check that this definition does not depend on the particular representation of $\Phi$ in (13.3). Note that $\int_0^T \Phi(t)\,dW_H(t)$ belongs to $L^p(\Omega, \mathscr{F}_T; E)$ for all $1 \leqslant p < \infty$, and satisfies

$$\mathbb{E}\int_0^T \Phi(t)\,dW_H(t) = 0.$$

The latter follows by linearity from

$$\begin{aligned}
&\mathbb{E}\big(1_{A_{mn}}(W_H(t_n)h_j - W_H(t_{n-1})h_j)\big)\\
&\qquad = \mathbb{E}\big(\mathbb{E}(1_{A_{mn}}(W_H(t_n)h_j - W_H(t_{n-1})h_j)|\mathscr{F}_{t_{n-1}})\big)\\
&\qquad = \mathbb{E}(1_{A_{mn}}\mathbb{E}(W_H(t_n)h_j - W_H(t_{n-1})h_j)|\mathscr{F}_{t_{n-1}}) = 0.
\end{aligned}$$

For each $\omega \in \Omega$ the trajectory $t \mapsto \Phi_\omega(t) := \Phi(t,\omega)$ is a finite rank step function and therewith defines an element $R_{\Phi_\omega}$ of $\gamma(L^2(0,T;H),E)$. This results in a simple random variable

$$R_\Phi : \Omega \to \gamma(L^2(0,T;H),E).$$

In order to extend the above stochastic integral to a more general class of $\mathscr{L}(H,E)$-valued processes we shall proceed as in Lecture 6 by estimating the $L^p(\Omega;E)$-norm of the stochastic integral in terms of $R_\Phi$. Due to the presence of the random variables $1_{A_{mn}}$, however, the Gaussian computation of Theorem 6.14 breaks down. In the proof of the next theorem we circumvent this problem by replacing $W_H$ by an independent copy $\widetilde{W}_H$ and use the decoupling estimate of Theorem 13.1.

**Theorem 13.2 (Itô isomorphism).** *Let $E$ be a UMD space and fix $1 < p < \infty$. For all finite rank adapted step processes $\Phi : (0,T) \times \Omega \to \mathscr{L}(H,E)$ we have*

$$\mathbb{E}\Big\| \int_0^T \Phi(t)\, dW_H(t) \Big\|^p \eqsim_{p,E} \mathbb{E}\|R_\Phi\|^p_{\gamma(L^2(0,T;H),E)},$$

*with constants depending only on $p$ and $E$.*

*Proof.* As in (13.1) we identify $W_H$ with an $H$-cylindrical Brownian motion on the product $\Omega \times \Omega$ and define an independent copy on $\widetilde{W}_H$ on $\Omega \times \Omega$ by putting

$$W_H(t)h(\omega,\widetilde{\omega}) := W_H(t)h(\omega), \quad \widetilde{W}_H(t)h(\omega,\widetilde{\omega}) := W_H(t)h(\widetilde{\omega}).$$

If $\Phi : (0,T) \times \Omega \to \mathscr{L}(H,E)$ is a finite rank adapted step process of the form (13.3), we define the decoupled stochastic integral

$$\int_0^T \Phi(t)\, d\widetilde{W}_H(t) := \sum_{n=1}^N \sum_{m=1}^M 1_{A_{mn}} \sum_{j=1}^k \big(\widetilde{W}_H(t_n)h_j - \widetilde{W}_H(t_{n-1})h_j\big) x_{jmn}.$$

The plan of the proof is to apply Theorem 13.1 to the real-valued sequence $(\xi_{jn})_{\substack{1\leqslant j\leqslant k\\ 1\leqslant n\leqslant N}}$ and the $E$-valued sequence $(v_{jn})_{\substack{1\leqslant j\leqslant k\\ 1\leqslant n\leqslant N}}$,

$$\xi_{jn} := W_H(t_n)h_j - W_H(t_{n-1})h_j, \quad v_{jn} := \sum_{m=1}^M 1_{A_{mn}} \otimes x_{jmn}.$$

With these notations,

$$\int_0^T \Phi(t)\, dW_H(t) = \sum_{n=1}^N \sum_{j=1}^k \xi_{jn} v_{jn}, \quad \int_0^T \Phi(t)\, d\widetilde{W}_H(t) = \sum_{n=1}^N \sum_{j=1}^k \widetilde{\xi}_{jn} v_{jn}.$$

We consider the filtration $(\mathscr{F}_{jn})_{\substack{1\leqslant j\leqslant k\\ 1\leqslant n\leqslant N}}$, where $\mathscr{F}_{jn}$ is the $\sigma$-algebra generated by all $\xi_{j'n'}$ with $(j',n')\leqslant(j,n)$; the pairs are ordered lexicographically according to the rule $(j',n')\leqslant(j,n)\Longleftrightarrow n'<n$ or $[n'=n \;\&\; j'\leqslant j]$.

With respect to this filtration, the sequence $(\xi_{jn})_{\substack{1\leqslant j\leqslant k\\ 1\leqslant n\leqslant N}}$ is centred and has the properties (1) and (2) stated at the beginning of Section 13.1 and $(v_{jn})^N_{\substack{1\leqslant j\leqslant k\\ 1\leqslant n\leqslant N}}$ is predictable.

Let us denote by $\mathbb{E}_1$ and $\mathbb{E}_2$ the expectations with respect to the first and second coordinate of $\Omega\times\Omega$. Applying successively Theorem 13.1, the Kahane-Khintchine inequality, and Theorem 6.14 (pointwise with respect to $\Omega_1$), we obtain

$$\begin{aligned}\mathbb{E}_1\mathbb{E}_2\Big\|\int_0^T \Phi(t)\,dW_H(t)\Big\|^p &\eqsim_{p,E} \mathbb{E}_1\mathbb{E}_2\Big\|\int_0^T \Phi(t)\,d\widetilde{W}_H(t)\Big\|^p\\ &\eqsim_{p,E} \mathbb{E}_1\Big(\mathbb{E}_2\Big\|\int_0^T \Phi(t)\,d\widetilde{W}_H(t)\Big\|^2\Big)^{\frac{p}{2}}\\ &\eqsim_{p,E} \mathbb{E}_1\|R_\Phi\|^p_{\gamma(L^2(0,T;H),E)}. \qquad \square\end{aligned}$$

**Definition 13.3.** *A random variable $R\in L^p(\Omega;\gamma(L^2(0,T;H),E))$ is called* adapted *if it belongs to the closure in $L^p(\Omega;\gamma(L^2(0,T;H),E))$ of the finite rank adapted step processes.*

The closed subspace in $L^p(\Omega;\gamma(L^2(0,T;H),E))$ of all adapted elements with be denoted by $L^p_{\mathbb{F}}(\Omega;\gamma(L^2(0,T;H),E))$. Theorem 13.2 shows that the stochastic integral extends uniquely to an isomorphic embedding

$$J_T^{W_H} : L^p_{\mathbb{F}}(\Omega;\gamma(L^2(0,T;H),E))\to L^p(\Omega;E).$$

**Definition 13.4.** *Let $E$ be a Banach space and fix $1<p<\infty$. A process $\Phi:(0,T)\times\Omega\to\mathscr{L}(H,E)$ is said to be* $L^p$-stochastically integrable *with respect to the $H$-cylindrical Brownian motion $W_H$ if there exists a sequence of finite rank adapted step processes $\Phi_n:(0,T)\times\Omega\to\mathscr{L}(H,E)$ such that:*

(1) *for all $h\in H$ we have $\lim_{n\to\infty}\Phi_n h=\Phi h$ in measure;*

(2) *there exists a random variable $X\in L^p(\Omega;E)$ such that $\lim_{n\to\infty}\int_0^T\Phi_n\,dW_H = X$ in $L^p(\Omega;E)$.*

*The* $L^p$-stochastic integral *of $\Phi$ is then defined as*

$$\int_0^T\Phi\,dW_H := \lim_{n\to\infty}\int_0^T\Phi_n\,dW_H.$$

The remarks (a) and (b) following Definition 6.15 extend to the present situation, but (c) is no longer automatic since stochastic integrals of step processes are no longer Gaussian. This is the reason why the adjective '$L^p$-' has been built into the definition.

**Theorem 13.5.** *Let $1 < p < \infty$. If $\Phi : (0,T) \times \Omega \to \mathscr{L}(H,E)$ is $L^p$-stochastically integrable with respect to $W_H$, then the stochastic integral process $\big(\int_0^t \Phi\, dW_H\big)_{t\in[0,T]}$ is an $E$-valued $L^p$-martingale which has a continuous version satisfying the maximal inequality*

$$\mathbb{E}\Big(\sup_{t\in[0,T]} \Big\| \int_0^t \Phi\, dW_H \Big\|^p\Big) \leqslant \big(\frac{p}{p-1}\big)^p\, \mathbb{E}\Big\| \int_0^T \Phi\, dW_H \Big\|^p.$$

*Proof.* Choose a sequence $(\Phi_n)_{n\geqslant 1}$ of finite rank adapted step processes such that the conditions of Definition 13.4 are satisfied and put $X_n(t) := \int_0^t \Phi_n\, dW_H$. Clearly, there exists a continuous version $\widetilde{X}_n$ of each $X_n$, and by the Pettis measurability theorem we have $\widetilde{X}_n \in L^p(\Omega; C([0,T];E))$. To see that this theorem can be applied in the present situation, first note that there exists a separable closed subspace $E_0$ of $E$ such that each $X_n$ has trajectories in $C([0,T];E_0)$. The space $C([0,T];E_0)$ is separable, and the linear span of the functionals $\delta_t \otimes x^*$ is norming in its dual; moreover, $\langle X_n, \delta_t \otimes x^*\rangle = \int_0^t \Phi_n^* x^*\, dW_H$ almost surely and the right hand side is measurable as a function on $\Omega$.

By Doob's maximal inequality (we use that the stochastic integral process is a martingale; see Exercise 3), for every choice of $0 \leqslant t_1 < \dots < t_N \leqslant T$ we have

$$\mathbb{E}\big(\sup_{j=1,\dots,N} \|\widetilde{X}_n(t_j) - \widetilde{X}_m(t_j)\|^p\big) \leqslant \big(\frac{p}{p-1}\big)^p\, \mathbb{E}\|X_n(T) - X_m(T)\|^p.$$

Hence, by path continuity and Fatou's lemma,

$$\mathbb{E}\big(\sup_{t\in[0,T]} \|\widetilde{X}_n(t) - \widetilde{X}_m(t)\|^p\big) \leqslant \big(\frac{p}{p-1}\big)^p\, \mathbb{E}\|X_n(T) - X_m(T)\|^p.$$

This inequality shows that the sequence $(\widetilde{X}_n)_{n\geqslant 1}$ is a Cauchy sequence in $L^p(\Omega; C([0,T];E))$. Since for all $t \in [0,T]$ we have $\lim_{n\to\infty} X_n(t) = X(t)$ in $L^p(\Omega;E)$, the limit $\widetilde{X} = \lim_{n\to\infty} \widetilde{X}_n$ defines a continuous version of $X$.

The final inequality follows from Doob's maximal inequality in the same way as above (replace $\widetilde{X}_n - \widetilde{X}_m$ by $\widetilde{X}$). $\square$

As in Lecture 6, in the special case $E = \mathbb{R}$ we may identify $\mathscr{L}(H,\mathbb{R})$ with $H$ and Theorem 13.2 reduces to the statement that the $L^p$-stochastic integral of an adapted step process $\phi : (0,T) \times \Omega \to H$ satisfies

$$\mathbb{E}\Big\| \int_0^T \phi\, dW_H \Big\|^p \eqsim_p \mathbb{E}\|\phi\|^p_{L^2(0,T;H)} \tag{13.4}$$

The constants depend only on $p$ since the UMD$_p$-constant of Hilbert spaces only depend on $p$. From this it is not hard to see (Exercise 2) that a strongly adapted measurable process $\phi : (0,T)\times\Omega \to H$ is $L^p$-stochastically integrable with respect to $W_H$ if and only if $\phi \in L^p(\Omega; L^2(0,T;H))$, and the isomorphism (13.4) extends to this situation.

**Definition 13.6.** *A process $\Phi : (0,T)\times\Omega \to \mathscr{L}(H,E)$ is called* $H$-strongly measurable *if for each $h \in H$ the process $\Phi h : (0,T)\times\Omega \to E$ is strongly measurable. Such a process $\Phi$ is called* adapted *if for each $h \in H$ the process $\Phi h$ is adapted.*

We are now in a position to state the main result of this section, which extends Theorem 6.17 to $\mathscr{L}(H,E)$-valued processes.

**Theorem 13.7.** *Let $E$ be a UMD space and fix $1 < p < \infty$. For an $H$-strongly measurable adapted process $\Phi : (0,T)\times\Omega \to \mathscr{L}(H,E)$ the following assertions are equivalent:*

(1) *$\Phi$ is $L^p$-stochastically integrable with respect to $W_H$;*
(2) *$\Phi^* x^* \in L^p(\Omega; L^2(0,T;H))$ for all $x^* \in E^*$, and there exists a random variable $X \in L^p(\Omega;E)$ such that for all $x^* \in E^*$,*

$$\langle X, x^*\rangle = \int_0^T \Phi^* x^* \, dW_H(t) \quad \text{in } L^p(\Omega).$$

(3) *$\Phi^* x^* \in L^p(\Omega; L^2(0,T;H))$ for all $x^* \in E^*$, and there exists a random variable $R \in L^p(\Omega; \gamma(L^2(0,T;H),E))$ such that for all $f \in L^2(0,T;H)$ and $x^* \in E^*$,*

$$\langle Rf, x^*\rangle = \int_0^T \langle \Phi(t) f(t), x^*\rangle \, dt \quad \text{in } L^p(\Omega).$$

*If these equivalent conditions are satisfied, the random variables $X$ and $R$ are uniquely determined, we have $X = \int_0^T \Phi \, dW_H$ in $L^p(\Omega;E)$, and*

$$\mathbb{E}\Big\| \int_0^T \Phi \, dW_H \Big\|^p \eqsim_{p,E} \mathbb{E}\|R\|^p_{\gamma(L^2(0,T;H),E)}.$$

*Moreover, $R \in L^p_{\mathbb{F}}(\Omega; \gamma(L^2(0,T;H),E))$, that is, $R$ is adapted.*

*Proof.* We sketch the main steps and refer to the Notes for more information.

(1)⇒(2): This is proved in the same way as in Theorem 6.17. Note that the stochastic integrals $\int_0^T \Phi^* x^* \, dW_H$ are well-defined by the above remarks.

(2)⇒(3): For the special case where $\mathbb{F}$ is the filtration generated by $W_H$, a proof will be outlined below.

(1)⇒(3): This is an immediate consequence of Theorem 13.2: if $(\Phi_n)_{n=1}^\infty$ is an approximating sequence for $\Phi$, then the operators $(R_{\Phi_n})_{n=1}^\infty$ form a Cauchy sequence in $L^p_{\mathbb{F}}(\Omega; \gamma(L^2(0,T;H),E))$ and its limit has the desired properties.

(3)⇒(1): First one shows that $R$ is adapted (see Exercise 1). Knowing this, the proof can be finished in the same way as the corresponding implication of Theorem 6.17. □

Unfortunately we are not able to give a fully self-contained proof of the implication (2)⇒(3). In the sequel we shall not need this implication; we only use the equivalence (1)⇔(3) which is the most useful part of the theorem. In spite of this we want to sketch a proof of (2)⇒(3) under the simplifying assumption that the filtration is the one generated by $W_H$. In this situation we can apply a version of the so-called martingale representation theorem for $H$-cylindrical Brownian motions $W_H$. In most textbook proofs, the integrator is a Brownian motion (or a more general martingale); the extension to cylindrical Brownian motions is obtained from it by an approximation argument as in the proof of the martingale convergence theorem (Theorem 11.21).

Recall that the filtration $\mathbb{F}^{W_H}$ has been defined in Exercise 11.4.

**Lemma 13.8.** *Let $1 < p < \infty$ and $\xi \in L^p(\Omega, \mathscr{F}_T^{W_H})$. There exists unique $\phi \in L^p_{\mathbb{F}^{W_H}}(\Omega; L^2(0,T;H))$ such that*

$$\xi = \mathbb{E}\xi + \int_0^T \phi \, dW_H.$$

The proof of this lemma is beyond the scope of these lectures. Roughly speaking it proceeds like this. First, we may assume that $\mathbb{E}\xi = 0$. By approximation we may further assume that $H$ is finite-dimensional. From

$$W_H(t)h = \int_0^T 1_{(0,t)} \otimes h \, dW_H(t)$$

we see that every $X$ in the linear span of the random variables $W_H(t)h$ can be represented by a stochastic integral. Since the stochastic integral defines an isomorphic embedding, it remains to show that this span is dense in the closed subspace of $L^p(\Omega, \mathscr{F}_T^{W_H}; H)$ consisting of all mean 0 elements.

The next result extends the lemma to UMD spaces. Recall that

$$J_T^{W_H} : L^p_{\mathbb{F}}(\Omega; \gamma(L^2(0,T;H), E)) \to L^p(\Omega; E)$$

is the isomorphic embedding of Theorem 13.2.

**Theorem 13.9.** *Let $E$ be a UMD space, let $1 < p < \infty$, and let $X \in L^p(\Omega, \mathscr{F}_T^{W_H}; E)$. There exists a unique $R \in L^p_{\mathbb{F}^{W_H}}(\Omega; \gamma(L^2(0,T;H), E))$ such that*

$$X = \mathbb{E}X + J_T^{W_H}(R).$$

*Proof.* Choose a sequence of simple $\mathscr{F}_T^{W_H}$-measurable random variables $X_n$ such that $\lim_{n\to\infty} X_n = X$ in $L^p(\Omega; E)$. Let us write $X_n = \sum_{m=1}^{M_n} 1_{A_{mn}} \otimes x_{mn}$.

By Lemma 13.8, there exist unique processes $\phi_{mn} \in L^p_{\mathbb{F}^{W_H}}(\Omega; L^2(0,T;H))$ such that

$$1_{A_{mn}} = \mathbb{E}1_{A_{mn}} + \int_0^T \phi_{mn}\, dW_H.$$

Put $\Phi_n(t)h := \sum_{m=1}^M [\phi_{mn}, h] x_{mn}$. The process $\Phi_n : (0,T)\times\Omega \to \mathscr{L}(H,E)$ is $L^p$-stochastically integrable with respect to $W_H$ and

$$X_n = \mathbb{E}X_n + \int_0^T \Phi_n\, dW_H.$$

Let $R_n \in L^p_{\mathbb{F}}(\Omega; \gamma(L^2(0,T;H),E)$ be defined by

$$R_n(\omega)f := \sum_{m=1}^{M_n} \phi_{mn}(\omega)\otimes x_{mn}, \quad f \in L^2(0,T;H).$$

Since $\lim_{n\to\infty} X_n = X$ in $L^p(\Omega;E)$, the isomorphism of Theorem 13.2 implies that the sequence $(R_n)_{n=1}^\infty$ is Cauchy in $L^p(\Omega;\gamma(L^2(0,T;H),E))$. The limit $R$ has the desired properties.

Uniqueness follows from the injectivity of $J_T^{W_H}$. □

As a corollary we observe that the stochastic integral defines an isomorphism of Banach spaces

$$J_T^{W_H} : L^p_{\mathbb{F}^{W_H}}(\Omega;\gamma(L^2(0,T;H),E)) \simeq L^p_0(\Omega,\mathscr{F}_T^{W_H};E),$$

where $L^p_0(\Omega,\mathscr{F}_T^{W_H};E)$ is the closed subspace of $L^p(\Omega,\mathscr{F}_T^{W_H};E)$ consisting of all elements with mean 0.

*Proof (Proof of Theorem 13.7 (2)⇒(3) for the filtration $\mathbb{F}^{W_H}$).* By the Pettis measurability theorem, the random variable $X$ belongs to $L^p_0(\Omega,\mathscr{F}_T^{W_H};E)$. The element $R$ provided by Theorem 13.9 has the desired properties. □

## 13.3 Stochastic integrability of $L^p$-martingales

We return to the setting where $W_H$ is an $H$-cylindrical Brownian motion, adapted to a filtration $\mathbb{F}$. The main result of this section states that if $E$ is a UMD space and $M$ is a $\gamma(H,E)$-valued $L^p$-martingale with respect to $\mathbb{F}$, then $M$ is $L^p$-stochastically integrable with respect to $W^H$. The proof has three ingredients: the characterisation of $L^p$-stochastic integrability (the equivalence (1)⇔(3) of Theorem 13.7)), the $\gamma$-multiplier theorem (Theorem 9.13), and the vector-valued Stein inequality (Theorem 12.15).

**Theorem 13.10.** *Let $E$ be a UMD space and fix $1 < p < \infty$. Let $W_H$ be an $H$-cylindrical Brownian motion, adapted to a filtration $\mathbb{F}$, and let $M$ :*

*$[0,T] \times \Omega \to \gamma(H,E)$ be an $L^p$-martingale with respect to $\mathbb{F}$. Then $M$ is $L^p$-stochastically integrable with respect to $W_H$ and we have*

$$\Big(\mathbb{E}\Big\|\int_0^T M(t)\,dW_H(t)\Big\|^p\Big)^{\frac{1}{p}} \lesssim_{p,E} \sqrt{T}\,\big(\mathbb{E}\|M(T)\|^p_{\gamma(H,E)}\big)^{\frac{1}{p}},$$

*with a constant depending only on $p$ and $E$.*

*Proof.* First we prove the result under the additional assumption that $M(T) \in L^\infty(\Omega;\gamma(H,E))$. By the $L^\infty$-contractivity of the conditional expectation we then have $M \in L^\infty((0,T)\times\Omega;\gamma(H,E))$. In particular, for all $x^* \in E^*$ we have $M^*x^* \in L^p(\Omega;L^2(0,T;H))$, and even $M^*x^* \in L^p_{\mathbb{F}}(\Omega;L^2(0,T;H))$ since $M$ is adapted.

Let us write $B := L^p(\Omega;E)$ for brevity. Define the bounded function $N : [0,T] \to \mathscr{L}(B)$ by

$$N(t)\xi := \mathbb{E}(\xi|\mathscr{F}_t), \qquad \xi \in B,\ t \in [0,T].$$

Since $E$ is a UMD space, by Theorem 12.15 the family $\{N(t) : t \in [0,T]\}$ is $R$-bounded on $B$, and therefore $\gamma$-bounded, with $\gamma$-bound depending only on $p$ and $E$. By Theorem 11.21, for every $\xi \in B$ the function $t \mapsto N(t)\xi$ has left limits at every point $[0,T]$. In particular, these functions are strongly measurable.

By the $\gamma$-Fubini isomorphism (Theorem 5.22), for each $t \in [0,T]$ we may identify the random variable $M(t) \in L^p(\Omega;\gamma(H,E))$ with a unique operator $\widetilde{M}(t) \in \gamma(H,B)$ by the formula $(\widetilde{M}(t)h)(\omega) = M(t,\omega)h$. Define a constant function $G : [0,T] \to \gamma(H,B)$ by

$$G(t) := \widetilde{M}(T), \quad t \in [0,T].$$

This function represents the element $R_G \in \gamma(L^2(0,T;H),B)$ satisfying

$$\|R_G\|_{\gamma(L^2(0,T;H),B)} = \sqrt{T}\,\|\widetilde{M}(T)\|_{\gamma(H,B)} \eqsim_p \sqrt{T}\,\big(\mathbb{E}\|M(T)\|^p_{\gamma(H,E)}\big)^{\frac{1}{p}},$$

where we used the result of Exercise 5.3.

By the martingale property, for all $t \in [0,T]$ we have $\widetilde{M}(t) = N(t)\widetilde{M}(T)$ in $B$. Now we apply Theorem 9.13 to conclude that $\widetilde{M}$ represents an element $R \in \gamma(L^2(0,T;H),B)$ satisfying

$$\|R\|_{\gamma(L^2(0,T;H),B)} \lesssim_{p,E} \|R_G\|_{\gamma(L^2(0,T;H),B)}.$$

Using Theorem 5.22 once more, we can identify $R$ with an element $X \in L^p(\Omega;\gamma(L^2(0,T;H),E))$ by the formula $X(\omega)f = (Rf)(\omega)$. Below we check that $X^*x^* = M^*x^*$ in $L^p(\Omega;L^2(0,T;H))$ for all $x^* \in E^*$. Assuming this for the moment, it follows from Theorem 13.7(3)⇒(1) that $M$ is $L^p$-stochastically integrable and satisfies

$$\Big(\mathbb{E}\Big\|\int_0^T M(t)\,dW_H(t)\Big\|^p\Big)^{\frac{1}{p}} \eqsim_{p,E} \big(\mathbb{E}\|X\|^p_{\gamma(L^2(0,T;H),E)}\big)^{\frac{1}{p}}$$
$$\eqsim_p \|R\|_{\gamma(L^2(0,T;H),B)} \lesssim_{p,E} \sqrt{T}\,\big(\mathbb{E}\|M(T)\|^p_{\gamma(H,E)}\big)^{\frac{1}{p}}.$$

To prove that $X^*x^* = M^*x^*$ for all $x^* \in E^*$, let $f \in L^2(0,T;H)$ and $x^* \in E^*$ be arbitrary and note that for all $A \in \mathscr{F}$,

$$\begin{aligned}
\mathbb{E}(\langle Mf, x^*\rangle 1_A) &= \int_\Omega \int_0^T \langle M(t,\omega)f(t), x^*\rangle 1_A(\omega)\,dt\,dP(\omega)\\
&= \int_0^T \int_\Omega \langle M(t,\omega)f(t), x^*\rangle 1_A(\omega)\,dP(\omega)\,dt\\
&= \mathbb{E}\int_0^T \langle \widetilde{M}(t)f(t), 1_A \otimes x^*\rangle\,dt\\
&= \mathbb{E}\langle Rf, 1_A\otimes x^*\rangle = \mathbb{E}(\langle Xf, x^*\rangle 1_A).
\end{aligned}$$

To conclude the proof we remove the assumption $M(T) \in L^\infty(\Omega;E)$. Choose a sequence of $\mathscr{F}_T$-measurable simple random variables $M_n(T)$ converging to $M(T)$ in $L^p(\Omega;E)$, and define $M_n(t) := \mathbb{E}(M_n(T)|\mathscr{F}_t)$. Since $M_n(T) \in L^\infty(\Omega;E)$, we may apply what we proved above to the martingales $M_n$. We obtain that each $M_n$ is $L^p$-stochastically integrable with respect to $W_H$ and

$$\Big(\mathbb{E}\Big\|\int_0^T M_n(t)\,dW_H(t)\Big\|^p\Big)^{\frac{1}{p}} \leqslant C\sqrt{T}\,\big(\mathbb{E}\|M_n(T)\|^p_{\gamma(H,E)}\big)^{\frac{1}{p}},$$

with a constant $C$ independent of $n$. Similarly, by the above applied to the martingales $M_n - M_m$, we find that the stochastic integrals $\int_0^T M_n(t)\,dW_H(t)$ are Cauchy in $L^p(\Omega;E)$. By the Itô isomorphism, this means that the corresponding elements $R_n \in L^p_{\mathbb{F}}(\Omega;\gamma(L^2(0,T;H),E))$ are Cauchy and therefore converge to a limit $R \in L^p_{\mathbb{F}}(\Omega;\gamma(L^2(0,T;H),E))$. Clearly, $R^*x^* = \lim_{n\to\infty} R_n^*x^* = \lim_{n\to\infty} M_n^*x^* = M^*x^*$ in $L^p(\Omega;L^2(0,T;H))$, and the conclusion of the theorem now follows via Theorem 13.7. $\square$

## 13.4 Exercises

In the exercises 1-3 we fix $1 < p < \infty$.

1. In this exercise we compare the two notions of adaptedness given in Definitions 13.3 and 13.6.
   a) Show that $R \in L^p(\Omega;\gamma(L^2(0,T;H),E))$ is adapted if and only if the random variables $R(1_{(0,t)}f) : \Omega \to E$ have strongly $\mathscr{F}_t$-measurable versions for all $t \in (0,T)$ and $f \in L^2(0,T;H)$.
   *Hint:* For the 'if' part, approximate with simple random variables and use that the finite rank step functions are dense in $\gamma(L^2(0,T;H),E)$. To secure adaptedness, build in a small shift before approximating.

Now suppose that $\Phi : (0,T) \to \mathscr{L}(H,E)$ is $H$-strongly measurable and assume that the conditions of Theorem 13.7 (3) be satisfied; let $R \in L^p(\Omega;\gamma(L^2(0,T;H),E))$ be as in (3).

b) Show that if $\Phi$ is adapted, then $R$ is adapted.

2. In the discussion after Definition 13.4 it was observed that a strongly measurable adapted process $\phi : (0,T)\times\Omega \to H$ is $L^p$-stochastically integrable with respect to $W_H$ if and only if $\phi \in L^p(\Omega;L^2(0,T;H))$. Prove this.
*Hint:* If $\phi \in L^p(\Omega;L^2(0,T;H))$, then by the previous exercise $\phi$ is adapted as an element of $L^p(\Omega;L^2(0,T;H))$.

3. Let $\Phi : (0,T)\times\Omega \to \mathscr{L}(H,E)$ be $L^p$-stochastically integrable with respect to $W_H$. Show that the stochastic integral process $\big(\int_0^t \Phi\, dW_H\big)_{t\in[0,T]}$ is a martingale.
*Hint:* Approximate with finite rank adapted step processes.

If $E$ is a UMD space with type 2 and $\Phi : (0,T)\times\Omega \to \gamma(H,E)$ is an adapted and strongly measurable process such that

$$\mathbb{E}\int_0^T \|\Phi(t)\|^2_{\gamma(H,E)}\, dt < \infty,$$

then $\Phi$ is stochastically integrable with respect to $H$-cylindrical Brownian motions $W_H$; this follows from Theorem 13.7 and Exercise 5.4. In the next two exercises we show that the UMD assumption can essentially be dropped from this statement.

4. A Banach space $E$ has *martingale type* $p \in [1,2]$ if there exists a constant $M_p(E)$ such that for any $L^p$-martingale sequence $(d_n)_{n=1}^N$ with values in $E$ we have
$$\mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|^p \leqslant (M_p(E))^p \sum_{n=1}^N \mathbb{E}\|d_n\|^p.$$
a) Show that every martingale type $p$ space has type $p$.
b) Show that every UMD space with type $p$ has martingale type $p$.

In both cases, give relations between the constants.

c) Deduce that $L^p$-spaces, $1<p<\infty$, have martingale type $\min\{p,p'\}$, where $\frac{1}{p}+\frac{1}{p'}=1$.

5. Let $E$ be a martingale type 2 space.
a) Show that if $W_H$ is an $H$-cylindrical Brownian motion and $\Phi : (0,T)\times\Omega \to \mathscr{L}(H,E)$ is an adapted finite rank step process, then
$$\mathbb{E}\Big\|\int_0^T \Phi\, dW_H\Big\|^2 \leqslant (M_2(E))^2\, \mathbb{E}\int_0^T \|\Phi(t)\|^2_{\gamma(H,E)}\, dt.$$

b) Conclude that if $\Phi : (0,T) \times \Omega \to \gamma(H,E)$ is an adapted strongly measurable process satisfying

$$\mathbb{E}\int_0^T \|\Phi(t)\|^2 \, dt < \infty,$$

then $\Phi$ is stochastically integrable with respect to $W_H$, with the same estimate as before.

**Notes.** A systematic treatment of decoupling inequalities is presented in the monograph of Gine and de la Peña [31]. The proof of the decoupling inequality (Theorem 13.1) is based on an idea of Montgomery-Smith [78].

The idea to use decoupling inequalities for obtaining bounds on stochastic integrals in UMD spaces was first used by Garling [40], who only considered step processes and used the resulting estimates to investigate certain geometric properties of UMD spaces. Using a more delicate decoupling result together with Burkholder's characterisation of UMD spaces through $\zeta$-convexity (Theorem 12.17), McConnell [75] proved that a UMD-valued process is stochastically integrable if almost surely its trajectories are stochastically integrable with respect to an independent copy of the Brownian motion. In view of Theorem 6.17 this result can be viewed as an 'almost sure' version of the implication (3)⇒(1) of Theorem 13.7.

Our approach to vector-valued stochastic integration in UMD spaces via $\gamma$-radonifying norms is taken from [82], where Theorem 13.7 was proved. In that paper, McConnell's result is recovered using a stopping time argument.

The equivalence of norms in (13.4) is a special case of an inequality of Burkholder, Davis, Gundy which, in the more general situation where the integrator is a continuous-time martingale $M$, relates the norms of stochastic integrals to the norms of the quadratic variation process of $M$. For more details we refer to Karatzas and Shreve [59], Revuz and Yor [94], or Kallenberg [55].

An alternative proof of the implication (2)⇒(3) of Theorem 13.7, which is based on finite-dimensional approximations, covariance domination, and the theorem of Hoffmann-Jorgensen and Kwapień (Theorem 5.9) is given in [82]. A detailed proof of the implication (3)⇒(1) is contained in [81].

A systematic theory of stochastic integration in martingale type 2 space has been developed by Neidhardt [85], Dettweiler [33], and Brzeźniak [13]. The first two authors assumed that $E$ be 2-uniformly smooth, a property which was subsequently shown to be equivalent to the martingale type 2 property by Pisier [91]. For an overview, see Brzeźniak [15].

# 14

# Linear equations with multiplicative noise

In this lecture we study stochastic evolution equations with multiplicative noise of the form

$$\begin{cases} dU(t) = AU(t)\,dt + B(U(t))\,dW_H(t), & t \in [0,T], \\ U(0) = u_0. \end{cases} \tag{SCP}$$

Under suitable assumptions on $E$, the semigroup $S$ generated by $A$ on $E$, and the function $B : E \to \gamma(H,E)$, we shall prove existence, uniqueness, and Hölder regularity of mild solutions. Such a solution is defined as an adapted process $U$ such that for all $t \in [0,T]$ we have

$$U(t) = S(t)u_0 + \int_0^t S(t-s)B(U(s))\,dW_H(s) \tag{14.1}$$

almost surely. Its existence and uniqueness is proved by a fixed point argument in the completion $V_\theta^p(\Omega;\gamma(L^2(0,T),E))$ of the space of all adapted finite rank step processes $\phi : (0,T) \times \Omega \to E$ such that

$$s \mapsto (t-s)^{-\theta}\phi(s) \text{ belongs to } L^p(\Omega;\gamma(L^2(0,t),E)),$$

uniformly with respect to $0 < t \leqslant T$. The reason for working in this complicated space is the fact that in many applications (e.g. when $S$ is an analytic semigroup) the set $\{t^\theta S(t) : \ t \in (0,T)\}$ is $\gamma$-bounded.

The strategy for the fixed point argument is as follows. First, we find conditions on $B$ which guarantee that it acts as a Lipschitz map from $V_\theta^p(\Omega;\gamma(L^2(0,T),E))$ to $V_\theta^p(\Omega;\gamma(L^2(0,T;H),E))$. Note that under these conditions, the stochastic integrals in (14.1) are well defined by the results of the previous lecture. Then, we prove that the process on the right hand side of (14.1) is in $V_\theta^p(\Omega;\gamma(L^2(0,T),E))$ again.

## 14.1 $\gamma$-Lipschitz functions

Let $H$ be a non-zero Hilbert space, let $E$ and $F$ be Banach spaces, and let $(\gamma_{mn})_{m,n=1}^\infty$ and $(\gamma_n)_{n=1}^\infty$ be Gaussian sequences.

**Proposition 14.1.** *Let $B : E \to \gamma(H,F)$ be a function such that $Bh : E \to F$ is strongly measurable for all $h \in H$, and let $C \geqslant 0$ be a constant. The following assertions are equivalent:*

(1) *for all orthonormal sequences $(h_m)_{m=1}^M$ in $H$ and all sequences $(x_n)_{n=1}^N$ and $(y_n)_{n=1}^N$ in $E$,*

$$\mathbb{E}\Big\| \sum_{m=1}^M \sum_{n=1}^N \gamma_{mn}(B(x_n) - B(y_n))h_m\Big\|^2 \leqslant C^2 \mathbb{E}\Big\| \sum_{n=1}^N \gamma_n (x_n - y_n)\Big\|^2;$$

(2) *for all simple functions $\phi_1, \phi_2 : (0,T) \to E$ we have $B(\phi_1)$, $B(\phi_2) \in \gamma(L^2(0,T;H),F)$ and*

$$\|B(\phi_1) - B(\phi_2)\|_{\gamma(L^2(0,T;H),F)} \leqslant C\|\phi_1 - \phi_2\|_{\gamma(L^2(0,T),E)};$$

(3) *for all $\sigma$-finite measure spaces $(A, \mathscr{A}, \mu)$ and all $\mu$-simple functions $\phi_1, \phi_2 : A \to E$ we have $B(\phi_1)$, $B(\phi_2) \in \gamma(L^2(A;H),F)$ and*

$$\|B(\phi_1) - B(\phi_2)\|_{\gamma(L^2(A;H),F)} \leqslant C\|\phi_1 - \phi_2\|_{\gamma(L^2(A),E)}.$$

Note that if $H$ is separable, then $Bh : E \to F$ is strongly measurable for all $h \in H$ if and only if $B : E \to \gamma(H,F)$ is strongly measurable; this is proved in Proposition 5.14 (with strong $\mu$-measurability replaced by strong measurability).

*Proof.* Let us first prove that (1) is equivalent to

(1′) for all orthonormal sequences $(h_m)_{m=1}^M$ in $H$, all sequences $(a_n)_{n=1}^N$ of positive real numbers and all sequences $(x_n)_{n=1}^N$ and $(y_n)_{n=1}^N$ in $E$,

$$\mathbb{E}\Big\| \sum_{m=1}^M \sum_{n=1}^N a_n\gamma_{mn}(B(x_n) - B(y_n))h_m\Big\|^2 \leqslant C^2 \mathbb{E}\Big\| \sum_{n=1}^N a_n\gamma_n (x_n - y_n)\Big\|^2.$$

For integers $a_n$, the equivalence follows by applying (1) with the $x_n$ and $y_n$ repeated $a_n$ times and noting that the sum of $a_n$ independent standard Gaussians $\gamma_n^{(1)} + \cdots + \gamma_n^{(a_n)}$ has the same distribution as $a_n\gamma_n$. The case of rational $a_n$ is readily reduced to this, and the general case follows by approximation.

The equivalence of (1′), (2), (3) follows from the following general observation. Let $(A, \mathscr{A}, \mu)$ be any $\sigma$-finite measure space. If $(h_m)_{m=1}^M$ is orthonormal in $H$ and $\phi_1 = \sum_{n=1}^N 1_{A_n} \otimes x_n$ and $\phi_2 = \sum_{n=1}^N 1_{A_n} \otimes y_n$ are $\mu$-simple functions with values in $E$, with the sets $A_n$ disjoint, then by Lemma 5.7 (noting that the sequence $(\frac{1}{\sqrt{\mu(A_n)}}1_{A_n})_{n=1}^N$ is orthonormal in $L^2(A)$),

$$\|B(\phi_1)-B(\phi_2)\|^2_{\gamma(L^2(A;H),F)} = \mathbb{E}\Big\| \sum_{m=1}^{M}\sum_{n=1}^{N} \sqrt{\mu(A_n)}\gamma_{nm}(B(x_n)-B(y_n))h_m\Big\|^2$$

and

$$\|\phi_1-\phi_2\|^2_{\gamma(L^2(A),E)} = \mathbb{E}\Big\| \sum_{n=1}^{N} \sqrt{\mu(A_n)}\gamma_n(x_n-y_n)\Big\|^2. \qquad \square$$

Note that if $(A,\mathscr{A},\mu)$ is a $\sigma$-finite measure space with $\mu(A)\neq 0$ and $B: E\to \mathscr{L}(H,F)$ is a function such that $B(\phi)\in\gamma(L^2(A;H),F)$ for every $\mu$-simple function $\phi: A\to E$, then $B(x)\in\gamma(H,F)$ for all $x\in E$. Indeed, consider any set $A_0\in\mathscr{A}$ with $0<\mu(A_0)<\infty$. Then $B(1_{A_0}\otimes x) = 1_{A_0}\otimes B(x)$ belongs to $\gamma(L^2(A;H),F)$, which is only possible if $B(x)\in\gamma(H,F)$. This explains why we restrict ourselves to functions $B: E\to\gamma(H,F)$.

**Definition 14.2.** *A strongly measurable function $B: E\to\gamma(H,F)$ is called* $\gamma$-Lipschitz continuous *if the equivalent conditions of Proposition 14.1 hold. The least possible constant in these conditions is denoted by* $\mathrm{Lip}_\gamma(B)$.

By taking $H=\mathbb{R}$ we obtain the notion of a $\gamma$-Lipschitz continuous function from $E$ to $F$. Clearly, every $\gamma$-Lipschitz continuous function $B: E\to F$ is Lipschitz continuous and we have $\mathrm{Lip}(B)\leqslant \mathrm{Lip}_\gamma(B)$.

It is a natural question whether conversely Lipschitz functions are automatically $\gamma$-Lipschitz. In this direction we have the following result (cf. Exercise 3), which gives a first example of $\gamma$-Lipschitz continuous mappings.

*Example 14.3.* If $F$ has type 2, then every Lipschitz function $B: E\to\gamma(H,F)$ is $\gamma$-Lipschitz continuous and we have $\mathrm{Lip}(B)\leqslant\mathrm{Lip}_\gamma(B)\leqslant T_2^\gamma\mathrm{Lip}(B)$, where $T_2^\gamma$ is the Gaussian type 2 constant of $F$.

This result actually characterises the type 2 property; see the Notes at the end of the lecture. Further examples of $\gamma$-Lipschitz continuous mappings, relevant for applications to stochastic PDEs, will be given in the next lecture.

## 14.2 Pisier's property

Our next aim is to prove certain weighted bounds for stochastic convolutions. In order to keep the technicalities at a reasonable level we shall assume an additional geometric property on the underlying Banach space $E$, first studied by Pisier.

Let $(r_j')_{j=1}^\infty$ and $(r_k'')_{k=1}^\infty$ be Rademacher sequences on probability spaces $(\Omega',\mathscr{F}',\mathbb{P}')$ and $(\Omega'',\mathscr{F}'',\mathbb{P}'')$, and let $(r_{jk})_{j,k=1}^\infty$ be a doubly indexed Rademacher sequence on a probability space $(\Omega,\mathscr{F},\mathbb{P})$. In the next result, recall that $(r_j'r_k'')_{j,k=1}^\infty$ is not a Rademacher sequence (see Exercise 3.2).

**Proposition 14.4.** *For a Banach space $E$ the following assertions are equivalent:*

(1) *there exists a constant* $0 < C < \infty$ *such that for all finite sequences* $(a_{jk})_{j,k=1}^n$ *in* $\mathbb{R}$ *and* $(x_{jk})_{j,k=1}^n$ *in* $E$ *we have*

$$\mathbb{E}'\mathbb{E}''\Big\|\sum_{j,k=1}^n a_{jk}r_j' r_k'' x_{jk}\Big\|^2 \leqslant C^2\big(\max_{1\leqslant j,k\leqslant n}|a_{jk}|\big)\mathbb{E}'\mathbb{E}''\Big\|\sum_{j,k=1}^n r_j' r_k'' x_{jk}\Big\|^2;$$

(2) *there exists a constant* $0 < C < \infty$ *such that for all finite sequences* $(x_{jk})_{j,k=1}^n$ *in* $E$ *we have*

$$\frac{1}{C^2}\mathbb{E}\Big\|\sum_{j,k=1}^n r_{jk}x_{jk}\Big\|^2 \leqslant \mathbb{E}'\mathbb{E}''\Big\|\sum_{j,k=1}^n r_j' r_k'' x_{jk}\Big\|^2 \leqslant C^2\mathbb{E}\Big\|\sum_{j,k=1}^n r_{jk}x_{jk}\Big\|^2.$$

Condition (1) means that the analogue of the Kahane contraction principle holds for double Rademacher sums in $E$.

*Proof.* (1)⇒(2): By randomisation and Fubini's theorem, from (1) we obtain

$$\begin{aligned}
&\mathbb{E}\Big\|\sum_{m=1}^M\sum_{n=1}^N r_{mn}x_{mn}\Big\|^2 \\
&\quad = \mathbb{E}'\mathbb{E}''\mathbb{E}\Big\|\sum_{m=1}^M\sum_{n=1}^N r_{mn}r_m' r_n'' x_{mn}\Big\|^2 = \mathbb{E}\mathbb{E}'\mathbb{E}''\Big\|\sum_{m=1}^M\sum_{n=1}^N r_{mn}r_m' r_n'' x_{mn}\Big\|^2 \\
&\quad \leqslant C^2\mathbb{E}\mathbb{E}'\mathbb{E}''\Big\|\sum_{m=1}^M\sum_{n=1}^N r_m' r_n'' x_{mn}\Big\|^2 = C^2\mathbb{E}'\mathbb{E}''\Big\|\sum_{m=1}^M\sum_{n=1}^N r_m' r_n'' x_{mn}\Big\|^2.
\end{aligned}$$

This gives the left hand side inequality in (2).

To prove the right hand side inequality in (2) we fix numbers $\varepsilon_{mn} \in \{-1,1\}$ and use (1) to obtain

$$\begin{aligned}
&\mathbb{E}'\mathbb{E}''\Big\|\sum_{m=1}^M\sum_{n=1}^N r_m' r_n'' x_{mn}\Big\|^2 \\
&\quad = \mathbb{E}'\mathbb{E}''\Big\|\sum_{m=1}^M\sum_{n=1}^N \varepsilon_{mn}^2 r_m' r_n'' x_{mn}\Big\|^2 \leqslant C^2\mathbb{E}'\mathbb{E}''\Big\|\sum_{m=1}^M\sum_{n=1}^N \varepsilon_{mn}r_m' r_n'' x_{mn}\Big\|^2.
\end{aligned}$$

Taking $\varepsilon_{mn} = r_{mn}(\omega)$ and taking expectations,

$$\begin{aligned}
\mathbb{E}'\mathbb{E}''\Big\|\sum_{m=1}^M\sum_{n=1}^N r_m' r_n'' x_{mn}\Big\|^2 &\leqslant C^2\mathbb{E}\mathbb{E}'\mathbb{E}''\Big\|\sum_{m=1}^M\sum_{n=1}^N r_{mn}r_m' r_n'' x_{mn}\Big\|^2 \\
&= C^2\mathbb{E}'\mathbb{E}''\mathbb{E}\Big\|\sum_{m=1}^M\sum_{n=1}^N r_{mn}r_m' r_n'' x_{mn}\Big\|^2 = C^2\mathbb{E}\Big\|\sum_{m=1}^M\sum_{n=1}^N r_{mn}x_{mn}\Big\|^2.
\end{aligned}$$

(2)⇒(1): This implication follows from the Kahane contraction principle, which may be applied to the outer terms in (2). □

It can be shown that in (1) and (2), the role of Rademacher variables may be replaced by Gaussian variables without changing the class of spaces under consideration; this only affects the numerical value of the constants in the inequalities (a proof of the easy implication is contained in the proof of Proposition 14.7 below). Furthermore, in both formulations the exponent 2 may be replaced by an arbitrary $p \in [1, \infty)$. For Rademacher variables this was shown in the solution to Exercise 3.3; the proof for Gaussian variables is the same.

**Definition 14.5.** *A Banach space is said to have* Pisier's property *if it satisfies the equivalent conditions of the proposition.*

*Example 14.6.* If $(A, \mathscr{A}, \mu)$ is a $\sigma$-finite measure space, then for all $1 \leqslant p < \infty$ the space $L^p(A)$ has Pisier's property. More generally, if $E$ has Pisier's property, then $L^p(A; E)$ has Pisier's property.

In view of the remarks preceding the definition, the second assertion follows by switching to power $p$ and using Fubini's theorem. For the first assertion it then remains to be verified that $\mathbb{R}$ has Pisier's property. But this is the content of Exercise 3.3; the same argument shows that every Hilbert space has Pisier's property.

The next proposition connects Pisier's property with the theory of $\gamma$-radonifying operators.

**Proposition 14.7.** *Let $H$ be a Hilbert space. If $E$ has Pisier's property, then one has a canonical isomorphism of Banach spaces*

$$\gamma(L^2(0, T), \gamma(L^2(0, T; H), E)) \simeq \gamma(L^2((0, T)^2; H), E).$$

*Proof.* As in the proof of Theorem 3.12, from the central limit theorem we deduce that condition (2) of Proposition 14.4 implies its Gaussian counterpart

$$\frac{1}{C^2}\mathbb{E}\Big\| \sum_{j,k=1}^{n} \gamma_{jk} x_{jk} \Big\|^2 \leqslant \mathbb{E}'\mathbb{E}''\Big\| \sum_{j,k=1}^{n} \gamma'_j \gamma''_k x_{jk} \Big\|^2 \leqslant C^2 \mathbb{E}\Big\| \sum_{j,k=1}^{n} \gamma_{jk} x_{jk} \Big\|^2.$$

Let the sets $A_j$ be measurable and disjoint and also let the sets $B_j$ be measurable and disjoint, and let $h_1, \dots, h_n$ be orthonormal in $H$. Consider the step function

$$f = \sum_{j,k,l=1}^{n} 1_{A_j} \otimes ((1_{B_k} \otimes h_l) \otimes x_{jkl}) = \sum_{j,k,l=1}^{n} ((1_{A_j} \otimes 1_{B_k}) \otimes h_l) \otimes x_{jkl}.$$

The first sum is interpreted as an element of $\gamma(L^2(0, T), \gamma(L^2(0, T; H), E))$ and the second as an element of $\gamma(L^2((0, T)^2; H), E)$. For such $f$, the estimate

$$\frac{1}{C}\|f\|_{\gamma(L^2((0,T)^2;H),E)} \leqslant \|f\|_{\gamma(L^2(0,T),\gamma(L^2(0,T;H),E))} \leqslant C\|f\|_{\gamma(L^2((0,T)^2;H),E)}$$

is a reformulation of the above Gaussian estimate. The result follows from this by an approximation argument. $\square$

## 14.3 Stochastic convolutions

We shall now apply Proposition 14.7 to estimate stochastic convolutions.

Let $S : (0,T) \to \mathscr{L}(E,F)$ be strongly measurable in the sense that $Sx$ is strongly measurable for all $x \in E$. Let $W_H$ be an $H$-cylindrical Brownian motion, adapted to a filtration $(\mathscr{F}_t)_{t\in[0,T]}$. Given an adapted operator-valued process $\Phi : (0,T) \times \Omega \to \mathscr{L}(H,E)$, we introduce the notation

$$(S \diamond \Phi)(t) := \int_0^t S(t-s)\Phi(s)\, dW_H(s), \quad t \in [0,T],$$

provided these stochastic integrals exist.

**Lemma 14.8.** *Let $E$ and $F$ be Banach spaces, where $F$ is UMD and has Pisier's property, and let $S : (0,T) \to \mathscr{L}(E,F)$ be as above. Let $\Phi : (0,T) \times \Omega \to \mathscr{L}(H,E)$ be $H$-strongly measurable and adapted. Let $1 < p < \infty$ and fix $0 \leqslant \theta < \frac{1}{2}$. Suppose that:*

(1) *the set $\{t^\theta S(t) :\ t \in [0,T]\}$ is $\gamma$-bounded in $\mathscr{L}(E,F)$;*
(2) *the process $t \mapsto (T-t)^{-\theta}\Phi(t)$ belongs to $L^p(\Omega;\gamma(L^2(0,T;H),E))$.*

*Then the process $t \mapsto (T-t)^{-\theta}(S\diamond\Phi)(t)$ is well defined, $H$-strongly measurable and adapted, and defines an element of $L^p(\Omega;\gamma(L^2(0,T),F))$. Moreover,*

$$\begin{aligned}&\|t \mapsto (T-t)^{-\theta}(S\diamond\Phi)(t)\|_{L^p(\Omega;\gamma(L^2(0,T),F))}\\ &\qquad \leqslant CT^{\frac{1}{2}-\theta}\|t \mapsto (T-t)^{-\theta}\Phi(t)\|_{L^p(\Omega;\gamma(L^2(0,T;H),E))},\end{aligned}$$

*where $C$ is independent of $T$ and $\Phi$.*

*Proof.* Let us first note that $s \mapsto S(t-s)\Phi(s)$ is $H$-strongly measurable and adapted on $(0,t)$. Moreover, by Theorem 9.13 and the assumptions (1) and (2), this function defines an element of $L^p(\Omega;\gamma(L^2(0,t;H),F))$. Theorem 13.7 therefore shows that it is $L^p$-stochastically integrable on $(0,t)$ with respect to $W_H$. This shows that the process $S\diamond\Phi$ is well-defined. The proof that it is $H$-strongly measurable and adapted is a bit tedious and is left to the reader.

Let $\Delta := \{(t,s) \in (0,T)^2 :\ 0 < s < t < T\}$. We estimate

$$\begin{aligned}&\|t \mapsto (T-t)^{-\theta}(S\diamond\Phi)(t)\|_{L^p(\Omega;\gamma(L^2(0,T),F))}\\
&\overset{\text{(i)}}{\eqsim} \Big\|t \mapsto (T-t)^{-\theta}\int_0^t S(t-s)\Phi(s)\,dW_H(s)\Big\|_{\gamma(L^2(0,T),L^p(\Omega;F))}\\
&\overset{\text{(ii)}}{\eqsim} \big\|t \mapsto (T-t)^{-\theta}[s \mapsto 1_{(0,t)}(s)S(t-s)\Phi(s)]\big\|_{\gamma(L^2(0,T),L^p(\Omega;\gamma(L^2(0,T;H),F)))}\\
&\overset{\text{(iii)}}{\eqsim} \big\|t \mapsto (T-t)^{-\theta}[s \mapsto 1_{(0,t)}(s)S(t-s)\Phi(s)]\big\|_{\gamma(L^2(0,T),\gamma(L^2(0,T;H),L^p(\Omega;F)))}\\
&\overset{\text{(iv)}}{\eqsim} \big\|(t,s) \mapsto 1_\Delta(t,s)(T-t)^{-\theta}S(t-s)\Phi(s)\big\|_{\gamma(L^2((0,T)^2;H),L^p(\Omega;F)))}\\
&\overset{\text{(v)}}{\lesssim} \big\|(t,s) \mapsto 1_\Delta(t,s)(T-t)^{-\theta}(t-s)^{-\theta}\Phi(s)\big\|_{\gamma(L^2((0,T)^2;H),L^p(\Omega;E)))}.\end{aligned}$$

The justification of these steps is as follows: (i) follows from the $\gamma$-Fubini isomorphism of Theorem 5.22, (ii) combines Theorem 13.7 with the observation that each bounded operator $S$ from $E_1$ to $F_1$ canonically induces a bounded operator from $\gamma(L^2(0,T),E_1)$ to $\gamma(L^2(0,T);F_1)$, (iii) follows again from the $\gamma$-Fubini isomorphism, (iv) uses Pisier's property of the space $L^p(\Omega;F)$ (cf. Example 14.6) through Proposition 14.7, and (v) follows from the $\gamma$-boundedness assumption.

Consider the operator $P : L^2(0,T;H) \to L^2((0,T)^2;H)$ defined by

$$(Pf)(t,s) := 1_\Delta(t,s)(T-t)^{-\theta}(t-s)^{-\theta}(T-s)^{\theta} f(s).$$

This operator is bounded of norm $\|P\| \leqslant CT^{\frac12-\theta}$, since

$$\begin{aligned}
&\int_0^T \int_0^t (T-t)^{-2\theta}(t-s)^{-2\theta}(T-s)^{2\theta}\|f(s)\|^2\,ds\,dt \\
&\quad = \int_0^T (T-s)^{2\theta}\|f(s)\|^2\Big(\int_s^T (T-t)^{-2\theta}(t-s)^{-2\theta}\,dt\Big)\,ds \\
&\quad = \int_0^T (T-s)^{1-2\theta}\|f(s)\|^2\Big(\int_0^1 (1-r)^{-2\theta}r^{-2\theta}\,dr\Big)\,ds \\
&\quad \leqslant C^2T^{1-2\theta}\int_0^T \|f(s)\|^2\,ds,
\end{aligned}$$

where $C^2 := \int_0^1 (1-r)^{-2\theta}r^{-2\theta}\,dr$ depends only on $\theta$. Using the right ideal property of Proposition 5.11 it follows that

$$\begin{aligned}
&\big\|(t,s)\mapsto 1_\Delta(t,s)(T-t)^{-\theta}(t-s)^{-\theta}\Phi(s)\big\|_{\gamma(L^2((0,T)^2;H),L^p(\Omega;E))} \\
&\quad \leqslant CT^{\frac12-\theta}\big\|s\mapsto (T-s)^{-\theta}\Phi(s)\big\|_{\gamma(L^2(0,T;H),L^p(\Omega;E))} \\
&\quad \eqsim CT^{\frac12-\theta}\big\|s\mapsto (T-s)^{-\theta}\Phi(s)\big\|_{L^p(\Omega;\gamma(L^2(0,T;H),E))}. \qquad \square
\end{aligned}$$

For $\theta \geqslant 0$ and $1 \leqslant p < \infty$ we define the Banach space

$$V_\theta^p(\Omega;\gamma(L^2(0,T;H),E))$$

as the completion of the space of all adapted finite rank step processes $\Phi : (0,T)\times\Omega \to \mathscr{L}(H,E)$ with respect to the norm

$$\|\Phi\|_{V_\theta^p(\Omega;\gamma(L^2(0,T;H),E))} := \sup_{t\in(0,T]} \|s\mapsto (t-s)^{-\theta}\Phi(s)\|_{L^p(\Omega;\gamma(L^2(0,t;H),E))}.$$

We write $V_\theta^p(\Omega;\gamma(L^2(0,T),E))$ instead of $V_\theta^p(\Omega;\gamma(L^2(0,T;\mathbb{R}),E))$.

By applying Lemma 14.8 to the subintervals $(0,t)$ we obtain:

**Proposition 14.9.** *Let $E$, $F$, $S$, $\theta$ be as in Lemma 14.8. Then for all $1 < p < \infty$ the stochastic convolution $\Phi \mapsto S \diamond \Phi$ acts as a bounded linear operator from $V_\theta^p(\Omega;\gamma(L^2(0,T),E))$ to $V_\theta^p(\Omega;\gamma(L^2(0,T;H),F))$ of norm $\leqslant CT^{\frac12-\theta}$.*

For $\gamma$-Lipschitz continuous mappings $B : E \to \mathscr{L}(H,F)$ we have the following mapping property:

**Proposition 14.10.** *If $B : E \to \mathscr{L}(H,F)$ is $\gamma$-Lipschitz continuous, then for all $\theta \geqslant 0$ and $1 \leqslant p < \infty$ the map $B$ acts as a Lipschitz continuous mapping from $V_\theta^p(\Omega;\gamma(L^2(0,T;H),E))$ to $V_\theta^p(\Omega;\gamma(L^2(0,T),F))$ with Lipschitz constant $\leqslant \mathrm{Lip}_\gamma(B)$.*

*Proof.* For $t \in (0,T)$ let $\mu_{t,\theta}$ be the finite Borel measure on $(0,t)$ defined by

$$\mu_{t,\theta}(A) = \int_A (t-s)^{-2\theta}\,ds, \quad A \in \mathscr{B}(0,t).$$

The result follows from Proposition 14.1 and the observation that for an $H$-strongly measurable function $\Psi : (0,T) \to \mathscr{L}(H,E)$ the following assertions are equivalent:

(1) $s \mapsto (t-s)^{-\theta}\Psi(s)$ defines an element of $\gamma(L^2(0,t;H),E)$;
(2) $s \mapsto \Psi(s)$ defines an element of $\gamma(L^2((0,t),\mu_{t,\theta};H),E)$.

This equivalence is a consequence of the fact that the functions $h_1,\dots,h_n$ are orthonormal in $L^2((0,t),\mu_{t,\theta};H)$ if and only if the functions $s \mapsto (t-s)^{-\theta}h_1(s),\ \dots,\ s \mapsto (t-s)^{-\theta}h_n(s)$ are orthonormal in $L^2(0,t;H)$. $\square$

## 14.4 Existence and uniqueness

After these preparations we are ready to prove existence and uniqueness of solutions for the problem (SCP).

We fix an initial value $u_0 \in L^p(\Omega,\mathscr{F}_0;E)$ and consider the fixed point map $L_T$, initially defined for step functions $\phi : (0,T) \to E$ by

$$L_T(\phi) := Su_0 + S \diamond B(\phi),$$

where for brevity we write $(Su_0)(t) := S(t)u_0$.

The next result formulates a set of conditions ensuring that $L_T$ be well-defined on $V_\theta^p(\Omega;\gamma(L^2(0,T),E))$.

**Proposition 14.11.** *Let $E$ be a UMD space with Pisier's property and let $1 < p < \infty$. Let $A$ be the generator of a $C_0$-semigroup $S$ on $E$ such that $\{t^\theta S(t) : \ t \in (0,T)\}$ is $\gamma$-bounded for some $0 \leqslant \theta < \frac{1}{2}$, and let $B : E \to \gamma(H,E)$ be $\gamma$-Lipschitz continuous. Then the mapping $L_T$ is well-defined and Lipschitz continuous on $V_\theta^p(\Omega;\gamma(L^2(0,T),E))$ and there exists a constant $C \geqslant 0$, independent of $T$ and $u_0$, such that:*

(1) *for all $\phi \in V_\theta^p(\Omega;\gamma(L^2(0,T),E))$,*

$$\|L_T(\phi)\|_{V_\theta^p(\Omega;\gamma(L^2(0,T),E))} \leqslant CT^{\frac{1}{2}-\theta}\big(T^{1-2\theta}+\|u_0\|_p+\|\phi\|_{V_\theta^p(\Omega;\gamma(L^2(0,T),E))}\big);$$

(2) *for all* $\phi_1, \phi_2 \in V_\theta^p(\Omega;\gamma(L^2(0,T),E))$,

$$\|L_T(\phi_1) - L_T(\phi_2)\|_{V_\theta^p(\Omega;\gamma(L^2(0,T),E))} \leqslant CT^{\frac{1}{2}-\theta}\|\phi_1 - \phi_2\|_{V_\theta^p(\Omega;\gamma(L^2(0,T),E))}.$$

*Proof.* We begin by estimating the initial value part. By Lemma 10.17 and Theorem 9.13, for all $t \in (0,T]$ the following estimate holds for almost all $\omega \in \Omega$:

$$\begin{aligned}\|s \mapsto (t-s)^{-\theta} S(s)u_0(\omega)\|_{\gamma(L^2(0,t),E)} &\lesssim t^\theta \|s \mapsto s^{-\theta}(t-s)^{-\theta}u_0(\omega)\|_{\gamma(L^2(0,t),E)} \\ &= t^{\frac{1}{2}-\theta}\|s \mapsto s^{-\theta}(t-s)^{-\theta}\|_{L^2(0,t)}\|u_0(\omega)\| \\ &\eqsim t^{\frac{1}{2}-\theta}\|u_0(\omega)\|,\end{aligned}$$

with a constant independent of $u_0$ and $t \in (0,T)$. In the third line, the equality follows Exercise 5.3. Hence,

$$\begin{aligned}&\|Su_0\|_{V_\theta^p(\Omega;\gamma(L^2(0,T),E))} \\ &\quad = \sup_{t\in(0,T]} \|s \mapsto (t-s)^{-\theta}S(s)u_0\|_{L^p(\Omega;\gamma(L^2(0,t;H),E))} \lesssim T^{\frac{1}{2}-\theta}\|u_0\|_p.\end{aligned} \tag{14.2}$$

Fix adapted step processes $\phi_1, \phi_2 : (0,T) \times \Omega \to E$. If $B$ is $\gamma$-Lipschitz, Propositions 14.9 and 14.10 show that $B(\phi_k) \in V_\theta^p(\Omega;\gamma(L^2(0,T;H),E))$, $S \diamond B(\phi_k) \in V_\theta^p(\Omega;\gamma(L^2(0,T),E))$, $k = 1,2$, and

$$\begin{aligned}\|S \diamond B(\phi_1) - S \diamond B(\phi_2))\|_{V_\theta^p(\Omega;\gamma(L^2(0,T),E))} &\leqslant CT^{\frac{1}{2}-\theta}\|B(\phi_1) - B(\phi_2)\|_{V_\theta^p(\Omega;\gamma(L^2(0,T;H),E))} \\ &\leqslant CT^{\frac{1}{2}-\theta}\mathrm{Lip}_\gamma(B)\|\phi_1 - \phi_2\|_{V_\theta^p(\Omega;\gamma(L^2(0,T),E))}.\end{aligned}$$

It follows from these estimates that $L_T$ has a unique extension to a Lipschitz continuous mapping on $V_\theta^p(\Omega;\gamma(L^2(0,T),E))$ which satisfies the estimate of (2). The estimate (1) follows from the identity $L_T(\phi) = L_T(0) + (L_T(\phi) - L_T(0))$ and (2), using that from (14.2) and Proposition 14.9 we obtain

$$\begin{aligned}\|L_T(0)\|_{V_\theta^p(\Omega;\gamma(L^2(0,T),E))} &\lesssim T^{\frac{1}{2}-\theta}\big(\|u_0\|_p + \|1_{(0,T)} \otimes B(0)\|_{V_\theta^p(\Omega;\gamma(L^2(0,T),E))}\big) \\ &\leqslant T^{\frac{1}{2}-\theta}\big(\|u_0\|_p + T^{1-2\theta}\|B(0)\|_{\gamma(H,E)}\big).\end{aligned}$$

□

After these preparations we are ready to formulate our main result for existence and uniqueness of mild solutions for the stochastic evolution equation (SCP). We denote by $S$ the $C_0$-semigroup generated by $A$.

**Definition 14.12.** *Let* $\theta \geqslant 0$ *and* $1 \leqslant p < \infty$. *A strongly measurable and adapted process* $U : [0,T] \times \Omega \to E$ *is called a* mild $V_\theta^p$-solution *of the problem* (SCP) *if it belongs to* $V_\theta^p(\Omega;\gamma(L^2(0,T),E))$ *and for all* $t \in [0,T]$ *the following identity holds almost surely:*

$$U(t) = S(t)u_0 + (S \diamond B(U))(t).$$

*A mild $V_0^p$-solution is called a* mild $L^p$-solution.

This definition is motivated by the formula $U(t) = S(t)u_0 + (S \diamond B)(t)$ for the unique weak solution of the problem $dU(t) = AU(t)\,dt + B\,dW_H(t)$ which was studied in Lectures 8–10 (and corresponds to the special case $B(x) \equiv B$).

**Theorem 14.13 (Existence and uniqueness).** *Let $E$ be a UMD space with Pisier's property and let $1 < p < \infty$. Suppose that $A$ is the generator of a $C_0$-semigroup $S$ on $E$ such that $\{t^\theta S(t) : t \in [0,T]\}$ is $\gamma$-bounded for some $0 \leqslant \theta < \frac{1}{2}$, let $B : E \to \gamma(H,E)$ be $\gamma$-Lipschitz continuous, and let $u_0 \in L^p(\Omega, \mathscr{F}_0; E)$. Then there exists a unique mild $V_\theta^p$-solution $U$ of* (SCP). *Moreover, there exists a constant $C_T \geqslant 0$, independent of $u_0$, such that*

$$\|U\|_{V_\theta^p(\Omega;\gamma(L^2(0,T;H),E))} \leqslant C_T(1 + \|u_0\|_p).$$

Here, uniqueness is understood in the sense of $V_\theta^p(\Omega;\gamma(L^2(0,T),E))$. By strong measurability, any two solutions representing the same element in this space are versions of each other.

*Proof.* By Proposition 14.11 we can find $0 < T_0 \leqslant T$, independent of $u_0$, such that

$$\|L_{T_0}(\phi_1) - L_{T_0}(\phi_2)\|_{V_\theta^p(\Omega;\gamma(L^2(0,T_0),E))} \leqslant \frac{1}{2}\|\phi_1 - \phi_2\|_{V_\theta^p(\Omega;\gamma(L^2(0,T_0),E))} \quad (14.3)$$

for all $\phi_1, \phi_2 \in V_\theta^p(\Omega;\gamma(L^2(0,T_0),E))$ and

$$\|L_{T_0}(\phi)\|_{V_\theta^p(\Omega;\gamma(L^2(0,T_0),E))} \leqslant \frac{1}{2}\big(1 + \|u_0\|_p + \|\phi\|_{V_\theta^p(\Omega;\gamma(L^2(0,T_0),E))}\big) \quad (14.4)$$

for $\phi \in V_\theta^p(\Omega;\gamma(L^2(0,T_0),E))$. By (14.3) and the Banach fixed point theorem, $L_{T_0}$ has a unique fixed point $\widetilde{U} \in V_\theta^p(\Omega;\gamma(L^2(0,T_0),E))$. Define the strongly measurable adapted process $U : [0,T_0] \times \Omega \to E$ by

$$U(t) := S(t)u_0 + (S \diamond B(\widetilde{U}))(t).$$

Then $U$ is a mild $V_\theta^p$-solution, and clearly we have $U = \widetilde{U}$ as elements of $V_\theta^p(\Omega;\gamma(L^2(0,T_0),E))$. Uniqueness in $V_\theta^p(\Omega;\gamma(L^2(0,T_0),E))$ follows from the uniqueness of the fixed point in that space. Noting that $\widetilde{U} = L_{T_0}(\widetilde{U})$, the estimate (14.4) implies the final estimate on the interval $[0,T_0]$.

Via a standard induction argument we now construct a mild solution on each of the intervals $[T_0, 2T_0], \ldots, [(n-1)T_0, nT_0], [nT_0, T]$, where $n$ is an appropriate integer. This results in a mild solution $U$ on $[0,T]$ of (SCP) with the properties as stated in the theorem. Uniqueness on $[0,T]$ follows by induction from the uniqueness on each of the subintervals. We leave the somewhat tedious details as an exercise (see Exercise 5). $\square$

Let us have a closer look at this theorem for the special case where $E$ is a Hilbert space. Then $E$ is a UMD space with Pisier's property, the family $\{S(t) :\ t \in [0,T]\}$ is $\gamma$-bounded (since in Hilbert spaces, uniformly bounded families are $\gamma$-bounded), and every Lipschitz continuous function $B : E \to \gamma(H,E) = \mathscr{L}_2(H,E)$ is $\gamma$-Lipschitz continuous (since Hilbert spaces have type 2, cf. Example 14.3); recall that $\mathscr{L}_2(H,E)$ denotes the space of all Hilbert-Schmidt operators from $H$ to $E$.

**Corollary 14.14 (Hilbert space case).** *Let $E$ be a Hilbert space and let $1 < p < \infty$. Suppose that $A$ is the generator of a $C_0$-semigroup on $E$, let $B : E \to \mathscr{L}_2(H,E)$ be Lipschitz continuous, and let $u_0 \in L^p(\Omega,\mathscr{F}_0;E)$. Then there exists a unique mild $L^p$-solution of* (SCP). *Moreover, there exists a constant $C_T \geqslant 0$, independent of $u_0$, such that*

$$\|U\|_{L^p(\Omega;\gamma(L^2(0,T),E))} \leqslant C_T(1 + \|u_0\|_p).$$

## 14.5 Space-time regularity

To motivate our approach we return to the proof Theorem 10.19, where space-time Hölder regularity of solutions was proved under the assumption that the semigroup $S$ generated by $A$ is analytic. The crucial ingredient was the $\gamma$-boundedness of the family $\{t^\theta S(t) :\ t \in (0,T)\}$ in $\mathscr{L}(E,E_\alpha)$ for $0 \leqslant \alpha < \theta < \frac{1}{2}$. Recall that the spaces $E_\alpha$ have been defined in Lecture 10 as the fractional domain spaces $\mathscr{D}((w-A)^\alpha)$.

As the next proposition shows, for processes in $V_\theta^p(\Omega;\gamma(L^2(0,T),E))$ the proof of Theorem 10.19 can be repeated.

**Proposition 14.15.** *Let $A$ be the generator of an analytic $C_0$-semigroup $S$ on a UMD space $E$. Suppose that $2 < p < \infty$ and $\frac{1}{p} < \theta < \frac{1}{2}$ and let $\alpha \geqslant 0$ and $\beta \geqslant 0$ satisfy $0 \leqslant \alpha + \beta < \theta - \frac{1}{p}$. If $\Phi \in V_\theta^p(\Omega;\gamma(L^2(0,t;H),E))$, then $S \diamond \Phi$ has a version with trajectories in $C^\beta([0,T];E_\alpha)$ and*

$$\|S \diamond \Phi\|_{L^p(\Omega;C^\beta([0,T];E_\alpha))} \leqslant C_T\|\Phi\|_{V_\theta^p(\Omega;\gamma(L^2(0,t;H),E))},$$

*where the constant $C_T \geqslant 0$ is independent of $\Phi$.*

*Proof.* First note that the $\gamma$-boundedness of $\{t^\theta S(t) :\ t \in (0,T)\}$ in $\mathscr{L}(E,E_\alpha)$ implies that $(S \diamond \Phi)(t) \in L^p(\Omega;E_\alpha)$ for all $t \in [0,T]$. Fix $0 < s < t \leqslant T$ and write

$$(\mathbb{E}\|S \diamond \Phi(t) - S \diamond \Phi(s)\|_{E_\alpha}^p)^{\frac{1}{p}} \leqslant R_1 + R_2,$$

where

$$R_1 = \Big(\mathbb{E}\Big\|\int_0^s S(t-r) - S(s-r)\Phi(r)\,dW_H(r)\Big\|_{E_\alpha}^p\Big)^{\frac{1}{p}},$$

$$R_2 = \Big(\mathbb{E}\Big\|\int_s^t S(t-r)\Phi(r)\,dW_H(r)\Big\|_{E_\alpha}^p\Big)^{\frac{1}{p}}.$$

Let $\beta' > \beta$ satisfy $\alpha + \beta' < \theta - \frac{1}{p}$ and put $\delta := \beta' + \frac{1}{p}$. Then $\alpha + \delta < \theta$ and $\beta < \delta - \frac{1}{p}$. By Lemma 10.17, and Theorems 9.13 and 13.7, for large $w$ we have

$$\begin{aligned}
R_1^p &\lesssim \mathbb{E}\|r \mapsto S(s-r)(S(t-s)-I)\Phi(r)\|^p_{\gamma(L^2(0,s;H),E_\alpha)} \\
&\eqsim \mathbb{E}\|r \mapsto S(s-r)(S(t-s)-I)(w-A)^{-\delta}\Phi(r)\|^p_{\gamma(L^2(0,s;H),E_{\alpha+\delta})} \\
&\lesssim T^{(\theta-\alpha-\delta)p}\mathbb{E}\|r \mapsto (s-r)^{-\theta}(S(t-s)-I)(w-A)^{-\delta}\Phi(r)\|^p_{\gamma(L^2(0,s;H),E)} \\
&\lesssim T^{(\theta-\alpha-\delta)p}(t-s)^{\delta p}\mathbb{E}\|r \mapsto (s-r)^{-\theta}\Phi(r)\|^p_{\gamma(L^2(0,s;H),E)} \\
&\lesssim T^{(\theta-\alpha-\delta)p}(t-s)^{\delta p}\|\Phi\|^p_{V^p_\theta(\Omega;\gamma(L^2(0,T;H),E))}.
\end{aligned}$$

In the second last estimate we used that $\|(S(t-s)-I)(w-A)^{-\delta}\| \lesssim |t-s|^\delta$ by the analyticity of $S$ (see Lemma 10.15). Similarly,

$$\begin{aligned}
R_2^p &\lesssim \mathbb{E}\|r \mapsto S(t-r)\Phi(r)\|^p_{\gamma(L^2(s,t;H),E_\alpha)} \\
&\lesssim T^{(\theta-\delta-\alpha)p}\mathbb{E}\|r \mapsto (t-r)^{-\theta+\delta}\Phi(r)\|^p_{\gamma(L^2(s,t;H),E)} \\
&\lesssim T^{(\theta-\delta-\alpha)p}(t-s)^{\delta p}\mathbb{E}\|r \mapsto (t-r)^{-\theta}\Phi(r)\|^p_{\gamma(L^2(s,t;H),E)} \\
&\lesssim T^{(\theta-\delta-\alpha)p}(t-s)^{\delta p}\|\Phi\|^p_{V^p_\theta(\Omega;\gamma(L^2(0,T;H),E))},
\end{aligned}$$

where the second last inequality follows by covariance domination. Combining these estimates with Kolmogorov's theorem (Theorem 6.9) and using that $\beta < (\delta p - 1)/p = \delta - \frac{1}{p}$, we obtain a version of $S \diamond \Phi$ which is $\beta$-Hölder continuous in $E_\alpha$. □

We are now ready to formulate our main regularity result for the mild solutions of problem (SCP).

**Theorem 14.16 (Hölder Regularity).** *Let $A$ be the generator of an analytic $C_0$-semigroup $S$ on a UMD space $E$ with Pisier's property. Suppose that $B : E \to \mathscr{L}(H,E)$ is $\gamma$-Lipschitz continuous and let $u_0 \in L^p(\Omega, \mathscr{F}_0; E)$. For all $\alpha, \beta, \theta \geqslant 0$ satisfying $\alpha + \beta < \theta < \frac{1}{2}$ and all $1 < p < \infty$, the unique mild $V^p_\theta$-solution $U$ of the problem* (SCP) *has a version for which $U - Su_0$ has trajectories in $C^\beta([0,T];E_\alpha)$.*

Note that if $u_0$ is sufficiently regular, this result implies that $U$ itself has a version with trajectories in $C^\beta([0,T];E_\alpha)$.

*Proof.* The existence of a unique mild $V^p_\theta$-solution follows from Theorem 14.13; the $\gamma$-boundedness assumption holds by the analyticity of $S$.

If $U$ is a mild $V^p_\theta$-solution and $\widetilde{U}$ is mild $V^q_\theta$-solution, where $1 < p \leqslant q < \infty$, then $\widetilde{U}$ is also a mild $V^p_\theta$-solution. Hence by uniqueness, $U$ and $\widetilde{U}$ are equal as elements of $V^p_\theta(\Omega;\gamma(L^2(0,T),E))$, and by strong measurability $U$ and $\widetilde{U}$ are versions of each other. Therefore it suffices to consider the case where $2 < p < \infty$ satisfies $\alpha + \beta < \theta - \frac{1}{p} < \frac{1}{2} - \frac{1}{p}$.

By Proposition 14.9, $S \diamond B(U)$ belongs to $V_\theta^p(\Omega;\gamma(L^2(0,t;H),E_\alpha))$, where $U$ is the mild $V_\theta^p$-solution $U$ of (SCP). Hence, by Proposition 14.15, $U - Su_0 = S \diamond B(U)$ has a version with trajectories in $C^\beta([0,T];E_\alpha)$ and

$$\begin{aligned}\mathbb{E}\|S \diamond B(U)\|^p_{C^\beta([0,T];E_\alpha)} &\leqslant C^p \|B(U)\|^p_{V_\theta^p(\Omega;\gamma(L^2(0,T;H),E))} \\ &\leqslant C^p(1 + \|U\|_{V_\theta^p(\Omega;\gamma(L^2(0,T),E))})^p \leqslant C^p(1+\|u_0\|_p)^p,\end{aligned}$$

where the last of these estimates follows from Theorem 14.13. $\square$

## 14.6 Exercises

1. Provide the details of the central limit argument in Proposition 14.7.
2. Show that if $E$ is a Banach space with the property that the mapping
$$\sum_{j,k=1}^n 1_{A_j}(1_{B_k} \otimes x_{jk}) \mapsto \sum_{j,k=1}^n (1_{A_j} 1_{B_k}) \otimes x_{jk}$$
(with notations as in Proposition 14.7) induces an isomorphism
$$\gamma(L^2(0,T), \gamma(L^2(0,T),E)) \simeq \gamma(L^2((0,T)^2),E),$$
then $E$ has Pisier's property (in the formulation using Gaussian random variables; as has been noted without proof, this formulation is equivalent to the one with Rademachers given in the text). This gives a converse to Proposition 14.7.
3. Let $H$ be a Hilbert space, $E$ and $F$ Banach spaces, and assume that $E$ has cotype 2 and $F$ has type 2. Show that every Lipschitz continuous function $B: E \to \gamma(H,F)$ is $\gamma$-Lipschitz continuous with
$$\mathrm{Lip}(B) \leqslant \mathrm{Lip}_\gamma(B) \leqslant C_2(E)T_2(F)\mathrm{Lip}(B),$$
where $C_2(E)$ and $T_2(F)$ denote the Gaussian cotype 2 constant of $E$ and the type 2 constant of $F$, respectively.
*Hint:* Use the results of Exercise 5.4.
4. Frequently, uniqueness proofs are based on Gronwall's inequality. The purpose of this exercise is to show that the '$\gamma$-Gronwall inequality' fails in spaces without type 2.
   a) Show that if $E$ is a Banach space without type 2, then there exist step functions $\phi_n : (\frac{1}{n+1}, \frac{1}{n}) \to E$ such that
$$\|\phi_n\|_{L^\infty(\frac{1}{n+1},\frac{1}{n};E)} \leqslant 1, \quad \inf_{n\geqslant 1} \|\phi_n\|_{\gamma(L^2(\frac{1}{n+1},\frac{1}{n}),E)} > 0.$$
   b) Prove that the following assertions are equivalent:

(i) the space $E$ has type 2;

(ii) whenever $\phi : (0,1) \to E$ is a strongly measurable function representing an element of $\gamma(L^2(0,1),E)$ and there exists a constant $C = C_\phi \geqslant 0$ such that

$$\|\phi(t)\| \leqslant C\|\phi\|_{\gamma(L^2(0,t),E)} \text{ for almost all } t \in (0,1),$$

we have $\phi = 0$ almost everywhere on $(0,1)$.

*Hint:* In one direction, consider the function $\phi(t) := \frac{1}{n^2}\phi_n(t)$ for $t \in (t_{n+1}, t_n]$, where $\phi_n$ is as in a). In the other direction, use Gronwall's inequality.

5. Provide the details of the induction argument that was used at the end of the proof of Theorem 14.13.

**Notes.** The material of Section 14.1 and 14.3 is based on the paper [83]. Exercise 3 is a variation on a result of that paper. In [80], the following converse is proved: if every Lipschitz function $B : E \to F$ is $\gamma$-Lipschitz, then $E$ has cotype 2 and $F$ has type 2.

Pisier's property was introduced, under the name 'property $(\alpha)$', by Pisier [92] who proved that a Banach lattice has this property if and only if it has finite cotype. Proposition 14.4 and the equivalence with its Gaussian formulation belong to mathematical folklore. It should be noted that the UMD property and Pisier's property are unrelated: the Schatten classes $C^p$ have the UMD property for $1 < p < \infty$ but fail Pisier's property unless $p = 2$, whereas $L^1$-spaces have Pisier's property but fail the UMD property unless they are finite-dimensional.

Proposition 14.7 is a special case of the more general statement that if $H_1$ and $H_2$ are Hilbert spaces and $E$ is a Banach space with Pisier's property, then

$$\gamma(H_1, \gamma(H_2, E)) \simeq \gamma(H_1 \widehat{\otimes} H_2, E)$$

isomorphically, where $H_1 \widehat{\otimes} H_2$ is the Hilbert space tensor product of $H_1$ and $H_2$. Exercise 2 can be formulated similarly. Both results are due to Kalton and Weis [58].

The use of Pisier's property can be avoided in Lemma 14.8 and all results depending on it, but it would take a full lecture to explain all the details. The interested reader is referred to [83]. Previous results along these lines for Hilbert spaces can be found in Da Prato and Zabczyk [27]; they were extended to martingale type 2 spaces by Brzeźniak [14]. In this context it should be noted that if $S$ is a $C_0$-*contraction* semigroup on a Hilbert space $E$, then by a result of Kotelenez [60] and Tubaro [104] the convolution process

$$t \mapsto \int_0^t S(t-s)\Phi(s)\, dW_H(s)$$

has a continuous version for all adapted and $H$-strongly measurable $\Phi : (0,T) \times \Omega \to \mathscr{L}_2(H,E)$; see also Da Prato and Zabczyk [27, Theorem

6.10]. As a result, in this situation the solution of Theorem 14.14 has a continuous version.

The results of Sections 14.4 and 14.5 are based on the paper [83]. The main results, Theorem 14.13 and 14.16, are variations of results in that paper and can be extended to semilinear parabolic equations with time-dependent coefficients of the form

$$\begin{cases} dU(t) = (AU(t) + F(t, U(t)))\,dt + B(t, U(t))\,dW_H(t), \\ U(0) = u_0. \end{cases}$$

Sufficient conditions for a mild solution to be a weak solution (which is defined in analogy to Lecture 8) and vice versa are given by Da Prato and Zabczyk [27] and Veraar [106].

# 15

# Applications to stochastic PDE

In this final lecture we present some applications of the theory developed in this course to stochastic partial differential equations. We concentrate on two specific examples: the wave equation and the heat equation.

## 15.1 Space-time white noise

It has been mentioned already in Lecture 6 that for $H = L^2(D)$, where $D$ is a domain in $\mathbb{R}^d$, $H$-cylindrical Brownian motions can be used to model space-time white noise on $D$. We begin by making this idea more precise.

**Definition 15.1.** *Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space and denote by $\mathscr{A}_0$ the collection of all $B \in \mathscr{A}$ such that $\mu(B) < \infty$. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. A* white noise *on $(A, \mathscr{A}, \mu)$ is a mapping $w : \mathscr{A}_0 \to L^2(\Omega)$ such that:*

(i) *each $w(B)$ is centred Gaussian with*

$$\mathbb{E}(w(B))^2 = \mu(B);$$

(ii) *if $B_1 \cap \cdot \cap B_N = \varnothing$, then $w(B_1), \dots, w(B_N)$ are independent and*

$$w\Big(\bigcup_{n=1}^{N} B_n\Big) = \sum_{n=1}^{N} w(B_n).$$

It follows from the general theory of Gaussian processes that such mappings always exist. We shall not go into the details of this, since in all applications the white noise is assumed to be given.

**Definition 15.2.** *A white noise $w$ on $[0,T] \times D$, where $D$ is a domain in $\mathbb{R}^d$, will be called a* space-time white noise *on $D$.*

Canonically associated with such $w$ is an $L^2(D)$-cylindrical Brownian motion $W$, defined by

$$W(t)1_B := w([0,t] \times B), \quad B \in \mathscr{B}_0(D);$$

this definition is extended to simple functions by linearity. To see that $W$ is indeed an $L^2(D)$-cylindrical Brownian motion note that for disjoint $B_1, \dots, B_N \in \mathscr{B}_0(D)$ and real numbers $c_1, \dots, c_N$ we have, by (i) and (ii),

$$\begin{aligned}\mathbb{E}\Big(W(t)\sum_{n=1}^N c_n 1_{B_n}\Big)^2 &= \sum_{n=1}^N c_n^2 \mathbb{E}(w([0,t] \times B_n))^2 \\ &= \sum_{n=1}^N c_n^2 t|B_n| = t\Big\|\sum_{n=1}^N c_n 1_{B_n}\Big\|_{L^2(D)}^2.\end{aligned}$$

## 15.2 The stochastic wave equation

In this section we study the stochastic wave equation with Dirichlet boundary conditions, driven by multiplicative space-time white noise:

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2}(t,\xi) = \Delta u(t,\xi) + B\Big(u(t,\xi), \dfrac{\partial u}{\partial t}(t,\xi)\Big)\dfrac{\partial w}{\partial t}(t,\xi), & \xi \in D, \ t \in [0,T], \\ u(t,\xi) = 0, & \xi \in \partial D, \ t \in [0,T], \\ u(0,\xi) = u_0(\xi), & \xi \in D, \\ \dfrac{\partial u}{\partial t}(0,\xi) = v_0(\xi), & \xi \in D. \end{cases} \tag{WE}$$

Here $w$ is a space-time white noise on a bounded domain $D$ in $\mathbb{R}^d$ with smooth boundary $\partial D$.

In order to keep the technicalities at a minimum we discuss two special cases in detail: the case where the operator-valued function $B$ is of rank one, which is equivalent to the formulation (WE1) below, and the case where $D$ is the unit interval in $\mathbb{R}$ and $B = I$.

### 15.2.1 Rank one multiplicative noise

We begin with the following special case of (WE):

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2}(t,\xi) = \Delta u(t,\xi) + b\Big(u(t,\xi), \dfrac{\partial u}{\partial t}(t,\xi)\Big)\dfrac{\partial W}{\partial t}(t,\xi), & \xi \in D, \ t \in [0,T], \\ u(t,\xi) = 0, & \xi \in \partial D, \ t \in [0,T], \\ u(0,\xi) = u_0(\xi), & \xi \in D, \\ \dfrac{\partial u}{\partial t}(0,\xi) = v_0(\xi), & \xi \in D, \end{cases} \tag{WE1}$$

where $W$ is a standard Brownian motion. We assume that the diffusion term $b : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfies the growth condition

$$|b(\xi_1, \xi_2)|^2 \leqslant C_1^2(|\xi_1|^2 + |\xi_2|^2)$$

and the Lipschitz condition

$$|b(\xi_1, \xi_2) - b(\eta_1, \eta_2)|^2 \leqslant C_2^2(|\xi_1 - \eta_1|^2 + |\xi_2 - \eta_2|^2).$$

The initial values $u_0$ and $v_0$ are taken in $W^{1,2}(D)$ and $L^2(D)$, respectively.

Writing the first equation as a system of two first order equations,

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,\xi) = v(t,\xi), \\ \dfrac{\partial v}{\partial t}(t,\xi) = \Delta u(t,\xi) + b(u(t,\xi), v(t,\xi))\dfrac{\partial W}{\partial t}(t,\xi), \end{cases} \qquad \xi \in D,\ t \in [0,T], \tag{15.1}$$

we reformulate the problem (WE) as a first order stochastic evolution equation as follows. Let $\Delta$ denote the Dirichlet Laplacian on $L^2(D)$ with domain $\mathscr{D}(\Delta) = W^{2,2}(D) \cap W_0^{1,2}(D)$; see Examples 7.21. On the Hilbert space

$$\mathscr{H} := \mathscr{D}((-\Delta)^{1/2}) \times L^2(D) = W^{1,2}(D) \times L^2(D)$$

we define the operator

$$A := \begin{bmatrix} 0 & I \\ \Delta & 0 \end{bmatrix}$$

with domain $\mathscr{D}(A) := \mathscr{D}(\Delta) \times \mathscr{D}((-\Delta)^{1/2}) = (W^{2,2}(D) \cap W_0^{1,2}(D)) \times W^{1,2}(D)$. As in Example 7.22, this operator is the generator of a bounded $C_0$-group on $\mathscr{H}$, and we may reformulate the problem (15.1) as an abstract stochastic evolution equation of the form

$$\begin{cases} dU(t) = AU(t)\,dt + B(U(t))\,dW(t), \\ U(0) = U_0, \end{cases} \tag{15.2}$$

where $W$ is a Brownian motion and the function $B : \mathscr{H} \to \mathscr{H}$ is the *Nemytskii map* associated with $b$,

$$B\begin{bmatrix} f \\ g \end{bmatrix} := \begin{bmatrix} 0 \\ b(f,g) \end{bmatrix}, \qquad \begin{bmatrix} f \\ g \end{bmatrix} \in \mathscr{H},$$

and $U_0 := \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \in \mathscr{H}$.

**Proposition 15.3.** *Under the above assumptions on $b$, the Nemytskii map $B : \mathscr{H} \to \mathscr{H}$ is well defined, Lipschitz continuous with* $\mathrm{Lip}(B) \leqslant \mathrm{Lip}(b)$, *and of linear growth.*

*Proof.* For all $(f,g) \in \mathscr{H}$ we have

$$\begin{aligned}\|B(f,g)\|_{\mathscr{H}}^2 &= \int_D |b(f(\xi),g(\xi))|^2\,d\xi \\ &\lesssim \int_D |f(\xi)|^2 + |g(\xi)|^2\,d\xi \lesssim \|f\|_2^2 + \|g\|_2^2 \leqslant \|(f,g)\|_{\mathscr{H}}^2.\end{aligned}$$

A similar estimate gives that $B$ is Lipschitz continuous from $\mathscr{H}$ to $\mathscr{H}$ with $\mathrm{Lip}(B) \leqslant \mathrm{Lip}(b)$. ☐

We say that a measurable adapted process $u : [0,T] \times \Omega \times D \to \mathbb{R}$ is a *mild $L^p$-solution* of (WE) if $U(t,\omega) := \begin{bmatrix} u(t,\omega,\cdot) \\ \frac{\partial u}{\partial t}u(t,\omega,\cdot)\end{bmatrix}$ belongs to $\mathscr{H}$ for all $(t,\omega) \in [0,T] \times \Omega$ and the resulting process $U : [0,T] \times \Omega \to \mathscr{H}$ is a mild $L^p$-solution of the problem (15.2).

**Theorem 15.4.** *Under the above assumptions, for all $1 < p < \infty$ the problem* (WE) *admits a unique mild $L^p$-solution.*

Here, uniqueness is understood in the sense of $L^p(\Omega;\gamma(L^2(0,T),\mathscr{H}))$.

*Proof.* By Proposition 15.3, the Nemytskii operator $B$ associated with $b$ is Lipschitz continuous. Moreover, as we have seen in Example 7.22, the operator $A$ is the generator of a $C_0$-group on $\mathscr{H}$. We have thus checked all assumptions of Corollary 14.14 (with $H = \mathbb{R}$ and $E = \mathscr{H}$) and conclude that for all $1 < p < \infty$ the problem (WE) admits a unique mild $L^p$-solution. ☐

### 15.2.2 Additive space-time white noise

Our next example concerns the stochastic wave equation with additive space-time white noise on the unit interval $(0,1)$ in $\mathbb{R}$:

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2}(t,\xi) = \Delta u(t,\xi) + \dfrac{\partial w}{\partial t}(t,\xi), & \xi \in (0,1), \quad t \in [0,T], \\ u(t,0) = u(t,1) = 0, & \qquad\qquad\quad t \in [0,T], \\ u(0,\xi) = u_0(\xi), & \xi \in (0,1), \\ \dfrac{\partial u}{\partial t}(0,\xi) = v_0(\xi), & \xi \in (0,1). \end{cases} \tag{WE2}$$

Here $w$ is a space-time white noise on $(0,1)$. We model this problem as an abstract stochastic evolution equation on the Hilbert space $\mathscr{H} = W^{1,2}(0,1) \times L^2(0,1)$ as

$$\begin{cases} dU(t) = AU(t)\,dt + d\begin{bmatrix} 0 \\ W_{L^2}(t)\end{bmatrix}, \\ U(0) = U_0, \end{cases} \tag{15.3}$$

where as before $A = \begin{bmatrix} 0 & I \\ \Delta & 0 \end{bmatrix}$, with $\Delta$ the Dirichlet Laplacian on $L^2(0,1)$, and $W_{L^2}$ is the $L^2(0,1)$-cylindrical Brownian motion canonically associated with $w$; see Section 15.1.

To analyse the problem (15.3) we use the functional calculus for self-adjoint operators. Using this calculus it can be checked that the $C_0$-group $S$ generated by $A$ is of the form

$$S(t) = \begin{bmatrix} \cos(t(-\Delta)^{1/2}) & (-\Delta)^{-1/2}\sin(t(-\Delta)^{1/2}) \\ -(-\Delta)^{1/2}\sin(t(-\Delta)^{1/2}) & \cos(t(-\Delta)^{1/2}) \end{bmatrix}.$$

By Theorem (8.6), the unique weak solution $U$ of (15.3) is given by

$$U(t) = \int_0^t S(t-s)\,d\begin{bmatrix} 0 \\ W(s) \end{bmatrix} = \int_0^t \begin{bmatrix} (-\Delta)^{-1/2}\sin((t-s)(-\Delta)^{1/2}) \\ \cos((t-s)(-\Delta)^{1/2}) \end{bmatrix} dW_{L^2}(s), \tag{15.4}$$

provided both integrands are stochastically integrable with respect to $W_{L^2}$. Noting that the trigonometric functions $h_n(\xi) := \sqrt{2}\sin(n\pi\xi)$, $n \geqslant 1$, form an orthonormal basis of eigenfunctions for $\Delta$, by using Theorems 5.19 and 6.17 this is the case if and only if the following two conditions are satisfied:

$$\begin{aligned} &\int_0^T \sum_{n=1}^\infty [\sin^2(t(-\Delta)^{1/2})h_n, h_n]\,dt < \infty, \\ &\int_0^T \sum_{n=1}^\infty [\cos^2(t(-\Delta)^{1/2})h_n, h_n]\,dt < \infty. \end{aligned} \tag{15.5}$$

But if these conditions hold, then by adding we obtain $\int_0^T \sum_{n=1}^\infty [h_n, h_n]\,dt < \infty$, which is obviously false. We conclude that the problem (15.3) fails to have a weak solution in $\mathscr{H}$.

Instead of looking for a solution in $\mathscr{H}$, we could try to look for a solution in the larger space

$$\mathscr{G} := L^2(0,1) \times W^{-1,2}(0,1),$$

where $W^{-1,2}(0,1)$ denotes the completion of $L^2(0,1)$ with respect to the norm $\|f\|_{W^{-1,2}(0,1)} := \|(-\Delta)^{-1/2}f\|$. This definition of $W^{-1,2}(0,1)$ is motivated by the fact that $W^{1,2}(0,1)$ can be characterised as the domain of $(-\Delta)^{1/2}$. The space $\mathscr{G}$ is the so-called *extrapolation space* of $\mathscr{H}$ with respect to $(-\Delta)^{1/2}$; we refer to Exercise 5 for a more systematic discussion.

As is easy to check, the semigroup $S$ extends to a $C_0$-semigroup on $\mathscr{G}$, and the stochastic convolution (15.4) is well defined in $\mathscr{G}$ if and only if

$$\begin{aligned} &\int_0^T \sum_{n=1}^\infty [(-\Delta)^{-1}\sin^2(t(-\Delta)^{1/2})h_n, h_n]\,dt < \infty, \\ &\int_0^T \sum_{n=1}^\infty [(-\Delta)^{-1}\cos^2(t(-\Delta)^{1/2})h_n, h_n]\,dt < \infty. \end{aligned} \tag{15.6}$$

These conditions are indeed satified, as is clear from the identity $(-\Delta)^{-1}h_n = (n\pi)^{-2}h_n$.

Let us call a measurable adapted process $u : [0,T] \times \Omega \times (0,1) \to \mathbb{R}$ an *extrapolated weak solution* of (WE2) if $U(t,\omega) := \begin{bmatrix} u(t,\omega,\cdot) \\ \frac{\partial u}{\partial t} u(t,\omega,\cdot) \end{bmatrix}$ belongs to $\mathscr{G}$ for all $(t,\omega) \in [0,T] \times \Omega$ and the resulting process $U : [0,T] \times \Omega \to \mathscr{G}$ is a weak solution of the problem (WE2). Summarising the above discussion, we have proved:

**Theorem 15.5.** *The stochastic wave equation* (WE2) *admits a unique extrapolated weak solution.*

Here, uniqueness is understood in the sense of $\gamma(L^2(0,T;L^2(0,1)),\mathscr{G})$.

## 15.3 The stochastic heat equation

Next we consider two stochastic heat equations with Dirichlet boundary values, driven by multiplicative space-time white noise on a domain $D$ in $\mathbb{R}^d$:

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,\xi) = \Delta u(t,\xi) + B(u(t,\xi))\dfrac{\partial w}{\partial t}(t,\xi), & \xi \in D, \quad t \in [0,T], \\ u(t,\xi) = 0, & \xi \in \partial D, \quad t \in [0,T], \\ u(0,\xi) = u_0(\xi), & \xi \in D, \end{cases}$$

Again we discuss two particular cases of this problem: multiplicative rank one noise and additive space-time white noise. In both cases, the proofs of the main results can only be sketched, as they depend on a fair amount of interpolation theory and results from the theory of PDE. We refer to the Notes for references on this material.

### 15.3.1 Rank one multiplicative noise

Let $D$ be a bounded domain in $\mathbb{R}^d$ with smooth boundary $\partial D$. Our first example concerns the following stochastic heat equation driven by a rank one multiplicative noise:

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,\xi) = \Delta u(t,\xi) + b(u(t,\xi))\dfrac{\partial W}{\partial t}(t), & \xi \in D, \quad t \in [0,T], \\ u(t,\xi) = 0, & \xi \in \partial D, \quad t \in [0,T], \\ u(0,\xi) = u_0(\xi), & \xi \in D. \end{cases} \tag{HE1}$$

Here $W$ is standard real-valued Brownian motion. We assume that the function $b : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous.

We fix $1 < p < \infty$ and assume that the initial value $u_0$ belongs to $L^p(D)$. We say that a measurable adapted process $u : [0,T] \times \Omega \times D \to \mathbb{R}$ is a *mild $V_\theta^p$-solution* of (HE1) if $\xi \mapsto u(t,\omega,\xi)$ belongs to $L^p(D)$ for all $(t,\omega) \in [0,T] \times \Omega$

and the resulting process $U : [0,T] \times \Omega \to L^p(D)$ is a mild $V^p_\theta$-solution of the stochastic evolution equation

$$\begin{cases} dU(t) = AU(t)\,dt + B(U(t))\,dW(t), \\ U(0) = u_0. \end{cases} \tag{15.7}$$

Here $A$ is the Dirichlet Laplacian on $L^p(D)$ and $B : L^p(D) \to L^p(D)$ is the Nemytskii map associated with $b$,

$$(B(u))(\xi) := b(u(\xi)).$$

**Proposition 15.6.** *Under the above assumptions on $b$, the Nemytskii map $B : L^p(D) \to L^p(D)$ is well defined and $\gamma$-Lipschitz continuous with $\mathrm{Lip}_\gamma(B) \leqslant C_p \mathrm{Lip}(b)$, where $C_p$ is a constant depending only on $p$.*

*Proof.* Let us first note that $B(f) \in L^p(D)$ for all $f \in L^p(D)$, so $B$ is well defined.

It follows from the Kahane-Khintchine inequality that for all $f_1, \dots, f_N$ and $g_1, \dots, g_N$ in $L^p(D)$,

$$\begin{aligned}
&\Big(\mathbb{E}\Big\| \sum_{n=1}^N \gamma_n (B(f_n) - B(g_n))\Big\|^2_{L^p(D)}\Big)^{\frac12} \\
&\qquad \eqsim_p \Big(\mathbb{E}\Big\| \sum_{n=1}^N \gamma_n (B(f_n) - B(g_n))\Big\|^p_{L^p(D)}\Big)^{\frac1p} \\
&\qquad = \Big(\int_D \mathbb{E}\Big| \sum_{n=1}^N \gamma_n (b(f_n(\xi)) - b(g_n(\xi)))\Big|^p \,d\xi\Big)^{\frac1p} \\
&\qquad \eqsim_p \Big(\int_D \Big(\mathbb{E}\Big| \sum_{n=1}^N \gamma_n (b(f_n(\xi)) - b(g_n(\xi)))\Big|^2\Big)^{\frac p2} \,d\xi\Big)^{\frac1p} \\
&\qquad = \Big(\int_D \Big( \sum_{n=1}^N |b(f_n(\xi)) - b(g_n(\xi))|^2\Big)^{\frac p2} \,d\xi\Big)^{\frac1p} \\
&\qquad \leqslant \mathrm{Lip}(b)\Big(\int_D \Big( \sum_{n=1}^N |f_n(\xi) - g_n(\xi)|^2\Big)^{\frac p2} \,d\xi\Big)^{\frac1p} \\
&\qquad \eqsim_p \mathrm{Lip}(b)\Big(\mathbb{E}\Big\| \sum_{n=1}^N \gamma_n (f_n - g_n)\Big\|^2_{L^p(D)}\Big)^{\frac12},
\end{aligned}$$

where the last equivalence is obtained by doing the same computation backwards. Now we apply Proposition 14.1 (with $H = \mathbb{R}$). $\square$

Note that this result can be extended to Nemytskii maps on spaces $L^p(A)$, where $(A, \mathscr{A}, \mu)$ is any $\sigma$-finite measure space.

Let us say that an adapted process $u : [0,T] \times \Omega \times D \to \mathbb{R}$ is a *mild $V_\theta^p$-solution* of the problem (HE1) if if $\xi \mapsto u(t,\omega,\xi)$ belongs to $V_\theta^p$ for all $(t,\omega) \in [0,T] \times \Omega$ and the resulting process $U : [0,T] \times \Omega \to V_\theta^p$ is a mild $V_\theta^p$-solution of the problem (15.7).

**Theorem 15.7.** *Let $1 < p < \infty$, $\alpha \geqslant 0$, $\beta \geqslant 0$, $\theta \geqslant 0$ be such that $\alpha + 2\beta + d/p < 2\theta < 1$. Then the problem* (HE1) *has a unique mild $V_\theta^p$-solution $u$. This solution has a version with the property that $u - Su_0$ has trajectories in $C^\beta([0,T]; C^\alpha(\overline{D}))$, where $S$ denotes the semigroup generated by the Dirichlet Laplacian on $L^p(D)$.*

*Proof (Sketch).* We check the conditions of Theorem 14.16.

The space $E = L^p(D)$ is UMD and has Pisier's property and by Proposition 15.6, $B$ is $\gamma$-Lipschitz continuous from $E$ to $E$.

The Dirichlet Laplacian $A$ generates an analytic $C_0$-semigroup $S$ on $E$ (see Exercise 3). Choose numbers $0 \leqslant \eta < \eta' < \frac{1}{2}$ such that $\alpha + d/p < 2\eta$ and $\eta' + \beta < \theta$. The fractional domain space $E_{\eta'}$ associated with $A$ equals, up to an equivalent norm, the complex interpolation space $[E, \mathscr{D}(A)]_{\eta'}$.

Let $W^{2\eta,p}(D)$ be the Sobolev-Slobodetskii space of all functions $f : D \to \mathbb{R}$ such that

$$\|f\|_{W^{2\eta,p}(D)}^d := \|f\|_{L^p(D)}^p + \int_D \int_D \frac{|f(\xi) - f(\eta)|^p}{|\xi - \eta|^{d+2\eta p}} \, d\xi \, d\eta < \infty.$$

This space equals, up to an equivalent norm, the real interpolation space $(E, \mathscr{D}(A))_{\eta,p}$.

By general results in interpolation theory, we have a continuous embedding $[E, \mathscr{D}(A)]_{\eta'} \hookrightarrow (E, \mathscr{D}(A))_{\eta',p}$. By the above identifications, this results in a continuous embedding $E_{\eta'} \hookrightarrow W^{2\eta,p}(D)$.

Now we apply Theorem 14.16, which tells us that $U - Su_0$ has a version in with trajectories in $C^\beta([0,T]; E_{\eta'})$. By the above, this space embeds into $C^\beta([0,T]; W^{2\eta,p}(D))$. The proof is finished by an appeal to the Sobolev embedding theorem, which asserts that for $0 \leqslant \alpha < 2\eta - d/p$ we have a continuous embedding $W^{2\eta,p}(D) \hookrightarrow C^\alpha(\overline{D})$. □

### 15.3.2 Additive space-time white noise

Our final example is the stochastic heat equation driven by an additive space-time white noise:

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,\xi) = \Delta u(t,\xi) + \dfrac{\partial w}{\partial t}(t,\xi), & \xi \in (0,1), \quad t \in [0,T], \\ u(t,0) = u(t,1) = 0, & \qquad\qquad\quad t \in [0,T], \\ u(0,\xi) = u_0(\xi), & \xi \in (0,1). \end{cases} \tag{HE2}$$

Here $w$ is a space-time white noise on the unit interval $(0,1)$.

We formulate the problem (HE2) as an abstract stochastic evolution equation in $L^2(0,1)$ of the form

$$\begin{cases} dU(t) = AU(t)\,dt + dW_{L^2}(t), & t \geqslant 0, \\ U(0) = u_0, \end{cases} \tag{15.8}$$

where $A$ is the Dirichlet Laplacian on $L^2 := L^2(0,1)$ and $W_{L^2}$ is the $L^2$-cylindrical Brownian motion canonically associated with $W$. By a computation similar to (15.9) below (see Exercise 3) it is easy to check that the assumptions of Theorem 8.6 are satisfied, and therefore for initial values $u_0 \in L^2$ we obtain the existence of a unique weak solution $U$ of (15.8) in $L^2$. Note that in contrast to the situation for the wave equation, here it is not necessary to pass to an extrapolation space. The reason behind this is that the regularising effect of the heat semigroup takes us back into $L^2$; the wave semigroup does not have any such effect. It is nevertheless useful to consider the equation in a suitable extrapolation scale, as this enables us to obtain precise Hölder regularity results.

To this end we shall apply Theorem 10.19 in a suitable extrapolation space of $L^p := L^p(0,1)$. Fix $\delta > \frac{1}{4}$ and let $L^p_{-\delta}$ denote the *extrapolation space* of order $\delta$ associated with the Dirichlet Laplacian $A_p$ on $L^p$, that is, $L^p_{-\delta}$ is the completion of $L^p$ with respect to the norm $\|x\|_{-\delta} := \|(-A_p)^{-\delta}x\|$. Since $A_p$ is invertible on $L^p$ (see Exercise 4), $(-A_p)^{\delta}$ acts as an isomorphism from $L^p$ onto $L^p_{-\delta}$. We will show next that the identity operator $I$ on $L^2$ extends to a bounded embedding from $L^2$ into $L^p_{-\delta}$ which is $\gamma$-radonifying. Then, we will exploit the regularising effect of the semigroup $S$ to get back into a suitable Sobolev space contained in $L^p$ and use this to deduce regularity properties of the solution.

As is well known,

$$H_1 := \mathscr{D}(A) = W^{2,2}(0,1) \cap W^{1,2}_0(0,1)$$

and

$$E_1 := \mathscr{D}(A_p) = W^{2,p}(0,1) \cap W^{1,p}_0(0,1)$$

with equivalent norms.

The functions $h_n(\xi) := \sqrt{2}\sin(n\pi\xi)$, $n \geqslant 1$, form an orthonormal basis in $L^2$ of eigenfunctions for $A$ with eigenvalues $-\lambda_n$, where $\lambda_n = (n\pi)^2$. If we endow $H_1$ with the equivalent Hilbert norm $\|f\|_{H_1} := \|Af\|_2$, the functions $\lambda_n^{-1}h_n$ form an orthonormal basis for $H_1$ and we have

$$\begin{aligned} \mathbb{E}\Big\| \sum_{n=M}^{N} \gamma_n \lambda_n^{-1} h_n \Big\|^2_{L^p_{1-\delta}} &= \mathbb{E}\Big\| \sum_{n=M}^{N} \gamma_n \lambda_n^{-1} (-A_p)^{1-\delta} h_n \Big\|^2_p \\ &= \mathbb{E}\Big\| \sum_{n=M}^{N} \gamma_n (n\pi)^{-2\delta} h_n \Big\|^2_p \overset{(*)}{\lesssim} \sum_{n=M}^{N} (n\pi)^{-4\delta}, \end{aligned} \tag{15.9}$$

where $(*)$ follows from a square function estimate as in the proof of Proposition 15.6 together with the fact that $\|h_n\|_p \leqslant \sqrt{2}$. The right hand side of (15.9) tends to 0 as $M, N \to \infty$ since we took $\delta > \frac{1}{4}$. It follows that the identity operator on $H_1$ extends to a continuous embedding from $H_1$ into $L^p_{1-\delta}$ which is $\gamma$-radonifying. Denoting this embedding by $i_{-\delta}$, we obtain a commutative diagram

$$\begin{array}{ccc} H & \xrightarrow{I_{-\delta}} & L^p_{-\delta} \\ {\scriptstyle A^{-1}}\Big\downarrow & & \Big\uparrow{\scriptstyle A_p} \\ H_1 & \xrightarrow{i_{-\delta}} & L^p_{1-\delta} \end{array}$$

where the top mapping $I_{-\delta} : H \to L^p_{-\delta}$ is injective and $\gamma$-radonifying by the ideal property.

We are now in a position to apply Theorem 10.19. As before we assume that $u_0 \in L^2$. We say that a measurable adapted process $u : [0,T]\times\Omega\times(0,1) \to \mathbb{R}$ is a *weak solution* of (HE2) if $\xi \mapsto u(t,\omega,\xi)$ belongs to $L^2$ for all $(t,\omega) \in [0,T]\times\Omega$ and the resulting process $U : [0,T]\times\Omega \to L^2$ is a weak solution of the problem (15.8).

**Theorem 15.8.** *The problem* (HE2) *admits a unique weak solution $u$. For all $\alpha \geqslant 0$ and $\beta \geqslant 0$ satisfying $\alpha+2\beta < \frac{1}{2}$, the process $u - Su_0$ has a version with trajectories in $C^\beta([0,T]; C^\alpha_0[0,1])$, where $S$ denotes the semigroup generated by the Dirichlet Laplacian on $L^2(0,1)$.*

*Proof (Sketch).* Fix arbitrary real numbers $\alpha \geqslant 0$ and $\beta \geqslant 0$ satisfying $\alpha + 2\beta < \frac{1}{2}$. Replacing $\delta$ by a smaller number if necessary, we can find $\theta \geqslant 0$ such that $\frac{1}{4} < \delta < \theta$, $\beta+\theta < \frac{1}{2}$, and $\alpha+2\delta < 2\theta$. Put $\eta := \theta-\delta$. As is easy to check, (the extrapolation of) $A_p$ generates an analytic $C_0$-semigroup in $L^p_{-\delta}$. Hence we may apply Theorem 10.19 in the space $L^p_{-\delta}$ to obtain a weak solution $U$ of the problem

$$\begin{cases} dU(t) = AU(t)\,dt + I_{-\delta}\,dW_H(t), \quad t \in [0,T], \\ U(0) = 0, \end{cases}$$

with paths in the space $C^\beta\big([0,T]; (L^p_{-\delta})_\theta\big) = C^\beta\big([0,T]; L^p_\eta\big)$; the identity $(L^p_{-\delta})_\theta = L^p_\eta$ is a generalisation of Lemma 10.8. Along the embedding $L^2 \hookrightarrow L^p_{-\delta}$, this solution is consistent with the weak solution $U$ of this problem in $L^2$.

Noting that $\alpha < 2\eta$ we choose $p$ so large that $\alpha + \frac{1}{p} < 2\eta$. We have

$$L^p_\eta = W^{2\eta,p}_0(0,1)$$

with equivalent norms, and by the Sobolev embedding theorem,

$$W^{2\eta,p}(0,1) \hookrightarrow C^\alpha[0,1]$$

with continuous inclusion. We denote $C_0^\alpha[0,1] = \{f \in C^\alpha[0,1] : f(0) = f(1) = 0\}$. Putting things together we obtain a continuous inclusion

$$L_\eta^p \hookrightarrow C_0^\alpha[0,1].$$

In particular it follows that $U$ takes values in $L^p$. Almost surely, the trajectories of $U$ belong to $C^\beta([0,T]; C_0^\alpha[0,1])$.

If we compare Theorems 15.7 and 15.8 (for $d = 1$ and $D = (0,1)$), we notice that we get better Hölder regularity for the former ($\alpha + 2\beta < 1$ in the limit $p \to \infty$) than for the latter ($\alpha + 2\beta < \frac{1}{2}$). The explanation for this is the additional $\delta > \frac{1}{4}$ needed in Theorem 15.8 to get the $\gamma$-radonification of $B := I_{-\delta}$. In Theorem 15.7, $\gamma$-radonification came for free.

## 15.4 Exercises

1. Let $(A, \mathscr{A}, \mu)$ be a $\sigma$-finite measure space and put $H := L^2(A)$. Let $(W_H(t))_{t\in[0,T]}$ be an $H$-cylindrical Brownian motion. Show that

$$w([0,t] \times B) := W_H(t)1_B, \quad t \in [0,T],\ B \in \mathscr{A},$$

uniquely defines a space-time white noise $w$ on $A$.

2. Check the computations leading to the conditions (15.5) and (15.6).
*Hint:* A bounded operator $R : H_1 \to H_2$, where $H_1$ and $H_2$ are separable Hilbert spaces, is Hilbert-Schmidt if and only if $RR^* : H_2 \to H_2$ has finite trace.

3. In this exercise we take a look at the following stochastic heat equation with additive space-time white noise on the domain $D = (0,1)^d$ in $\mathbb{R}^d$.

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,\xi) = \Delta u(t,\xi) + \dfrac{\partial W}{\partial t}(t,\xi), & \xi \in D, \quad t \in [0,T], \\ u(t,\xi) = 0, & \xi \in \partial D, \quad t \in [0,T], \\ u(0,\xi) = u_0(\xi), & \xi \in D. \end{cases}$$

We model this problem as a stochastic evolution equation of the form (15.8).
 a) Prove that the Dirichlet Laplacian generates an analytic $C_0$-semigroup on $L^2(D)$.
 b) Show that the problem (15.8) has a weak solution in $L^2(D)$ if and only if $d = 1$.

*Hint for a) and b):* Find an orthonormal basis of eigenvectors.

4. Show that the heat semigroup generated by the Dirichlet Laplacian on $L^2(0,1)$ extends to an analytic $C_0$-semigroup on $L^p(0,1)$, $1 < p < \infty$, and show that its generator is invertible.

5. In this exercise we take a closer look at extrapolation spaces. Let $A$ be a densely defined closed operator on a Banach space $E$ and denote by $\mathscr{G}(A)$ its graph,
$$\mathscr{G}(A) = \{(x, Ax) \in E \times E : \ x \in \mathscr{D}(A)\}.$$
Define the *extrapolation space* of $E$ with respect to $A$ as the quotient space
$$E_{-1} := (E \times E)/\mathscr{G}(A).$$
   a) Show that the mapping $x \mapsto (0, x)$ defines a bounded dense embedding $E \hookrightarrow E_{-1}$.
   b) Show that $A_{-1} : x \mapsto (-x, 0)$ defined a bounded operator from $E$ to $E_{-1}$ which extends $A$.
   c) Show that if $\lambda \in \varrho(A)$, then the identity map on $E$ extends to an isomorphism of Banach spaces $E_{-1} \simeq E_{-1}^{\lambda}$, where the latter is defined as the completion of $E$ with respect to the norm $\|x\|_{E_{-1}^{\lambda}} := \|R(\lambda, A)x\|$.

**Notes.** The literature on stochastic partial differential equations is enormous and various approaches are possible. The functional analytic approach taken here, where the equation is reformulated as a stochastic evolution equation on some infinite-dimensional state space, goes back to Hille and Phillips in the deterministic case and give rise to the theory of $C_0$-semigroups. In the setting of Hilbert spaces, the theory of stochastic evolution equations was pioneered by Da Prato and Zabczyk and their schools. We refer to their monograph [27] for further references. See also Curtain and Pritchard [25] for some earlier references.

Our definition of a space-time white noise in Section 15.2 follows the lecture notes of Walsh [107], where also Theorem 15.5 can be found.

The presentation of Section 15.2.2 follows Da Prato and Zabczyk [27, Example 5.8].

Concerning problem (HE2), the existence of a solution in $C^{\alpha}([0,T] \times [0,1])$ for $0 \leqslant \alpha < \frac{1}{4}$ was proved by Da Prato and Zabczyk by very different methods; see [27, Theorem 5.20]. Theorem 15.8 was obtained by Brzeźniak [14] under more general assumptions. The approach taken here is from [34].

The results on interpolation theory needed in the proofs of Theorems 15.8 and 15.7 can be found in the book of Triebel [103].

# References


1. F. Albiac and N.J. Kalton, "Topics in Banach Space Theory", Graduate Texts in Mathematics, vol. 233, Springer, New York, 2006.
2. T.W. Anderson, *The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities*, Proc. Amer. Math. Soc. **6** (1955), 170–176.
3. W. Arendt, C.J.K. Batty, M. Hieber, and F. Neubrander, "Vector-valued Laplace Transforms and Cauchy problems", Monographs in Mathematics, vol. 96, Birkhäuser Verlag, Basel, 2001.
4. W. Arendt and S. Bu, *The operator-valued Marcinkiewicz multiplier theorem and maximal regularity*, Math. Z. **240** (2002), no. 2, 311–343.
5. J.M. Ball, *Strongly continuous semigroups, weak solutions, and the variation of constants formula*, Proc. Amer. Math. Soc. **63** (1977), no. 2, 370–373.
6. E. Berkson and T.A. Gillespie, *Spectral decompositions and harmonic analysis on UMD spaces*, Studia Math. **112** (1994), no. 1, 13–49.
7. P. Billingsley, "Convergence of Probability Measures", second ed., Wiley Series in Probability and Statistics, John Wiley & Sons Inc., New York, 1999.
8. V.I. Bogachev, "Gaussian Measures", Mathematical Surveys and Monographs, vol. 62, American Mathematical Society, Providence, RI, 1998.
9. V.I. Bogachev, "Measure Theory", Stringer-Verlag, Berlin, 2006.
10. J. Bourgain, *Some remarks on Banach spaces in which martingale difference sequences are unconditional*, Ark. Mat. **21** (1983), no. 2, 163–168.
11. J. Bourgain, *Real isomorphic complex Banach spaces need not be complex isomorphic*, Proc. Amer. Math. Soc. **96** (1986), no. 2, 221–226.
12. J. Bourgain, *Vector-valued singular integrals and the $H^1$-BMO duality*, in: "Probability Theory and Harmonic Analysis" (Cleveland, Ohio, 1983), Monogr. Textbooks Pure Appl. Math., vol. 98, Dekker, New York, 1986, pp. 1–19.
13. Z. Brzeźniak, *Stochastic partial differential equations in M-type* 2 *Banach spaces*, Potential Anal. **4** (1995), no. 1, 1–45.
14. Z. Brzeźniak, *On stochastic convolution in Banach spaces and applications*, Stochastics Stochastics Rep. **61** (1997), no. 3-4, 245–295.
15. Z. Brzeźniak, *Some remarks on Itô and Stratonovich integration in 2-smooth Banach spaces*, Probabilistic methods in fluids, World Sci. Publ., River Edge, NJ, 2003, pp. 48–69.

16. Z. Brzeźniak and J.M.A.M. van Neerven, *Stochastic convolution in separable Banach spaces and the stochastic linear Cauchy problem*, Studia Math. **143** (2000), no. 1, 43–74.
17. D.L. Burkholder, *A geometrical characterization of Banach spaces in which martingale difference sequences are unconditional*, Ann. Probab. **9** (1981), no. 6, 997–1011.
18. D.L. Burkholder, *A geometric condition that implies the existence of certain singular integrals of Banach-space-valued functions*, in: "Conference on Harmonic Analysis in Honor of Antoni Zygmund", Vol. I, II (Chicago, Ill., 1981), Wadsworth Math. Ser., Wadsworth, Belmont, CA, 1983, pp. 270–286.
19. D.L. Burkholder, *Martingales and Fourier analysis in Banach spaces*, in: "Probability and Analysis" (Varenna, 1985), Lecture Notes in Math., vol. 1206, Springer, Berlin, 1986, pp. 61–108.
20. D.L. Burkholder, *Martingales and singular integrals in Banach spaces*, in: "Handbook of the Geometry of Banach Spaces", Vol. I, North-Holland, Amsterdam, 2001, pp. 233–269.
21. K.L. Chung, "A Course in Probability Theory", third ed., Academic Press Inc., San Diego, CA, 2001.
22. K.L. Chung and R.J. Williams, "Introduction to Stochastic Integration", second ed., Probability and its Applications, Birkhäuser Boston Inc., Boston, MA, 1990.
23. Z. Ciesielski, *Modulus of smoothness of the Brownian paths in the $l^p$ norm*, in: "Constructive Theory of Functions (Varna, 1991), pp. 71–75.
24. P. Clément, B. de Pagter, F. A. Sukochev, and H. Witvliet, *Schauder decompositions and multiplier theorems*, Studia Math. **138** (2000), no. 2, 135–163.
25. R.F. Curtain and A.J. Pritchard, "Infinite Dimensional Linear Systems Theory, Lecture Notes in Control and Information Sciences, vol. 8, Springer-Verlag, Berlin, 1978.
26. G. Da Prato, S. Kwapień, and J. Zabczyk, *Regularity of solutions of linear stochastic equations in Hilbert spaces*, Stochastics **23** (1987), no. 1, 1–23.
27. G. Da Prato and J. Zabczyk, "Stochastic Equations in Infinite Dimensions", Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992.
28. G. Da Prato and J. Zabczyk, "Ergodicity for Infinite-Dimensional Systems", London Mathematical Society Lecture Note Series, vol. 229, Cambridge University Press, Cambridge, 1996.
29. E.B. Davies, "One-Parameter Semigroups", London Mathematical Society Monographs, vol. 15, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1980.
30. W.J. Davis, N. Ghoussoub, W.B. Johnson, S. Kwapień, and B. Maurey, *Weak convergence of vector valued martingales*, in: "Probability in Banach Spaces 6" (Sandbjerg, 1986), Progr. Probab., vol. 20, Birkhäuser Boston, Boston, MA, 1990, pp. 41–50.
31. V.H. de la Peña and E. Giné, "Decoupling", Probability and its Applications (New York), Springer-Verlag, New York, 1999.
32. R. Denk, M. Hieber, and J. Prüss, *R-boundedness, Fourier multipliers and problems of elliptic and parabolic type*, Mem. Amer. Math. Soc. **166** (2003), no. 788, viii+114.

33. E. Dettweiler, *Stochastic integration relative to Brownian motion on a general Banach space*, Doğa Mat. **15** (1991), no. 2, 58–97.
34. J. Dettweiler, J.M.A.M. van Neerven, and L. Weis, *Space-time regularity of solutions of the parabolic stochastic Cauchy problem*, Stoch. Anal. Appl. **24** (2006), no. 4, 843–869.
35. J. Diestel, H. Jarchow, and A. Tonge, "Absolutely Summing Operators", Cambridge Studies in Advanced Mathematics, vol. 43, Cambridge University Press, Cambridge, 1995.
36. J. Diestel and J.J. Uhl, Jr., "Vector Measures", American Mathematical Society, Providence, R.I., 1977.
37. N. Dunford and J.T. Schwartz, "Linear Operators, Part I: General Theory", Wiley Classics Library, John Wiley & Sons Inc., New York, 1988.
38. K.-J. Engel and R. Nagel, "One-Parameter Semigroups for Linear Evolution Equations", Graduate Texts in Mathematics, vol. 194, Springer-Verlag, New York, 2000.
39. X. Fernique, *Intégrabilité des vecteurs gaussiens*, C. R. Acad. Sci. Paris Sér. A-B **270** (1970), A1698–A1699.
40. D.J.H. Garling, *Brownian motion and UMD-spaces*, in: "Probability and Banach Spaces" (Zaragoza, 1985), Lecture Notes in Math., vol. 1221, Springer, Berlin, 1986, pp. 36–49.
41. J.A. Goldstein, "Semigroups of Linear Operators and Applications", Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1985.
42. B. Goldys and J.M.A.M. van Neerven, *Transition semigroups of Banach space-valued Ornstein-Uhlenbeck processes*, Acta Appl. Math. **76** (2003), no. 3, 283–330, Updated version on arXiv:math/0606785.
43. L. Gross, *Measurable functions on Hilbert space*, Trans. Amer. Math. Soc. **105** (1962), 372–390.
44. B.H. Haak, J.M.A.M. van Neerven, and M.C. Veraar, *A stochastic Datko-Pazy theorem*, J. Math. Anal. Appl. **329** (2007), no. 2, 1230–1239.
45. M. Haase, "The Functional Calculus for Sectorial Operators", Operator Theory: Advances and Applications, vol. 169, Birkhäuser Verlag, Basel, 2006.
46. M. Haase, *Convexity inequalities for positive operators*, Positivity **11** (2007), no. 1, 57–68.
47. E. Hille, "Functional Analysis and Semi-Groups", American Mathematical Society Colloquium Publications, vol. 31, American Mathematical Society, New York, 1948.
48. E. Hille and R.S. Phillips, "Functional Analysis and Semi-Groups", American Mathematical Society Colloquium Publications, vol. 31, American Mathematical Society, Providence, R. I., 1957, rev. ed.
49. J. Hoffmann-Jorgensen and G. Pisier, *The law of large numbers and the central limit theorem in Banach spaces*, Ann. Probability **4** (1976), no. 4, 587–599.
50. T. Hytönen, "R-Boundedness and Multiplier Theorems", Helsinki University of Technology, 2001.
51. T. Hytönen and M.C. Veraar, *On Besov regularity of Brownian motions in infinite dimensions*, Probab. Math. Statist **28** (2008), 143–162.
52. K. Itô and M. Nisio, *On the convergence of sums of independent Banach space valued random variables*, Osaka J. Math. **5** (1968), 35–48.

53. S. Janson, "Gaussian Hilbert Spaces", Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997.
54. J.-P. Kahane, "Some Random Series of Functions", second ed., Cambridge Studies in Advanced Mathematics, vol. 5, Cambridge University Press, Cambridge, 1985.
55. O. Kallenberg, "Foundations of Modern Probability", second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002.
56. N.J. Kalton, *An elementary example of a Banach space not isomorphic to its complex conjugate*, Canad. Math. Bull. **38** (1995), no. 2, 218–222.
57. N.J. Kalton and G. Lancien, *A solution to the problem of $L^p$-maximal regularity*, Math. Z. **235** (2000), no. 3, 559–568.
58. N.J. Kalton and L. Weis, *The $H^\infty$-calculus and square function estimates*, in preparation.
59. I. Karatzas and S.E. Shreve, "Brownian Motion and Stochastic Calculus, second ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991.
60. P. Kotelenez, *A maximal inequality for stochastic convolution integrals on Hilbert spaces and space-time regularity of linear stochastic partial differential equations*, Stochastics **21** (1987), no. 4, 345–358.
61. P.C. Kunstmann and L. Weis, *Maximal $L_p$-regularity for parabolic equations, Fourier multiplier theorems and $H^\infty$-functional calculus*, In: "Functional Analytic Methods for Evolution Equations", Lecture Notes in Math., vol. 1855, Springer, Berlin, 2004, pp. 65–311.
62. H.-H. Kuo, "Gaussian Measures in Banach Spaces", Springer-Verlag, Berlin, 1975, Lecture Notes in Mathematics, Vol. 463.
63. H.-H. Kuo, "Introduction to Stochastic Integration", Universitext, Springer, New York, 2006.
64. S. Kwapień, *Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients*, Studia Math. **44** (1972), 583–595.
65. S. Kwapień and W.A. Woyczyński, "Random Series and Stochastic Integrals: Single and Multiple", Probability and its Applications, Birkhäuser Boston Inc., Boston, MA, 1992.
66. S. Lang, "Real and Functional Analysis", Graduate Texts in Mathematics, vol. 142, Springer-Verlag, Berlin, 1993.
67. R. Latala and K. Oleszkiewicz, *On the best constant in the Khinchin-Kahane inequality*, Studia Math. **109** (1994), no. 1, 101–104.
68. M. Ledoux, "The Concentration of Measure Phenomenon", Mathematical Surveys and Monographs, vol. 89, American Mathematical Society, Providence, RI, 2001.
69. M. Ledoux and M. Talagrand, "Probability in Banach Spaces: Isoperimetry and Processes", Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 23, Springer-Verlag, Berlin, 1991.
70. V. Linde and A. Pietsch, *Mappings of Gaussian measures of cylindrical sets in Banach spaces*, Teor. Verojatnost. i Primenen. **19** (1974), 472–487, English translation in: Theory Probab. Appl. 19 (1974), 445–460.
71. G. Lumer and R.S. Phillips, *Dissipative operators in a Banach space*, Pacific J. Math. **11** (1961), 679–698.
72. A. Lunardi, "Analytic Semigroups and Optimal Regularity in Parabolic Problems", Progress in Nonlinear Differential Equations and their Applications, 16, Birkhäuser Verlag, Basel, 1995.

73. B. Maurey, *Système de Haar*, In: Séminaire Maurey-Schwartz 1974–1975: Espaces $L^p$, Applications Radonifiantes et Géométrie des Espaces de Banach", Exp. Nos. I et II, Centre Math., École Polytech., Paris, 1975, pp. 26 pp. (erratum, p. 1).
74. T.R. McConnell, *On Fourier multiplier transformations of Banach-valued functions*, Trans. Amer. Math. Soc. **285** (1984), no. 2, 739–757.
75. T.R. McConnell, *Decoupling and stochastic integration in UMD Banach spaces*, Probab. Math. Statist. **10** (1989), no. 2, 283–295.
76. M. Métivier and J. Pellaumail, "Stochastic Integration", Academic Press, New York, 1980.
77. A. Millet and W. Smoleński, *On the continuity of Ornstein-Uhlenbeck processes in infinite dimensions*, Probab. Theory Related Fields **92** (1992), no. 4, 529–547.
78. S. Montgomery-Smith, *Concrete representation of martingales*, Electron. J. Probab. **3** (1998), No. 15, 15 pp. (electronic).
79. G.A. Muñoz, Y. Sarantopoulos, and A. Tonge, *Complexifications of real Banach spaces, polynomials and multilinear maps*, Studia Math. **134** (1999), no. 1, 1–33.
80. J.M.A.M. van Neerven and M.C. Veraar, *On the action of Lipschitz functions on vector-valued random sums*, Arch. Math. (Basel) **85** (2005), no. 6, 544–553.
81. J.M.A.M. van Neerven, M.C. Veraar, and L. Weis, *Conditions for stochastic integrability in UMD Banach spaces*, in: "Banach Spaces and their Applications in Analysis", De Gruyter Proceedings in Mathematics, De Gruyter, 2007, pp. 125–146.
82. J.M.A.M. van Neerven, M.C. Veraar, and L. Weis, *Stochastic integration in UMD Banach spaces*, Annals Probab. **35** (2007), 1438–1478.
83. J.M.A.M. van Neerven, M.C. Veraar, and L. Weis, *Stochastic evolution equations in UMD Banach spaces*, J. Funct. Anal. **255** (2008), no. 4, 940–993.
84. J.M.A.M. van Neerven and L. Weis, *Stochastic integration of functions with values in a Banach space*, Studia Math. **166** (2005), no. 2, 131–170.
85. A.L. Neidhardt, "Stochastic Integrals in 2-Uniformly Smooth Banach Spaces", Ph.D. thesis, University of Wisconsin, 1978.
86. B. Oksendal, "Stochastic Differential Equations", 6th ed., Universitext, Springer-Verlag, Berlin, 2003.
87. B. de Pagter, "Martingales, UMD-Spaces, and the Vector-Valued Hilbert Transform", Lecture notes, Delft University of Technology, 1993/94.
88. K.R. Parthasarathy, "Probability Measures on Metric Spaces", Academic Press, New York-London, 1967.
89. A. Pazy, "Semigroups of Linear Operators and Applications to Partial Differential Equations", Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983.
90. B.J. Pettis, *On integration in vector spaces*, Trans. Amer. Math. Soc. **44** (1938), no. 2, 277–304.
91. G. Pisier, *Martingales with values in uniformly convex spaces*, Israel J. Math. **20** (1975), no. 3-4, 326–350.
92. G. Pisier, *Some results on Banach spaces without local unconditional structure*, Compositio Math. **37** (1978), no. 1, 3–19.
93. G. Pisier, *Holomorphic semigroups and the geometry of Banach spaces*, Ann. of Math. (2) **115** (1982), no. 2, 375–392.

94. D. Revuz and M. Yor, "Continuous Martingales and Brownian Motion", third ed., Grundlehren der Mathematischen Wissenschaften, vol. 293, Springer-Verlag, Berlin, 1999.
95. L.C.G. Rogers and D. Williams, "Diffusions, Markov processes, and Martingales, Vol. 1", Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2000, Foundations, Reprint of the second (1994) edition.
96. J. Rosiński and Z. Suchanecki, *On the space of vector-valued functions integrable with respect to the white noise*, Colloq. Math. **43** (1980), no. 1, 183–201 (1981).
97. J.L. Rubio de Francia, *Martingale and integral transforms of Banach space valued functions*, in: "Probability and Banach Spaces" (Zaragoza, 1985), Lecture Notes in Math., vol. 1221, Springer, Berlin, 1986, pp. 195–222.
98. L. Schwartz, *Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants)*, J. Analyse Math. **13** (1964), 115–256.
99. J.M. Steele, "Stochastic Calculus and Financial Applications", Applications of Mathematics (New York), vol. 45, Springer-Verlag, New York, 2001.
100. E.M. Stein, "Topics in Harmonic Analysis Related to the Littlewood-Paley Theory", Annals of Mathematics Studies, No. 63, Princeton University Press, Princeton, N.J., 1970.
101. K.R. Stromberg, "Probability for Analysts", Chapman & Hall Probability Series, Chapman & Hall, New York, 1994.
102. N. Tomczak-Jaegermann, "Banach-Mazur Distances and Finite-Dimensional Operator Ideals", Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 38, Longman Scientific & Technical, Harlow, 1989.
103. H. Triebel, "Interpolation Theory, Function Spaces, Differential Operators", North-Holland Mathematical Library, vol. 18, North-Holland Publishing Co., Amsterdam, 1978.
104. L. Tubaro, *An estimate of Burkholder type for stochastic processes defined by the stochastic integral*, Stochastic Anal. Appl. **2** (1984), no. 2, 187–192.
105. N.N. Vakhania, V.I. Tarieladze, and S.A. Chobanyan, "Probability Distributions on Banach Spaces", Mathematics and its Applications, vol. 14, D. Reidel Publishing Co., Dordrecht, 1987.
106. M.C. Veraar, "Stochastic Integrals in Banach Spaces and Applications to Parabolic Evolution Equations", Ph.D. thesis, Delft University of Technology, 2006.
107. J.B. Walsh, *An introduction to stochastic partial differential equations*, in: "École d'Été de Probabilités de Saint-Flour", XIV—1984, Lecture Notes in Math., vol. 1180, Springer, Berlin, 1986, pp. 265–439.
108. L. Weis, *Operator-valued Fourier multiplier theorems and maximal $L_p$-regularity*, Math. Ann. **319** (2001), no. 4, 735–758.
109. D. Williams, "Probability with Martingales", Cambridge Mathematical Textbooks, Cambridge University Press, Cambridge, 1991.
110. M. Yor, *Sur les intégrales stochastiques à valeurs dans un espace de Banach*, Ann. Inst. H. Poincaré Sect. B (N.S.) **10** (1974), 31–36.
111. K. Yosida, *On the differentiability and the representation of one-parameter semi-group of linear operators*, J. Math. Soc. Japan **1** (1948), 15–21.