Karlsruhe Institute of Technology

Stochastic Integration in L^p Spaces

Diploma Thesis
of

Markus Antoni

at the Institute for Analysis
of the Department of Mathematics

Reviewer: Prof. Dr. Lutz Weis
Second reviewer: Prof. Dr. Roland Schnaubelt

Contents

Introduction

In 1923, a stochastic integral $\int_0^1 f \,\mathrm{d}\beta$ was first introduced by Norbert Wiener in [16] for functions $f \in L^2([0,1])$ with respect to a Brownian motion β. In this case the following isometry holds

$$\mathbb{E}\Big|\int_0^1 f \,\mathrm{d}\beta\Big|^2 = \|f\|^2_{L^2([0,1])}.$$

21 years later, Kiyoshi Itô first presented a stochastic integral in [6] for adapted processes $f \in L^2([0,1]\times\Omega)$, where he showed the famous *Itô isometry*

$$\mathbb{E}\Big|\int_0^1 f \,\mathrm{d}\beta\Big|^2 = \|f\|^2_{L^2([0,1]\times\Omega)}.$$

From this theory the stochastic calculus arose, greatly influenced by *Itô's formula*. Today, these results are widely applied in various fields, especially in financial mathematics.

Based on this theory, it is now more or less easy to verify that these stochastic integrals and their isometries can be generalized to the Hilbert space valued setting. This is due to the fact that the norm comes from an inner product. For functions with values in an arbitrary Banach space stochastic integration is more difficult. In [11] by Jan van Neerven and Lutz Weis and in [10] by Jan van Neerven, Marc Veraar, and Lutz Weis, the authors studied stochastic integrability of operator-valued functions $\Phi\colon [0,T]\times\Omega \to \mathscr{B}(H,E)$. Here, H is a separable Hilbert space and E is a UMD space. An example of the latter one are the L^p spaces for $1 < p < \infty$. In general, UMD spaces are those Banach spaces E which satisfy

$$\mathbb{E}\Big\|\sum_{n=1}^{N} \varepsilon_n d_n\Big\|_E^q \lesssim_{q,E} \mathbb{E}\Big\|\sum_{n=1}^{N} d_n\Big\|_E^q$$

for any sequence $(\varepsilon_n)_{n=1}^N \subseteq [-1,1]$, each martingale difference sequence $(d_n)_{n=1}^N$, and some (and hence all) $1 < q < \infty$.

With this in mind, one may ask whether we could get new results or less elaborate proofs of the existent results if we restrict ourselves to L^p spaces by taking advantage of the structure we have there. Developing the theory of stochastic integration in these spaces will be the subject of this thesis. Due to the fact that, in what follows, the number p is reserved for another purpose, we will from now on replace the p with an r and speak of L^r spaces. More precisely, we will consider the space $L^r(U, \Sigma, \mu)$ for $1 \leq r < \infty$ (or $1 < r < \infty$, respectively), where (U, Σ, μ) is a σ-finite measure space and Σ is countably generated. Note that in this case $L^r(U, \Sigma, \mu)$ is separable.

The results of this thesis are based on various results from stochastic analysis and martingale theory. Among other things, we will give a less elaborate proof of the Kahane inequalities for Gaussian and Rademacher sums, and we will show a representation theorem for Gaussian random variables. This will be the topic of Chapter 1, which will be completed by an introduction of the Brownian motion. In Chapter 2, we take a closer look on martingale difference sequences and the UMD property. The highlight of this chapter may be a new (stronger) version of Doob's maximal inequality, which states that for $1 < p, r < \infty$ and suitable $L^r(U, \Sigma, \mu)$-valued martingales $(M_n)_{n=1}^N$ we have

$$\mathbb{E}\Big\|\max_{n=1}^{N} |M_n|\Big\|_r^p \lesssim_{p,r} \mathbb{E}\, \|M_N\|_r^p \,.$$

This will then also lead to a new version of the Burkholder-Gundy inequality stating that

$$\mathbb{E}\Big\|\max_{n=1}^{N} |M_n|\Big\|_r^p \eqsim_{p,r} \mathbb{E}\Big\| \Big(\sum_{n=1}^{N} |d_n|^2\Big)^{\frac{1}{2}} \Big\|_r^p$$

for $1 < p, r < \infty$. Here, $(d_n)_{n=1}^N$ is the martingale difference sequence of the $L^r(U, \Sigma, \mu)$-valued martingale $(M_n)_{n=1}^N$.

Having laid these foundations, we will continue with the theory of stochastic integration in Chapter 3. We will first construct a stochastic integral for functions $f\colon [0, T] \to L^r(U, \Sigma, \mu)$, where $1 \leq r < \infty$, and we will see that those integrals belong to the family of Gaussian random variables. In the next step, we present the stochastic integral for appropriate processes $f\colon [0, T] \times \Omega \to L^r(U, \Sigma, \mu)$, now assuming that $1 < r < \infty$. Based on that, we will study the integral process $\big(\int_0^t f \,\mathrm{d}\beta\big)_{t\in[0,T]}$, which happens to be a martingale. Applying the results from Chapter 2 will therefore lead to a Burkholder-Gundy inequality for stochastic integrals, i.e.,

$$\mathbb{E}\Big\| \sup_{t\in[0,T]} \Big| \int_0^t f \,\mathrm{d}\beta \Big| \Big\|_r^p \eqsim_{p,r} \mathbb{E}\Big\| \Big(\int_0^T |f|^2 \,\mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r^p$$

for $1 < p < \infty$. In the last part of this chapter, we will finally use a localization argument to augment the class of integrands significantly.

In the final chapter we consider *Itô processes*, i.e., processes of the form

$$X(t) = x_0 + \int_0^t f \,\mathrm{d}s + \sum_{n=1}^{\infty} \int_0^t b_n \,\mathrm{d}\beta_n, \quad t \in [0,T],$$

for suitable functions f and b_n, $n \in \mathbb{N}$. Here, $(\beta_n)_{n=1}^{\infty}$ is a sequence of independent Brownian motions. At this point, we will show an interesting connection to the operator-valued stochastic integral we mentioned above. We also prove the Itô formula, which will subsequently be used to solve an abstract stochastic evolution equation.

Finally, in the appendix we provide an introduction to the theory of integration in Banach spaces, present some facts about Gaussian random variables, and show some results regarding vector-valued conditional expectations and martingales that are used in this thesis. Especially for readers who have not collected some experience in these fields we recommend to read the appendix first.

Notations. In this thesis, (U, Σ, μ) will always be a σ-finite measure space with countably generated σ-algebra Σ. For a number $1 \le r \le \infty$, we define $L^r(U) := L^r(U, \Sigma, \mu)$ as a Banach space over $\mathbb{R}$, and the Hölder conjugate of r by $r' := \frac{r}{r-1}$ (with $1' := \infty$ and $\infty' := 1$). The norm of this space will be abbreviated by $\|\cdot\|_r$. Moreover, for $1 \le r < \infty$, we identify the duality space $L^r(U)^*$ with the space $L^{r'}(U)$ since every continuous functional $T \in L^r(U)^*$ is given by a multiplication operator with unique kernel $g \in L^{r'}(U)$, i.e., $T_g(f) := T(f) = \int_U fg \,\mathrm{d}\mu$. Thus, by

$$\langle f, g \rangle := \langle f, T_g \rangle = T_g(f) = \int_U fg \,\mathrm{d}\mu$$

we denote the duality pairing of the elements $f \in L^r(U)$ and $g \in L^{r'}(U)$. If $a \le C(q)\, b$ for nonnegative numbers a and b and a constant $C(q) > 0$ depending only on the variable q, we write $a \lesssim_q b$. Additionally, we write $a \eqsim_q b$ if $a \lesssim_q b$ and $b \lesssim_q a$. Finally, for real numbers x and y, we define $x \vee y := \max\{x, y\}$ and $x \wedge y := \min\{x, y\}$.

Acknowledgements. First of all, I would like to express my deepest appreciation to my advisor Lutz Weis for providing outstanding support and guidance at every stage of this thesis. Further, I would like to thank him, Maria Girardi, and Holly Watson for the kind hospitality during my stay at the University of South Carolina, where parts of this thesis were made. Many thanks go to Roland Schnaubelt for the pleasure to enjoy four years of excellent teaching. Finally, I would like to express my gratitude to my fellow students, especially Christiane Görgen, Jörg Bäuerle, and Alexander Weber, for five years of great assistance throughout my studies.

Chapter 1

Random Series

In this chapter we collect some estimates for sums of independent random variables. More precisely, we concentrate on sums of the form $\sum_{n=1}^{N} \gamma_n x_n$ and $\sum_{n=1}^{N} r_n x_n$, where the γ_n are $\mathbb{R}$-valued Gaussian variables, r_n are Rademacher variables, and the x_n are elements of $L^r(U)$. In section 1.2 we will then see that such sums are the basic modules for general $L^r(U)$-valued Gaussian random variables, and in Chapter 3 we again run across these sums since stochastic integrals of $L^r(U)$-valued step functions are of this form. In section 1.3 we investigate the Brownian motion process and use results of section 1.2 to prove certain path properties of Brownian motions.

1.1 The Kahane Inequality

Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space. We call a random variable $r \colon \Omega \to \{-1, +1\}$ a *Rademacher variable* (or *Bernoulli variable*) if

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

A random variable $\gamma \colon \Omega \to \mathbb{R}$ is said to be *Gaussian* if its distribution has density

$$f_\gamma \colon \mathbb{R} \to \mathbb{R}, \quad f_\gamma(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\tfrac{1}{2}\tfrac{t^2}{\sigma^2}\right)$$

with respect to the Lebesgue measure on $\mathbb{R}$ for some $\sigma \geq 0$. Note that a Gaussian variable is always centered, i.e., the mean value of a Gaussian variable is 0. If $\sigma = 1$, γ is called a *standard Gaussian variable.*

For the rest of this section, $(r_n)_{n=1}^{\infty}$ and $(\gamma_n)_{n=1}^{\infty}$ will always denote a sequence of independent Rademacher variables and a sequence of independent standard Gaussian variables, respectively.

Lemma 1.1. *For all $1 \leq p < \infty$ and all finite sequences $\alpha_1, \dots, \alpha_N \in \mathbb{R}$, we have*

$$\Big(\mathbb{E}\Big|\sum_{n=1}^{N} \alpha_n \gamma_n\Big|^p\Big)^{\frac{1}{p}} = c(p)\Big(\sum_{n=1}^{N} |\alpha_n|^2\Big)^{\frac{1}{2}},$$

where $c(p) = \sqrt{2}\pi^{-\frac{1}{2p}}\Gamma\big(\frac{p+1}{2}\big)^{\frac{1}{p}}$ and Γ is the Gamma function defined by

$$\Gamma\colon (0,\infty) \to \mathbb{R}, \quad \Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,\mathrm{d}t.$$

Proof. First we compute for $r > -1$

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |t|^r \exp\big(-\tfrac{t^2}{2}\big)\,\mathrm{d}t &= \sqrt{\tfrac{2}{\pi}} \int_0^\infty t^r \exp\big(-\tfrac{t^2}{2}\big)\,\mathrm{d}t \\ &= \sqrt{\tfrac{2}{\pi}}\, 2^{\frac{r-1}{2}} \int_0^\infty s^{\frac{r-1}{2}} \exp(-s)\,\mathrm{d}s \\ &= \sqrt{\tfrac{2}{\pi}}\, 2^{\frac{r-1}{2}} \Gamma\big(\tfrac{r+1}{2}\big), \end{aligned}$$

where we used the substitution $s = \frac{t^2}{2}$ in the second equality. If $\alpha_n = 0$ for each $1 \leq n \leq N$, there is nothing to proof. So we can assume that there exists at least one α_k with $\alpha_k \neq 0$. Next, by the independence of the Gaussian sequence, we remark that the linear combination $\sum_{n=1}^{N} \alpha_n \gamma_n$ is again a Gaussian variable with variance $\sum_{n=1}^{N} \alpha_n^2$. Using the foregoing computation, we obtain

$$\begin{aligned} \mathbb{E}\Big|\sum_{n=1}^{N} \alpha_n \gamma_n\Big|^p &= \frac{1}{\sqrt{2\pi \sum_{n=1}^{N} \alpha_n^2}} \int_{-\infty}^{\infty} |t|^p \exp\Big(-\tfrac{1}{2}\tfrac{t^2}{\sum_{n=1}^{N} \alpha_n^2}\Big)\,\mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}} \Big(\sum_{n=1}^{N} \alpha_n^2\Big)^{\frac{p}{2}} \int_{-\infty}^{\infty} |s|^p \exp\big(-\tfrac{s^2}{2}\big)\,\mathrm{d}s \\ &= \sqrt{\tfrac{2}{\pi}}\, 2^{\frac{p-1}{2}} \Gamma\big(\tfrac{p+1}{2}\big) \Big(\sum_{n=1}^{N} \alpha_n^2\Big)^{\frac{p}{2}}, \end{aligned}$$

applying the substitution $s = t\,\big(\sum_{n=1}^{N} \alpha_n^2\big)^{-\frac{1}{2}}$ in the second equality. ■

If the Gaussian sequence in Lemma 1.1 is replaced by a Rademacher sequence, we do not get equality but at least some sort of equivalence. For the proof of this estimate, we first need the following proposition.

Proposition 1.2. *Let $f\colon \Omega \to \mathbb{R}$ be a measurable function. Then for all $1 \le p < \infty$ we have*

$$\int_\Omega |f|^p \,\mathrm{d}\mathbb{P} = \int_0^\infty p\lambda^{p-1}\mathbb{P}\big(|f| \ge \lambda\big)\,\mathrm{d}\lambda.$$

Proof. With Fubini's theorem we get

$$\begin{aligned}\int_0^\infty p\lambda^{p-1}\mathbb{P}\big(|f|\ge\lambda\big)\,\mathrm{d}\lambda &= \int_0^\infty p\lambda^{p-1}\int_\Omega \mathbb{1}_{\{|f|\ge\lambda\}}\,\mathrm{d}\mathbb{P}\,\mathrm{d}\lambda = \int_\Omega\int_0^{|f|} p\lambda^{p-1}\,\mathrm{d}\lambda\,\mathrm{d}\mathbb{P}\\ &= \int_\Omega |f|^p\,\mathrm{d}\mathbb{P}.\end{aligned}$$

■

Theorem 1.3 (Khintchine inequality). *For all $1 \le p < \infty$ and all finite sequences $\alpha_1, \dots, \alpha_N \in \mathbb{R}$, we have*

$$\left(\mathbb{E}\Big|\sum_{n=1}^N \alpha_n r_n\Big|^p\right)^{\frac{1}{p}} \eqsim_p \left(\sum_{n=1}^N |\alpha_n|^2\right)^{\frac{1}{2}}.$$

Proof. (1) We first consider the upper bound. In view of

$$2^n n! \le (2n)! \quad \text{for all } n \in \mathbb{N}_0,$$

we get for all $x \in \mathbb{R}$

$$\begin{aligned}\frac{1}{2}\big(\exp(x)+\exp(-x)\big) &= \frac{1}{2}\left(\sum_{n=0}^\infty \frac{x^n}{n!} + \sum_{n=0}^\infty \frac{(-x)^n}{n!}\right) = \sum_{n=0}^\infty \frac{x^{2n}}{(2n)!}\\ &\le \sum_{n=0}^\infty \frac{\left(\frac{1}{2}x^2\right)^n}{n!} = \exp\big(\tfrac{1}{2}x^2\big).\end{aligned} \tag{$\triangle$}$$

Now let $t > 0$ be fixed and define $\Omega_+ := \big\{t\sum_{n=1}^N \alpha_n r_n \ge 0\big\}$. Then

$$\begin{aligned}\mathbb{E}\exp\left(\Big|t\sum_{n=1}^N \alpha_n r_n\Big|\right) &= \int_{\Omega_+} \exp\Big(t\sum_{n=1}^N\alpha_n r_n\Big)\,\mathrm{d}\mathbb{P} + \int_{\Omega_+^C}\exp\Big(-t\sum_{n=1}^N \alpha_n r_n\Big)\,\mathrm{d}\mathbb{P}\\ &\le \mathbb{E}\exp\Big(t\sum_{n=1}^N\alpha_n r_n\Big) + \mathbb{E}\exp\Big(-t\sum_{n=1}^N\alpha_n r_n\Big)\\ &= \mathbb{E}\left(\prod_{n=1}^N \exp(t\alpha_n r_n)\right) + \mathbb{E}\left(\prod_{n=1}^N\exp(-t\alpha_n r_n)\right) = (\star).\end{aligned}$$

Next, using the stochastic independence of the Rademacher sequence and ($\triangle$), we obtain

$$\begin{aligned}
(\star) &= \prod_{n=1}^{N} \mathbb{E}\exp(t\alpha_n r_n) + \prod_{n=1}^{N} \mathbb{E}\exp(-t\alpha_n r_n) \\
&= \prod_{n=1}^{N} \frac{1}{2}\big(\exp(t\alpha_n) + \exp(-t\alpha_n)\big) + \prod_{n=1}^{N} \frac{1}{2}\big(\exp(-t\alpha_n) + \exp(t\alpha_n)\big) \\
&\leq 2\prod_{n=1}^{N} \exp\big(\tfrac{t^2}{2}\alpha_n^2\big) = 2\exp\Big(\tfrac{t^2}{2}\sum_{n=1}^{N}\alpha_n^2\Big).
\end{aligned}$$

Therefore, by Chebyshev's inequality, we deduce that

$$\begin{aligned}
\mathbb{P}\Big(\Big|\sum_{n=1}^{N}\alpha_n r_n\Big| \geq \lambda\Big) &\leq \exp(-t\lambda)\mathbb{E}\exp\Big(\Big|t\sum_{n=1}^{N}\alpha_n r_n\Big|\Big) \\
&\leq 2\exp\Big(-t\lambda + \tfrac{t^2}{2}\sum_{n=1}^{N}\alpha_n^2\Big)
\end{aligned}$$

for any $\lambda > 0$. Taking $t := \frac{\lambda}{\sum_{n=1}^{N}\alpha_n^2}$ leads to

$$\mathbb{P}\Big(\Big|\sum_{n=1}^{N}\alpha_n r_n\Big| \geq \lambda\Big) \leq 2\exp\Big(-\frac{1}{2}\frac{\lambda^2}{\sum_{n=1}^{N}\alpha_n^2}\Big).$$

Finally, Proposition 1.2 yields

$$\begin{aligned}
\mathbb{E}\Big|\sum_{n=1}^{N}\alpha_n r_n\Big|^p &= \int_0^\infty p\lambda^{p-1}\mathbb{P}\Big(\Big|\sum_{n=1}^{N}\alpha_n r_n\Big| \geq \lambda\Big)\,\mathrm{d}\lambda \\
&\leq 2\int_0^\infty p\lambda^{p-1}\exp\Big(-\frac{1}{2}\frac{\lambda^2}{\sum_{n=1}^{N}\alpha_n^2}\Big)\,\mathrm{d}\lambda \\
&= p\,2^{\frac{p}{2}}\Big(\sum_{n=1}^{N}\alpha_n^2\Big)^{\frac{p}{2}}\int_0^\infty s^{\frac{p}{2}-1}\exp(-s)\,\mathrm{d}s \\
&= p\,2^{\frac{p}{2}}\,\Gamma\big(\tfrac{p}{2}\big)\Big(\sum_{n=1}^{N}\alpha_n^2\Big)^{\frac{p}{2}},
\end{aligned}$$

where we used the substitution $s = \frac{1}{2}\frac{\lambda^2}{\sum_{n=1}^{N}\alpha_n^2}$.

(2) First, notice that for each $n = 1, \ldots, N$ we have

$$\mathbb{E}r_n = 0 \quad \text{and} \quad \mathbb{E}r_n^2 = 1.$$

Therefore, using the stochastic independence of the Rademacher sequence, we obtain

$$\begin{aligned}\mathbb{E}\Big|\sum_{n=1}^N \alpha_n r_n\Big|^2 &= \sum_{n=1}^N\sum_{m=1}^N \mathbb{E}(\alpha_n\alpha_m r_n r_m)\\ &= \sum_{n=1}^N \mathbb{E}(\alpha_n r_n)^2 + \sum_{n=1}^N \sum_{m=1,m\neq n}^N \alpha_n\alpha_m\,\mathbb{E}r_n\,\mathbb{E}r_m\\ &= \sum_{n=1}^N \alpha_n^2.\end{aligned}$$

By Hölder's inequality and step **(1)** (with $p=4$) we get

$$\begin{aligned}\sum_{n=1}^N \alpha_n^2 = \mathbb{E}\Big|\sum_{n=1}^N \alpha_n r_n\Big|^2 &= \mathbb{E}\Big|\sum_{n=1}^N \alpha_n r_n\Big|^{\frac{2}{3}}\Big|\sum_{n=1}^N \alpha_n r_n\Big|^{\frac{4}{3}}\\ &\le \Big(\mathbb{E}\Big|\sum_{n=1}^N \alpha_n r_n\Big|\Big)^{\frac{2}{3}}\Big(\mathbb{E}\Big|\sum_{n=1}^N \alpha_n r_n\Big|^4\Big)^{\frac{1}{3}}\\ &\le \Big(\mathbb{E}\Big|\sum_{n=1}^N \alpha_n r_n\Big|\Big)^{\frac{2}{3}} 16^{\frac{1}{3}}\Big(\sum_{n=1}^N \alpha_n^2\Big)^{\frac{2}{3}},\end{aligned}$$

and thus we obtain

$$\Big(\sum_{n=1}^N \alpha_n^2\Big)^{\frac{1}{2}} \le 4\,\mathbb{E}\Big|\sum_{n=1}^N \alpha_n r_n\Big| \le 4\,\Big(\mathbb{E}\Big|\sum_{n=1}^N \alpha_n r_n\Big|^p\Big)^{\frac{1}{p}}$$

for all $1\le p<\infty$. ∎

In what follows, we assume that $1\le r<\infty$ is fixed. Next, we consider finite sequences $x_1,\dots,x_N\in L^r(U)$ instead of coefficients $\alpha_1,\dots,\alpha_N\in\mathbb{R}$.

Theorem 1.4 (Kahane inequalities). *Let $1\le p<\infty$ and set $q:=p\vee r$. Then we have for all $x_1,\dots,x_N\in L^r(U)$*

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^N \gamma_n x_n\Big\|_r^p\Big)^{\frac{1}{p}} \eqsim_q \Big\|\Big(\sum_{n=1}^N |x_n|^2\Big)^{\frac{1}{2}}\Big\|_r$$

and

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^N r_n x_n\Big\|_r^p\Big)^{\frac{1}{p}} \eqsim_q \Big\|\Big(\sum_{n=1}^N |x_n|^2\Big)^{\frac{1}{2}}\Big\|_r.$$

Proof. Let $(\xi_n)_{n=1}^N \in \big\{(\gamma_n)_{n=1}^N, (r_n)_{n=1}^N\big\}$. Due to Lemma 1.1 or Theorem 1.3, respectively, we have

$$\Big(\mathbb{E}\Big|\sum_{n=1}^N \xi_n x_n(u)\Big|^p\Big)^{\frac{1}{p}} \eqsim_p \Big(\sum_{n=1}^N |x_n(u)|^2\Big)^{\frac{1}{2}}$$

for each fixed $u \in U$. Thus, we obtain

$$\Big\|\Big(\mathbb{E}\Big|\sum_{n=1}^N \xi_n x_n\Big|^p\Big)^{\frac{1}{p}}\Big\|_r \eqsim_p \Big\|\Big(\sum_{n=1}^N |x_n|^2\Big)^{\frac{1}{2}}\Big\|_r.$$

(1) For $p \geq r$, using Minkowski's integral inequality, we have

$$\begin{aligned}\Big(\mathbb{E}\Big\|\sum_{n=1}^N \xi_n x_n\Big\|_r^p\Big)^{\frac{1}{p}} &= \Big(\int_\Omega \Big(\int_U \Big|\sum_{n=1}^N \xi_n x_n\Big|^r \mathrm{d}\mu\Big)^{\frac{p}{r}} \mathrm{d}\mathbb{P}\Big)^{\frac{1}{p}} \\ &\leq \Big(\int_U \Big(\int_\Omega \Big|\sum_{n=1}^N \xi_n x_n\Big|^p \mathrm{d}\mathbb{P}\Big)^{\frac{r}{p}} \mathrm{d}\mu\Big)^{\frac{1}{r}} \\ &= \Big\|\Big(\mathbb{E}\Big|\sum_{n=1}^N \xi_n x_n\Big|^p\Big)^{\frac{1}{p}}\Big\|_r \eqsim_p \Big\|\Big(\sum_{n=1}^N |x_n|^2\Big)^{\frac{1}{2}}\Big\|_r.\end{aligned}$$

For $p < r$, Hölder's inequality and the foregoing estimate lead to

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^N \xi_n x_n\Big\|_r^p\Big)^{\frac{1}{p}} \leq \Big(\mathbb{E}\Big\|\sum_{n=1}^N \xi_n x_n\Big\|_r^r\Big)^{\frac{1}{r}} \lesssim_r \Big\|\Big(\sum_{n=1}^N |x_n|^2\Big)^{\frac{1}{2}}\Big\|_r.$$

(2) For the converse inequality, we apply Proposition A.10 and again Hölder's inequality to get

$$\Big\|\Big(\sum_{n=1}^N |x_n|^2\Big)^{\frac{1}{2}}\Big\|_r \lesssim \Big\|\mathbb{E}\Big|\sum_{n=1}^N \xi_n x_n\Big|\Big\|_r \leq \mathbb{E}\Big\|\sum_{n=1}^N \xi_n x_n\Big\|_r \leq \Big(\mathbb{E}\Big\|\sum_{n=1}^N \xi_n x_n\Big\|_r^p\Big)^{\frac{1}{p}}$$

for all $1 \leq p < \infty$. ■

Remark 1.5. As an immediate consequence of Theorem 1.4, we obtain the following estimate

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^N \xi_{1n} x_n\Big\|_r^p\Big)^{\frac{1}{p}} \eqsim_{p,q,r} \Big(\mathbb{E}\Big\|\sum_{n=1}^N \xi_{2n} x_n\Big\|_r^q\Big)^{\frac{1}{q}}$$

for all $1 \leq p, q < \infty$ and every sequence $x_1, \ldots, x_N \in L^r(U)$, where $(\xi_{1n})_{n=1}^N$, $(\xi_{2n})_{n=1}^N \in \big\{(\gamma_n)_{n=1}^N, (r_n)_{n=1}^N\big\}$ can be chosen arbitrarily. ■

The next corollary shows that, in the situation of a convergent Gaussian or Rademacher sum, we even have unconditional convergence.

Corollary 1.6 (Contraction principle). *Let* $(\varepsilon_n)_{n=1}^N \subseteq [-1,1]$ *and* $1 \leq p < \infty$. *Then we have for all* $x_1, \dots, x_N \in L^r(U)$

$$\left(\mathbb{E}\Big\| \sum_{n=1}^N \varepsilon_n \gamma_n x_n \Big\|_r^p\right)^{\frac{1}{p}} \lesssim_{p,r} \left(\mathbb{E}\Big\| \sum_{n=1}^N \gamma_n x_n \Big\|_r^p\right)^{\frac{1}{p}}$$

and

$$\left(\mathbb{E}\Big\| \sum_{n=1}^N \varepsilon_n r_n x_n \Big\|_r^p\right)^{\frac{1}{p}} \lesssim_{p,r} \left(\mathbb{E}\Big\| \sum_{n=1}^N r_n x_n \Big\|_r^p\right)^{\frac{1}{p}}.$$

1.2 The Karhunen-Loève Representation Theorem

In this section we will discuss some properties of $L^r(U)$-valued Gaussian random variables, which are defined as a function $X\colon \Omega \to L^r(U)$ on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$ such that the $\mathbb{R}$-valued random variable $\langle X, g\rangle$ is Gaussian for all $g \in L^{r'}(U)$, where $1 \leq r < \infty$ is again fixed. The main result will be a representation theorem, which states that every Gaussian random variable can be represented as a (convergent) sum of the form $\sum_{n=1}^\infty \gamma_n x_n$. Note that some auxiliary results regarding Gaussian random variables are outlined in the Appendix.

First we consider the fundamental integrability result for Gaussian random variables due to Fernique.

Theorem 1.7 (Fernique). *Let* X *be an* $L^r(U)$*-valued Gaussian random variable. Then there exists a constant* $\beta > 0$ *such that*

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

Proof. (1) Let Y be an independent copy of X, and let $t \geq s > 0$ be fixed. Then, by Proposition A.15,

$$V := \frac{X+Y}{\sqrt{2}} \quad \text{and} \quad W := \frac{X-Y}{\sqrt{2}}$$

are independent and have the same distribution as X and Y. Moreover, observe that the set

$$\Big\{(x,y)\in\mathbb{R}_+^2 : |x-y|\le s\sqrt{2},\ x+y>t\sqrt{2}\Big\}$$

is a subset of

$$\Big\{(x,y)\in\mathbb{R}_+^2 : x>\frac{t-s}{\sqrt{2}},\ y>\frac{t-s}{\sqrt{2}}\Big\}.$$

Therefore, we obtain

$$\begin{aligned}
\mathbb{P}\big(\|X\|_r\le s\big)\cdot\mathbb{P}\big(\|Y\|_r>t\big) &= \mathbb{P}\Big(\Big\|\frac{X-Y}{\sqrt{2}}\Big\|_r\le s\Big)\cdot\mathbb{P}\Big(\Big\|\frac{X+Y}{\sqrt{2}}\Big\|_r>t\Big)\\
&\le \mathbb{P}\Big(\Big|\frac{\|X\|_r-\|Y\|_r}{\sqrt{2}}\Big|\le s,\ \frac{\|X\|_r+\|Y\|_r}{\sqrt{2}}>t\Big)\\
&\le \mathbb{P}\Big(\|X\|_r>\frac{t-s}{\sqrt{2}},\|Y\|_r>\frac{t-s}{\sqrt{2}}\Big)\\
&= \mathbb{P}\Big(\|X\|_r>\frac{t-s}{\sqrt{2}}\Big)\cdot\mathbb{P}\Big(\|Y\|_r>\frac{t-s}{\sqrt{2}}\Big).
\end{aligned}$$

Since X and Y have the same distribution, this leads to

$$\mathbb{P}\big(\|X\|_r\le s\big)\cdot\mathbb{P}\big(\|X\|_r>t\big)\le\mathbb{P}\Big(\|X\|_r>\frac{t-s}{\sqrt{2}}\Big)^2.$$

(2) Choose $s_0>0$ such that $\mathbb{P}\big(\|X\|_r\le s_0\big)\ge\frac{2}{3}$, and define

$$t_0:=s_0\quad\text{and}\quad t_n:=s_0+\sqrt{2}\,t_{n-1}\quad\text{for } n\ge 1.$$

By induction we then have $t_n=s_0\frac{\sqrt{2}^{n+1}-1}{\sqrt{2}-1}$, and since $\sqrt{2}-1>\frac{1}{4}\sqrt{2}$, we get $t_n\le s_0\sqrt{2}^{n+4}$. For $n\in\mathbb{N}$, we define

$$\alpha_n:=\frac{\mathbb{P}\big(\|X\|_r>t_n\big)}{\mathbb{P}\big(\|X\|_r\le s_0\big)}.$$

By construction we have $\alpha_0\le(1-\frac{2}{3})/\frac{2}{3}=\frac{1}{2}$, and by step **(1)** (with $s=s_0$ and $t=t_{n+1}$) we obtain

$$\alpha_{n+1}=\frac{\mathbb{P}\big(\|X\|_r>t_{n+1}\big)}{\mathbb{P}\big(\|X\|_r\le s_0\big)}\le\frac{\mathbb{P}\big(\|X\|_r>\frac{t_{n+1}-s_0}{\sqrt{2}}\big)^2}{\mathbb{P}\big(\|X\|_r\le s_0\big)^2}=\frac{\mathbb{P}\big(\|X\|_r>t_n\big)^2}{\mathbb{P}\big(\|X\|_r\le s_0\big)^2}=\alpha_n^2.$$

It follows that

$$\mathbb{P}\big(\|X\|_r>t_n\big)=\alpha_n\mathbb{P}\big(\|X\|_r\le s_0\big)\le\alpha_0^{2^n}\le 2^{-2^n}.$$

Using these estimates, we obtain

$$\begin{aligned}\mathbb{E}\exp\big(\beta\|X\|_r^2\big) &\le \mathbb{P}\big(\|X\|_r \le t_0\big)\exp(\beta t_0^2)\\ &\quad + \sum_{n=0}^{\infty}\mathbb{P}\big(t_n < \|X\|_r \le t_{n+1}\big)\exp(\beta t_{n+1}^2)\\ &\le \exp(\beta s_0^2) + \sum_{n=0}^{\infty} 2^{-2^n}\exp(\beta s_0^2\, 2^{n+5})\\ &= \exp(\beta s_0^2) + \sum_{n=0}^{\infty}\exp\big(2^n(-\log 2 + 32\beta s_0^2)\big) =: C.\end{aligned}$$

Now take $\beta < \frac{1}{32 s_0^2}\log 2$ and apply Cauchy's root test to see that $C < \infty$. ■

Remark 1.8. Let $1 \le p < \infty$, and take $\alpha := 1 + \log(\lceil p\rceil !)$. By an elementary computation we have $x^p \le e^{\alpha x^2}$ for every $x \ge 0$. As a consequence of Theorem 1.7, we thus get

$$\sqrt{\frac{\beta}{\alpha}}^{\,p}\,\mathbb{E}\|X\|_r^p \le \mathbb{E}\exp\big(\beta\,\|X\|_r^2\big) < \infty,$$

and hence

$$\mathbb{E}\|X\|_r^p < \infty,$$

where the bounding constant does not depend on r. ■

As a simple corollary to Fernique's theorem, we get the following result.

Corollary 1.9. *If X is an $L^r(U)$-valued Gaussian random variable, then we have $\mathbb{E}X = 0$.*

Proof. By Remark 1.8 we have $\mathbb{E}\|X\|_r < \infty$, and therefore

$$\langle \mathbb{E}X, g\rangle = \mathbb{E}\langle X, g\rangle = 0 \quad \text{for all } g \in L^{r'}(U).$$

The claim now follows by the Hahn-Banach theorem. ■

Example 1.10. Let X be an $L^r(U)$-valued random variable of the form $X = \sum_{n=1}^{\infty}\gamma_n x_n$, where $(\gamma_n)_{n=1}^{\infty}$ is a sequence of independent Gaussian variables and $(x_n)_{n=1}^{\infty}$ is a (finite or infinite) sequence in $L^r(U)$. Obviously, $X_N := \sum_{n=1}^{N}\gamma_n x_n$ is a Gaussian random variable for each $N \in \mathbb{N}$, and the sequence $(X_N)_{N=1}^{\infty}$ satisfies

$$\lim_{N\to\infty}\langle X_N, g\rangle = \langle X, g\rangle \quad \text{almost surely for all } g \in L^{r'}(U).$$

Hence, by Proposition A.17, X is a Gaussian random variable. ■

This example shows that every (convergent) Gaussian sum is a Gaussian random variable. Next, we turn our attention to the question if every $L^r(U)$-valued Gaussian random variable can be represented as a Gaussian sum of the form $\sum_{n=1}^{\infty} \gamma_n x_n$. The next example shows that this is true in the case $L^r(U) = \mathbb{R}^N$.

Example 1.11. The theory of $\mathbb{R}^N$-valued Gaussian random variables and the notions we will use in this example can be found in many introducing works about stochastic calculus, for example in [9].

Assume that X is an $\mathbb{R}^N$-valued Gaussian random variable with positive-definite covariance matrix $\operatorname{Cov}(X) \in \mathbb{R}^{N\times N}$. Since $\operatorname{Cov}(X)$ is symmetric, there exists an orthonormal basis $\{v_1, \ldots, v_N\} \subseteq \mathbb{R}^N$ of eigenvectors of $\operatorname{Cov}(X)$ with corresponding eigenvalues $\lambda_1, \ldots, \lambda_N > 0$. Set

$$A := (v_1|\ldots|v_N)\cdot \operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_N}).$$

Then A is invertible and $\operatorname{Cov}(X) = A\,A^T$. Now set $Y := A^{-1}X$. Then Y is a Gaussian random variable with $\operatorname{Cov}(Y) = I_d$, which means that the components of Y are independent standard Gaussian variables. Write $Y = (\gamma_1, \ldots, \gamma_N)$ for a Gaussian sequence $(\gamma_n)_{n=1}^N$. Then we obtain the following representation of X

$$X = A\,Y = \sum_{n=1}^{N} \gamma_n \sqrt{\lambda_n} v_n. \qquad \blacksquare$$

The next result shows that the desired estimate is also true in arbitrary L^r spaces.

Theorem 1.12 (Karhunen-Loève). *Let X be an $L^r(U)$-valued random variable. Then X is Gaussian if and only if it has the form*

$$X = \sum_{n=1}^{\infty} \gamma_n x_n,$$

where convergence holds almost surely and in $L^p(\Omega; L^r(U))$ for all $1 \le p < \infty$. Here, $(\gamma_n)_{n=1}^{\infty}$ is a sequence of independent standard Gaussian variables and $(x_n)_{n=1}^{\infty} \subseteq L^r(U)$. If this is the case, we further have

$$\big(\mathbb{E}\|X\|_r^p\big)^{\frac{1}{p}} \eqsim_{r,p} \Big\| \Big(\sum_{n=1}^{\infty} |x_n|^2\Big)^{\frac{1}{2}} \Big\|_r < \infty$$

for one (or all) $1 \le p < \infty$.

For the proof of this theorem we need the following convergence result.

Proposition 1.13. *Fix* $1 \le p < \infty$. *Let* $(X_n)_{n=1}^{\infty}$ *be a sequence of independent, integrable, and centered* $L^r(U)$*-valued random variables. Put* $S_N := \sum_{n=1}^{N} X_n$, *and let* $S \in L^p(\Omega; L^r(U))$ *satisfying* $\lim_{N\to\infty} \langle S_N, g\rangle = \langle S, g\rangle$ *in* $L^1(\Omega)$ *for all* $g \in L^{r'}(U)$. *Then we have*

$$\lim_{N\to\infty} S_N = S \quad \text{almost surely and in } L^p(\Omega; L^r(U)).$$

Proof. Since $\lim_{N\to\infty} \langle S_N, g\rangle = \langle S, g\rangle$ in $L^1(\Omega)$ for all $g \in L^{r'}(U)$, we have

$$\lim_{N\to\infty} \Big| \int_B \langle S_N, g\rangle - \langle S, g\rangle \, \mathrm{d}\mathbb{P} \Big| \le \lim_{N\to\infty} \int_\Omega \big| \langle S_N, g\rangle - \langle S, g\rangle \big| \, \mathrm{d}\mathbb{P} = 0$$

for every $B \in \mathscr{A}$. Next, define $\mathscr{F}_N := \sigma(X_1, \ldots, X_N) \subseteq \mathscr{A}$. Then every S_N is $\mathscr{F}_N$-measurable, and X_n is independent of $\mathscr{F}_M$ whenever $n > M$. Hence, using (A.2),(A.5), and the fact that $\mathbb{E}X_n = 0$ for each n, we get for $N > M$

$$\mathbb{E}[S_N | \mathscr{F}_M] = \sum_{n=1}^{M} X_n + \sum_{n=M+1}^{N} \mathbb{E}X_n = \sum_{n=1}^{M} X_n = S_M.$$

So, $(S_N)_{N=1}^{\infty}$ is a martingale with respect to $(\mathscr{F}_N)_{N=1}^{\infty}$. Now fix a $K \in \mathbb{N}$. Then, by the foregoing convergence result, we obtain

$$\begin{aligned} \Big\langle \int_B S \, \mathrm{d}\mathbb{P}, g \Big\rangle &= \int_B \langle S, g\rangle \, \mathrm{d}\mathbb{P} = \lim_{N\to\infty} \int_B \langle S_N, g\rangle \, \mathrm{d}\mathbb{P} \\ &= \lim_{N\to\infty} \Big\langle \int_B \mathbb{E}[S_N | \mathscr{F}_K] \, \mathrm{d}\mathbb{P}, g \Big\rangle \\ &= \Big\langle \int_B S_K \, \mathrm{d}\mathbb{P}, g \Big\rangle \end{aligned}$$

for all $g \in L^{r'}(U)$ and every $B \in \mathscr{F}_K$. By the Hahn-Banach theorem and Theorem A.21 we almost surely have $S_K = \mathbb{E}[S | \mathscr{F}_K]$ for all $K \in \mathbb{N}$. Since $S \in L^p(\Omega; L^r(U))$, the martingale convergence theorem yield

$$\lim_{N\to\infty} S_N = \lim_{N\to\infty} \mathbb{E}[S | \mathscr{F}_N] = \mathbb{E}[S | \mathscr{F}_\infty]$$

almost surely and in $L^p(\Omega; L^r(U))$. Moreover, we have for all $g \in L^{r'}(U)$

$$\big\langle \mathbb{E}[S | \mathscr{F}_\infty], g \big\rangle = \lim_{N\to\infty} \langle S_N, g\rangle = \langle S, g\rangle \quad \text{almost surely,}$$

and hence $\mathbb{E}[S | \mathscr{F}_\infty] = S$ almost surely by Corollary A.8. ■

Proof (of Theorem 1.12). If $X = \sum_{n=1}^{\infty} \gamma_n x_n$, then Example 1.10 implies that X is Gaussian.

For the converse direction, we define

$$G_X := \overline{\big\{\langle X, g\rangle : g \in L^{r'}(U)\big\}}^{L^2(\Omega)} \subseteq L^2(\Omega).$$

Then $\big(G_X, \|\cdot\|_{L^2(\Omega)}\big)$ is a (separable) Hilbert space and, by Proposition A.17, every random variable in G_X is Gaussian. Choose an orthonormal basis $(\gamma_n)_{n=1}^{\infty}$ of G_X, and observe that the orthogonality of $(\gamma_n)_{n=1}^{\infty}$ in G_X implies stochastic independence (cf. Proposition A.16). Next, consider the linear mapping

$$\Phi_X \colon G_X \to L^r(U), \quad \Phi_X(\gamma) = \mathbb{E}\gamma X,$$

which is well-defined and bounded. In fact, by applying Hölder's inequality and Fernique's theorem, we obtain

$$\|\Phi_X(\gamma)\|_r \leq \mathbb{E}|\gamma|\|X\|_r \leq \big(\mathbb{E}|\gamma|^2\big)^{\frac{1}{2}}\big(\mathbb{E}\|X\|_r^2\big)^{\frac{1}{2}} \lesssim_X \big(\mathbb{E}|\gamma|^2\big)^{\frac{1}{2}}.$$

Now define $x_n := \Phi_X(\gamma_n)$, and set $X_N := \sum_{n=1}^{N} \gamma_n x_n$ for $N \in \mathbb{N}$. We then get for all $g \in L^{r'}(U)$

$$\lim_{N\to\infty} \langle X_N, g\rangle = \lim_{N\to\infty} \sum_{n=1}^{N} \gamma_n \langle \Phi_X(\gamma_n), g\rangle = \lim_{N\to\infty} \sum_{n=1}^{N} \gamma_n \mathbb{E}\gamma_n \langle X, g\rangle = \langle X, g\rangle \text{ in } G_X.$$

By Fernique's theorem we have $X \in L^p(\Omega; L^r(U))$ for each $1 \leq p < \infty$. Therefore, Proposition 1.13 implies that

$$\lim_{N\to\infty} X_N = X \qquad \text{almost surely and in } L^p(\Omega; L^r(U))$$

for all $1 \leq p < \infty$. Finally, by the Kahane inequality and Fernique's theorem,

$$\big(\mathbb{E}\|X\|_r^p\big)^{\frac{1}{p}} = \Big(\mathbb{E}\Big\|\sum_{n=1}^{\infty} \gamma_n x_n\Big\|_r^p\Big)^{\frac{1}{p}} \eqsim_{p,r} \Big\|\Big(\sum_{n=1}^{\infty} |x_n|^2\Big)^{\frac{1}{2}}\Big\|_r < \infty. \qquad \blacksquare$$

As a consequence of Theorem 1.12, we get the following corollary.

Corollary 1.14 (Kahane inequality for Gaussian random variables). *Let X be an $L^r(U)$-valued Gaussian random variable. Then we have for all $1 \leq p, q < \infty$*

$$\big(\mathbb{E}\|X\|_r^p\big)^{\frac{1}{p}} \eqsim_{p,q,r} \big(\mathbb{E}\|X\|_r^q\big)^{\frac{1}{q}}.$$

As an application of this corollary, we want to show that if a sequence of Gaussian random variables converges in probability, then it converges in L^p for all $1 \le p < \infty$. For this purpose we need the following lemma.

Lemma 1.15 (Paley-Zygmund inequality). *Let ξ be a non-negative random variable satisfying*

$$0 < \mathbb{E}\xi^2 \le c\big(\mathbb{E}\xi\big)^2 < \infty$$

for some $c > 0$. Then, for all $0 < r < 1$ we have

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

Proof. Since ξ is non-negative, we have

$$(1-r)\mathbb{E}\xi = \mathbb{E}(\xi - r\,\mathbb{E}\xi) \le \mathbb{E}\big(\mathbb{1}_{\{\xi > r\mathbb{E}\xi\}}(\xi - r\,\mathbb{E}\xi)\big) \le \mathbb{E}(\mathbb{1}_{\{\xi > r\mathbb{E}\xi\}}\xi).$$

Therefore, using the Cauchy-Schwarz inequality, we get

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

Finally, dividing both sides by $\mathbb{E}\xi^2$, we obtain

$$\mathbb{P}(\xi > r\,\mathbb{E}\xi) \ge (1-r)^2\frac{(\mathbb{E}\xi)^2}{\mathbb{E}\xi^2} \ge \frac{(1-r)^2}{c}. \qquad \blacksquare$$

Theorem 1.16. *For a sequence $(X_n)_{n=1}^\infty$ of $L^r(U)$-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 $1 \le p < \infty$, the sequence $(X_n)_{n=1}^\infty$ converges in $L^p(\Omega; L^r(U))$ to a random variable X;

(3) for all $1 \le p < \infty$, the sequence $(X_n)_{n=1}^\infty$ converges in $L^p(\Omega; L^r(U))$ to a random variable X.

In this case, the random variable X is Gaussian.

Proof. Since the applications (3) $\Rightarrow$ (2) $\Rightarrow$ (1) are clear, we only have to prove that $\lim_{n\to\infty} X_n = X$ in probability implies $\lim_{n\to\infty} X_n = X$ in $L^p(\Omega; L^r(U))$ for all $1 \le p < \infty$. Note that X is Gaussian by Proposition A.17.

(1) Fix $1 \le q < \infty$. By Fernique's theorem we have $\mathbb{E}\|X_n\|_r^q < \infty$ for all $n \in \mathbb{N}$. We now want to show that even

$$\sup_{n\in\mathbb{N}} \mathbb{E}\|X_n\|_r^q < \infty.$$

By Corollary 1.14 (with $p = 4$ and $q = 2$), there exists a constant $c := c(r, 4, 2) > 0$ such that

$$0 < \mathbb{E}\|X_n\|_r^4 \le c^4 \big(\mathbb{E}\|X_n\|_r^2\big)^2 \quad \text{for all } n \in \mathbb{N}.$$

Therefore, we can apply the Paley-Zygmund inequality and obtain

$$\mathbb{P}\big(\|X_n\|_r^2 > \tfrac{1}{2}\mathbb{E}\|X_n\|_r^2\big) \ge \frac{1}{4\,c^4}. \tag{$\triangle$}$$

Now let $\varepsilon > 0$ be arbitrary. Since $\lim_{n\to\infty} X_n = X$ in probability, we can find for any $t > 0$ an index $N \in \mathbb{N}$ such that for all $n \ge N$ we have

$$\begin{aligned}\mathbb{P}\big(\|X_n\|_r^2 > t\big) &\le \mathbb{P}\big(\|X\|_r > \tfrac{1}{2}\sqrt{t}\,\big) + \mathbb{P}\big(\|X_n - X\|_r > \tfrac{1}{2}\sqrt{t}\,\big)\\ &\le \mathbb{P}\big(\|X\|_r > \tfrac{1}{2}\sqrt{t}\,\big) + \varepsilon,\end{aligned}$$

where we used that

$$\big\{\|X\|_r \le \tfrac{1}{2}\sqrt{t},\ \|X_n - X\|_r \le \tfrac{1}{2}\sqrt{t}\,\big\} \subseteq \big\{\|X_n\|_r^2 \le t\big\}.$$

Hence, for $t_0 > 0$ large enough, we find an index $N_0 > 0$ such that for all $n \ge N_0$

$$\mathbb{P}\big(\|X_n\|_r^2 > t_0\big) < 2\varepsilon.$$

Next, assume that there exists a subsequence such that $\lim_{k\to\infty} \mathbb{E}\|X_{n_k}\|_r^2 = \infty$. Then, for all sufficiently large k, we obtain

$$\mathbb{P}\big(\|X_{n_k}\|_r^2 > \tfrac{1}{2}\mathbb{E}\|X_{n_k}\|_r^2\big) \le \mathbb{P}\big(\|X_{n_k}\|_r^2 > t_0\big) < 2\varepsilon,$$

which contradicts ($\triangle$). Therefore, $\sup_{n\in\mathbb{N}} \mathbb{E}\,\|X_n\|_r^2 < \infty$. Using Corollary 1.14 again, we finally get

$$\sup_{n\in\mathbb{N}} \mathbb{E}\|X_n\|_r^q \lesssim_{r,q,2} \sup_{n\in\mathbb{N}} \mathbb{E}\|X_n\|_r^2 < \infty.$$

(2) Fix $1 \le p < q < \infty$. By step **(1)** and Fernique's theorem, there exists a constant $C > 0$ such that

$$\sup_{n\in\mathbb{N}} \big(\mathbb{E}\|X_n - X\|_r^q\big)^{\frac{1}{q}} \le \sup_{n\in\mathbb{N}} \big(\mathbb{E}\|X_n\|_r^q\big)^{\frac{1}{q}} + \big(\mathbb{E}\|X\|_r^q\big)^{\frac{1}{q}} \le C.$$

Let $\varepsilon > 0$ be fixed. Using Hölders inequality (with $\frac{1}{p} = \frac{1}{q} + \frac{1}{s}$), we obtain

$$\begin{aligned}
\mathbb{E}\|X_n - X\|_r^p &= \mathbb{E}\big(\mathbb{1}_{\{\|X_n - X\|_r \le \varepsilon\}}\|X_n - X\|_r^p\big) + \mathbb{E}\big(\mathbb{1}_{\{\|X_n - X\|_r > \varepsilon\}}\|X_n - X\|_r^p\big) \\
&\le \varepsilon^p + \mathbb{E}\big(\mathbb{1}_{\{\|X_n - X\|_r > \varepsilon\}}\|X_n - X\|_r^p\big) \\
&\le \varepsilon^p + \big(\mathbb{E}\,\mathbb{1}_{\{\|X_n - X\|_r > \varepsilon\}}\big)^{\frac{p}{s}}\big(\mathbb{E}\|X_n - X\|_r^q\big)^{\frac{p}{q}} \\
&\le \varepsilon^p + C^p\,\mathbb{P}\big(\|X_n - X\|_r > \varepsilon\big)^{\frac{p}{s}}.
\end{aligned}$$

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

$$\lim_{n\to\infty} \mathbb{E}\|X_n - X\|_r^p < \varepsilon^p.$$

This being true for all $\varepsilon > 0$, we get $\lim_{n\to\infty} \mathbb{E}\|X_n - X\|_r^p = 0$. ■

1.3 Brownian Motion

In this section we want to give an introduction to the *Brownian motion*, which plays an important role in the theory of stochastic integration. Let E be an arbitrary Banach space. Then an E-valued *stochastic process*, indexed by a set I, is a family of E-valued random variables $\big(X(i)\big)_{i\in I}$ defined on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$.

Definition 1.17. *An* $\mathbb{R}$*-valued process* $\big(\beta(t)\big)_{t\in[0,T]}$ *is called* Brownian motion *(or* Wiener process*) if the following properties hold:*

(1) $\beta(0) = 0$ *almost surely;*

(2) for all $0 \le s \le t \le T$, $\beta(t) - \beta(s)$ *is Gaussian with variance* $t - s$*;*

(3) for all $0 \le s \le t \le T$, $\beta(t) - \beta(s)$ *is independent of* $\{\beta(r) : 0 \le r \le s\}$.

Remark 1.18. The definition of a Brownian motion often include a fourth property, which states that:

(4) the function $t \mapsto \beta(t)$ is almost surely continuous.

As we will see in Remark 1.23, we can drop this property since every Brownian motion with the properties (1)–(3) has a version which even has Hölder continuous trajectories. ■

Remark 1.19. Every Brownian motion $\big(\beta(t)\big)_{t\in[0,T]}$ is a martingale with respect to the filtration $(\mathscr{F}_t^\beta)_{t\in[0,T]}$ defined by

$$\mathscr{F}_t^\beta := \sigma\big(\beta(s) : s \in [0,T]\big).$$

To see this, we first observe that $\big(\beta(t)\big)_{t\in[0,T]}$ is adapted to $(\mathscr{F}_t^\beta)_{t\in[0,T]}$ and the random variables $\beta(t)$ are integrable. So it remains to show that $\mathbb{E}[\beta(t)|\mathscr{F}_s^\beta] = \beta(s)$ almost surely for all $0 \le s \le t \le T$. Since $\beta(s)$ is $\mathscr{F}_s^\beta$-measurable and $\beta(t)-\beta(s)$ is independent of $\mathscr{F}_s^\beta$ (this follows from the definition of the Brownian motion), (A.2) and (A.5) lead to

$$\begin{aligned}\mathbb{E}[\beta(t)|\mathscr{F}_s^\beta] &= \mathbb{E}[\beta(s)|\mathscr{F}_s^\beta] + \mathbb{E}[\beta(t)-\beta(s)|\mathscr{F}_s^\beta] \\ &= \beta(s) + \mathbb{E}\big(\beta(t)-\beta(s)\big) = \beta(s).\end{aligned}$$

■

We next want to investigate the existence of Brownian motions. Therefore, let $(\gamma_n)_{n=1}^\infty$ be a sequence of independent standard Gaussian variables, and let $(g_n)_{n=1}^\infty$ be an orthonormal basis in $L^2([0,T])$. For $n \in \mathbb{N}$ we then define

$$G_n\colon [0,T] \to \mathbb{R}, \quad G_n(t) = \int_0^t g_n \,\mathrm{d}s.$$

For $0 < \gamma < 1$, we denote by $C^\gamma([0,T])$ the space of all γ-Hölder continuous functions, which is a Banach space when endowed with the norm

$$\|f\|_{C^\gamma} := \sup_{s,t\in[0,T]} \frac{|f(t)-f(s)|}{|t-s|^\gamma} + |f(0)|.$$

Theorem 1.20. *The series* $\sum_{n=1}^\infty G_n\gamma_n$ *converges almost surely in* $C^\gamma([0,T])$ *and in* $L^p(\Omega; C^\gamma([0,T]))$ *for* $\gamma < \frac{1}{2}$. *Additionally,* $\big(\sum_{n=1}^\infty G_n(t)\gamma_n\big)_{t\in[0,T]}$ *is a Brownian motion.*

To prove Theorem 1.20, we first have to show the next two propositions.

Proposition 1.21. *Define the* Beta function *by*

$$B\colon (0,\infty)\times(0,\infty) \to \mathbb{R}, \quad B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,\mathrm{d}t.$$

Then we have

$$\Gamma(x)\Gamma(y) = \Gamma(x+y)B(x,y),$$

where Γ *is the Gamma function from Lemma 1.1.*

Proof. Define $\phi\colon (0,\infty)\times(0,1)\to(0,\infty)\times(0,\infty)$ by $\phi(u,v)=\big(uv,u(1-v)\big)$. Then $|\det\phi'(u,v)| = u > 0$, and ϕ is surjective, which means that ϕ is a diffeomorphism. Applying the change of variables formula and Fubini's theorem, we obtain

$$\begin{aligned}\Gamma(x)\Gamma(y) &= \left(\int_0^\infty s^{x-1}e^{-s}\,\mathrm{d}s\right)\left(\int_0^\infty t^{y-1}e^{-t}\,\mathrm{d}t\right)\\ &= \int_{(0,\infty)\times(0,\infty)} s^{x-1}t^{y-1}e^{-x-y}\,\mathrm{d}(s,t)\\ &= \int_{(0,\infty)\times(0,1)} u^{x-1}v^{x-1}u^{y-1}(1-v)^{y-1}e^{-u}u\,\mathrm{d}(u,v)\\ &= \left(\int_0^\infty u^{x+y-1}e^{-u}\,\mathrm{d}u\right)\left(\int_0^1 v^{x-1}\,(1-v)^{y-1}\,\mathrm{d}v\right)\\ &= \Gamma(x+y)B(x,y).\end{aligned}$$

■

Proposition 1.22. *For $0<\alpha\le 1$ and suitable functions $f\colon[0,T]\to\mathbb{R}$ we define the operator D^α by*

$$(D^\alpha f)(t) := [k_\alpha * f](t) = \int_0^t k_\alpha(t-s)f(s)\,\mathrm{d}s, \quad 0\le t\le T,$$

where $k_\alpha(s) := \frac{1}{\Gamma(\alpha)}s^{\alpha-1}$. Then we have:

(1) $D^\alpha D^\beta = D^{\alpha+\beta}$ for $\alpha,\beta>0$ with $0<\alpha+\beta\le 1$;

(2) $D^\alpha\colon L^2([0,T])\to L^r([0,T])$ is continuous for $\frac12<\alpha\le 1$ and $1\le r<\infty$.

Further let Let $0\le\gamma<\frac12$. Then

(3) $D^\beta\colon L^r([0,T])\to C^\gamma([0,T])$ is continuous for $\gamma+\frac1r<\beta\le 1$ and $2<r<\infty$.

Proof. **(1)** Let $\alpha,\beta>0$ with $0<\alpha+\beta\le 1$. Then, with Proposition 1.21,

$$\begin{aligned}(k_\alpha * k_\beta)(t) &= \int_0^t k_\alpha(t-s)k_\beta(s)\,\mathrm{d}s = \frac{1}{\Gamma(\alpha)\Gamma(\beta)}\int_0^t (t-s)^{\alpha-1}s^{\beta-1}\,\mathrm{d}s\\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)}\int_0^t t^{\alpha-1}\left(1-\frac{s}{t}\right)^{\alpha-1}s^{\beta-1}\,\mathrm{d}s\\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)}t^{\alpha+\beta-1}\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\,\mathrm{d}u\\ &= \frac{B(\alpha,\beta)}{\Gamma(\alpha)\Gamma(\beta)}t^{\alpha+\beta-1} = \frac{1}{\Gamma(\alpha+\beta)}t^{\alpha+\beta-1} = k_{\alpha+\beta}(t),\end{aligned}$$

where we used the substitution $u = 1 - \frac{s}{t}$ in the fourth equation. Using this together with the associativity of the convolution, we get

$$D^\alpha D^\beta f = \big(k_\alpha * (k_\beta * f)\big) = \big((k_\alpha * k_\beta) * f\big) = (k_{\alpha+\beta} * f) = D^{\alpha+\beta} f$$

for appropriate functions f.

(2) Now let $\frac{1}{2} < \alpha \le 1$ and $1 \le r < \infty$. Note that D^α is linear and, by the Cauchy-Schwarz inequality,

$$\begin{aligned}
\|D^\alpha f\|_r^r &= \int_0^T \big|(D^\alpha f)(t)\big|^r \,\mathrm{d}t \\
&\le \int_0^T \left(\int_0^t \big|k_\alpha(t-s) f(s)\big| \,\mathrm{d}s\right)^r \mathrm{d}t \\
&\le \int_0^T \left(\int_0^t k_\alpha(t-s)^2 \,\mathrm{d}s\right)^{\frac{r}{2}} \left(\int_0^t f(s)^2 \,\mathrm{d}s\right)^{\frac{r}{2}} \mathrm{d}t \\
&= \int_0^T \left(\frac{1}{\Gamma(\alpha)\sqrt{2\alpha-1}} t^{\alpha-\frac{1}{2}}\right)^r \left(\int_0^t f(s)^2 \,\mathrm{d}s\right)^{\frac{r}{2}} \mathrm{d}t \\
&\le c_1(\alpha, r, T)^r \left(\int_0^T f(s)^2 \,\mathrm{d}s\right)^{\frac{r}{2}} \\
&= c_1(\alpha, r, T)^r \|f\|_2^r,
\end{aligned}$$

where $c_1(\alpha, r, T) = \frac{1}{\Gamma(\alpha)\sqrt{2\alpha-1}} T^{\alpha-\frac{1}{2}+\frac{1}{r}}$. Thus, $D^\alpha\colon L^2([0,T]) \to L^r([0,T])$ is continuous.

(3) Now fix $0 < \gamma < \frac{1}{2}$, and let $2 < r < \infty$ and $\gamma + \frac{1}{r} < \beta \le 1$. Then, for all $0 \le s \le t \le T$,

$$\begin{aligned}
&\big|(D^\beta f)(t) - (D^\beta f)(s)\big| \\
&\quad = \frac{1}{\Gamma(\beta)} \left| \int_0^t (t-u)^{\beta-1} f(u) \,\mathrm{d}u - \int_0^s (s-u)^{\beta-1} f(u) \,\mathrm{d}u \right| \\
&\quad \le \frac{1}{\Gamma(\beta)} \int_0^s \big|\big((t-u)^{\beta-1} - (s-u)^{\beta-1}\big) f(u)\big| \,\mathrm{d}u \\
&\qquad + \frac{1}{\Gamma(\beta)} \int_s^t \big|(t-u)^{\beta-1} f(u)\big| \,\mathrm{d}u.
\end{aligned}$$

Next, we estimate each summand separately. Observe that for all $x, y \ge 0$ and all $q \ge 1$ we have

$$|x-y|^q \le |x^q - y^q|.$$

In fact, for $x > y \ge 0$ we have

$$(x-y)^q = x^q\big(1 - \tfrac{y}{x}\big)^q \le x^q\big(1 - \tfrac{y}{x}\big) = x^q - x^q \tfrac{y}{x} \le x^q - x^q\big(\tfrac{y}{x}\big)^q = x^q - y^q,$$

which gives the desired estimate. Using this together with Hölder's inequality, we obtain

$$\begin{aligned}
&\int_0^s \big|\big((t-u)^{\beta-1}-(s-u)^{\beta-1}\big)f(u)\big|\,\mathrm{d}u \\
&\leq \left(\int_0^s \big|(t-u)^{\beta-1}-(s-u)^{\beta-1}\big|^{r'}\,\mathrm{d}u\right)^{\frac{1}{r'}}\|f\|_r \\
&\leq \left(\int_0^s \big|(t-u)^{(\beta-1)r'}-(s-u)^{(\beta-1)r'}\big|\,\mathrm{d}u\right)^{\frac{1}{r'}}\|f\|_r \\
&= \left(\int_0^s (s-u)^{(\beta-1)r'}-(t-u)^{(\beta-1)r'}\,\mathrm{d}u\right)^{\frac{1}{r'}}\|f\|_r \\
&= \left(\frac{1}{(\beta-1)r'+1}\big((t-s)^{(\beta-1)r'+1}+s^{(\beta-1)r'+1}-t^{(\beta-1)r'+1}\big)\right)^{\frac{1}{r'}}\|f\|_r \\
&\leq \left(\frac{1}{(\beta-1)r'+1}(t-s)^{(\beta-1)r'+1}\right)^{\frac{1}{r'}}\|f\|_r.
\end{aligned}$$

And similarly,

$$\begin{aligned}
\int_s^t \big|(t-u)^{\beta-1}f(u)\big|\,\mathrm{d}u &\leq \left(\int_s^t \big|(t-u)^{\beta-1}\big|^{r'}\,\mathrm{d}u\right)^{\frac{1}{r'}}\|f\|_r \\
&= \left(\int_s^t (t-u)^{(\beta-1)r'}\,\mathrm{d}u\right)^{\frac{1}{r'}}\|f\|_r \\
&= \left(\frac{1}{(\beta-1)r'+1}(t-s)^{(\beta-1)r'+1}\right)^{\frac{1}{r'}}\|f\|_r.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\big|(D^\beta f)(t)-(D^\beta f)(s)\big| &\leq \frac{2}{\Gamma(\beta)}\left(\frac{1}{(\beta-1)r'+1}\right)^{\frac{1}{r'}}(t-s)^{\beta-1+\frac{1}{r'}}\|f\|_r \\
&= \frac{2}{\Gamma(\beta)}\left(\frac{1}{(\beta-1)r'+1}\right)^{\frac{1}{r'}}(t-s)^{\beta-\frac{1}{r}}\|f\|_r.
\end{aligned}$$

Using this estimate and observing that $\beta-\frac{1}{r}-\gamma>0$, we obtain

$$\begin{aligned}
\|D^\beta f\|_{C^\gamma} &= \sup_{s,t\in[0,T]}\frac{|(D^\beta f)(t)-(D^\beta f)(s)|}{|t-s|^\gamma}+|(D^\beta f)(0)| \\
&\leq \frac{2}{\Gamma(\beta)}\left(\frac{1}{(\beta-1)r'+1}\right)^{\frac{1}{r'}}\sup_{s,t\in[0,T]}|t-s|^{\beta-\frac{1}{r}-\gamma}\|f\|_r \\
&\leq c_2(\beta,\gamma,r,T)\|f\|_r,
\end{aligned}$$

with $c_2(\beta,\gamma,r,T):=\frac{2}{\Gamma(\beta)}\big(\frac{1}{(\beta-1)r'+1}\big)^{\frac{1}{r'}}T^{\beta-\frac{1}{r}-\gamma}$. This concludes the proof. ■

Proof (of Theorem 1.20). (1) Let $2 < r < \infty$ such that $\gamma + \frac{1}{r} < \frac{1}{2}$, and choose $\frac{1}{2} < \alpha < 1$ and $\gamma + \frac{1}{r} < \beta < \frac{1}{2}$ such that $\alpha + \beta = 1$. Then, by Proposition 1.22 (1), we have $G_n = D^\beta D^\alpha g_n$, and by Proposition 1.22 (2) and (3) it suffices to prove that $\sum_{n=1}^\infty (D^\alpha g_n)\gamma_n$ converges almost surely in $L^r([0,T])$ and in $L^p(\Omega; L^r([0,T]))$. By the Bessel inequality we obtain for all $t \in [0,T]$

$$\begin{aligned}\Big(\sum_{n=1}^N (D^\alpha g_n)(t)^2\Big)^{\frac{1}{2}} &\le \Big(\sum_{n=1}^\infty \Big(\int_0^T \mathbb{1}_{(0,t)}(s) k_\alpha(t-s) g_n(s)\,\mathrm{d}s\Big)^2\Big)^{\frac{1}{2}} \\ &\le \big\|\mathbb{1}_{(0,t)} k_\alpha(t-\cdot)\big\|_2 = \frac{1}{\Gamma(\alpha)}\Big(\int_0^t s^{2(\alpha-1)}\,\mathrm{d}s\Big)^{\frac{1}{2}} \\ &= \frac{1}{\Gamma(\alpha)\sqrt{2\alpha-1}} t^{\alpha-\frac{1}{2}},\end{aligned}$$

which is integrable. Moreover, this estimate shows that $\sum_{n=1}^\infty (D^\alpha g_n)(t)^2$ converges absolutely for each $t \in [0,T]$. By the Kahane inequality and the dominated convergence theorem, we thus obtain

$$\lim_{N,M\to\infty}\Big(\mathbb{E}\Big\|\sum_{n=N}^M (D^\alpha g_n)\gamma_n\Big\|_r^r\Big)^{\frac{1}{r}} \eqsim_r \lim_{N,M\to\infty}\Big\|\Big(\sum_{n=N}^M (D^\alpha g_n)^2\Big)^{\frac{1}{2}}\Big\|_r = 0,$$

which implies that $\sum_{n=1}^\infty (D^\alpha g_n)\gamma_n$ converges in $L^r(\Omega; L^r([0,T]))$. Combining Theorem 1.12 and Theorem 1.16, we infer that $\sum_{n=1}^\infty (D^\alpha g_n)\gamma_n$ converges almost surely in $L^r([0,T])$ and in $L^p(\Omega; L^r([0,T]))$ for all $1 \le p < \infty$.

(2) By step **(1)**, $\big(X(t)\big)_{t\in[0,T]}$ defined by $X(t) := \sum_{n=1}^\infty G_n(t)\gamma_n$ is an $\mathbb{R}$-valued stochastic process, which satisfies $X(0) = 0$ almost surely. Fix $0 \le s \le t \le T$. Using the dominated convergence theorem, the L^2-orthogonality of the Gaussian sequence $(\gamma_n)_{n=1}^\infty$, and the Parseval identity, we obtain

$$\begin{aligned}\mathbb{E}\big(X(s)X(t)\big) &= \sum_{n=1}^\infty\sum_{m=1}^\infty G_n(s)G_m(t)\mathbb{E}(\gamma_n\gamma_m) = \sum_{n=1}^\infty G_n(s)G_n(t) \\ &= \sum_{n=1}^\infty\Big(\int_0^T \mathbb{1}_{(0,s)}(u)g_n(u)\,\mathrm{d}u\Big)\Big(\int_0^T \mathbb{1}_{(0,t)}(u)g_n(u)\,\mathrm{d}u\Big) \\ &= \int_0^T \mathbb{1}_{(0,s)}(u)\mathbb{1}_{(0,t)}(u)\,\mathrm{d}u = s.\end{aligned}$$

By Theorem 1.12, $X(t) - X(s)$ is Gaussian with variance

$$\mathbb{E}\big(X(t)-X(s)\big)^2 = \mathbb{E}X(t)^2 - 2\mathbb{E}\big(X(s)X(t)\big) + \mathbb{E}X(s)^2 = t - 2s + s = t - s.$$

Next, we want to show that $X(t) - X(s)$ is independent of $\big(X(r_1),\dots,X(r_N)\big)$ whenever $0 \le r_1 \le \ldots \le r_N \le s$. For this it suffices to prove that the random variables $X(r_1), X(r_2) - X(r_1), \dots, X(r_N) - X(r_{N-1}), X(t) - X(s)$ are independent. By construction and Theorem 1.12,

$$\sum_{n=1}^{N} a_n\big(X(r_n) - X(r_{n-1})\big) + a_{N+1}\big(X(t) - X(s)\big)$$

is Gaussian for all sequences $(a_n)_{n=1}^{N+1}$ (where we put $r_0 := 0$). By Proposition A.16 we therefore obtain the desired result if we show their orthogonality in $L^2(\Omega)$. As above, we compute

$$\begin{aligned}&\mathbb{E}\big[\big(X(t) - X(s)\big)\big(X(r_n) - X(r_{n-1})\big)\big]\\ &\quad= \mathbb{E}X(t)X(r_n) - \mathbb{E}X(t)X(r_{n-1}) - \mathbb{E}X(s)X(r_n) + \mathbb{E}X(s)X(r_{n-1})\\ &\quad= r_n - r_{n-1} - r_n + r_{n-1} = 0\end{aligned}$$

for all $n = 1, \dots, N$. Similarly, we get for all $1 \le m < n \le N$

$$\mathbb{E}\big[\big(X(r_n) - X(r_{n-1})\big)\big(X(r_m) - X(r_{m-1})\big)\big] = 0.$$

Thus, by definition, $\big(X(t)\big)_{t\in[0,T]}$ is a Brownian motion. ■

Remark 1.23. Assume that $\big(\beta(t)\big)_{t\in[0,T]}$ is a Brownian motion. By Fubini's theorem, we have

$$\mathbb{E}\int_0^T \beta(t)^2\,\mathrm{d}t = \int_0^T \mathbb{E}\beta(t)^2\,\mathrm{d}t = \int_0^T t\,\mathrm{d}t = \frac{1}{2}T^2 < \infty.$$

Therefore, $\beta \in L^2([0,T])$ almost surely. Moreover,

$$\widetilde{C}^\gamma([0,T]) := \big\{f \in L^2([0,T]) : \exists\, \widetilde{f} \in f \text{ with } \widetilde{f} \in C^\gamma([0,T])\big\} \in \mathscr{B}\big(L^2([0,T])\big).$$

Let us prove this. We define for $m \in \mathbb{N}$,

$$U_m := \{f \in L^2([0,T]) : \exists\, \widetilde{f} \in f \text{ with } \|\widetilde{f}\|_{C^\gamma} \le m\}.$$

Now fix an $m \in \mathbb{N}$ and let $(f_n)_{n=1}^\infty \subseteq U_m$ converge to some $f \in L^2([0,T])$. Then there exists a subsequence $(f_{n_k})_{k=1}^\infty$ that converges pointwise almost everywhere to f. Therefore,

$$\big|f(t) - f(s)\big| = \lim_{k\to\infty}\big|f_{n_k}(t) - f_{n_k}(s)\big| \le m|t-s|^\gamma \quad \text{almost everywhere.}$$

The continuous extension $\widetilde{f}$ of f then satisfies

$$\big|\widetilde{f}(t) - \widetilde{f}(s)\big| \le m|t-s|^\gamma \quad \text{for all } s, t \in [0,T].$$

We infer that $f \in U_m$, and this implies that U_m is closed in $L^2([0,T])$. Thus, we have

$$\widetilde{C}^\gamma([0,T]) = \bigcup_{m\in\mathbb{N}} U_m \in \mathscr{B}\big(L^2([0,T])\big).$$

Let $\big(X(t)\big)_{t\in[0,T]}$ be the Brownian motion constructed in Theorem 1.20. Then the random variables β and X have the same distribution and the Borel measures $\mathbb{P}_\beta$ and $\mathbb{P}_X$ agree on $\mathscr{B}\big(L^2([0,T])\big)$. Hence, by Theorem 1.20,

$$\mathbb{P}\big(\beta \in \widetilde{C}^\gamma([0,T])\big) = \mathbb{P}\big(X \in \widetilde{C}^\gamma([0,T])\big) = 1,$$

which means that $\big(\beta(t)\big)_{t\in[0,T]}$ has a version with γ-Hölder continuous trajectories for any exponent $\gamma < \frac{1}{2}$. ■

Chapter 2

Martingale Inequalities

This chapter is devoted to the study of the UMD property and a new version of Doob's martingale inequality. Both proves are based on a *good-λ inequality* for which it will be necessary to run a reduction process first. The purpose of this process is to verify that it is sufficient to prove the estimates only for a special class of martingales. Subsequently, we will apply these results to obtain a stronger version of the Burkholder-Gundy inequality as well as a Decoupling theorem. The latter one will play an important role in the construction of a stochastic integral for processes. But first of all, we take a look at *martingale difference sequences.*

In this chapter we may always assume that $1 < r < \infty$ is fixed.

2.1 Martingale Difference Sequences

We start with the definition.

Definition 2.1. *Let $(M_n)_{n=1}^N$ be an $L^r(U)$-valued martingale. The sequence $(d_n)_{n=1}^N$ defined by $d_n := M_n - M_{n-1}$ (with $M_0 = 0$) is called the* martingale difference sequence *associated with $(M_n)_{n=1}^N$. Additionally, we call $(d_n)_{n=1}^N$ an L^p* martingale difference sequence *if it is the difference sequence of an L^p martingale.*

Remark 2.2. (1) 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 $(\mathscr{F}_n)_{n=1}^N$.

(2) For $1 \leq m < n \leq N$, we have

$$\mathbb{E}[d_n|\mathscr{F}_m] = \mathbb{E}[M_n|\mathscr{F}_m] - \mathbb{E}[M_{n-1}|\mathscr{F}_m] = M_m - M_m = 0. \qquad \blacksquare$$

The next proposition shows that the properties of Remark 2.2 already give a characterization of martingale difference sequences.

Proposition 2.3. *Let $(d_n)_{n=1}^N$ be a sequence of integrable $L^r(U)$-valued random variables satisfying the properties of Remark 2.2. Then $(M_n)_{n=1}^N$ defined by $M_n := \sum_{i=1}^n d_i, 1 \leq n \leq N$, is a martingale with respect to $(\mathscr{F}_n)_{n=1}^N$, and $(d_n)_{n=1}^N$ is the martingale difference sequence associated with $(M_n)_{n=1}^N$.*

Proof. Let $1 \leq m < n \leq N$ be fixed. Since $d_1, \dots, d_n$ are $\mathscr{F}_n$-measurable and integrable, also $M_n = \sum_{i=1}^n d_i$ is $\mathscr{F}_n$-measurable and integrable. Using (A.2) and the fact that $\mathbb{E}[d_i|\mathscr{F}_m] = 0$ for all $m < i \leq N$, we obtain

$$\mathbb{E}[M_n|\mathscr{F}_m] = \sum_{i=1}^n \mathbb{E}[d_i|\mathscr{F}_m] = \sum_{i=1}^m \mathbb{E}[d_i|\mathscr{F}_m] = \sum_{i=1}^m d_i = M_m.$$

Hence, $(M_n)_{n=1}^N$ is a martingale with respect to the filtration $(\mathscr{F}_n)_{n=1}^N$ and, clearly, $M_n - M_{n-1} = d_n$ for each $n = 1, \dots, N$. ■

This fact allows us to talk about martingale difference sequences without mentioning the associated martingale. Next, we will take a closer look on the case $r = 2$, i.e., the Hilbert space case.

Proposition 2.4. *Let $(d_n)_{n=1}^N$ be an $L^2(U)$-valued L^2 martingale difference sequence, then the following assertions hold:*

(1) $\mathbb{E} \int_U d_n d_m \, \mathrm{d}\mu = 0$ *for all* $n \neq m$*;*

(2) $\mathbb{E}\big\|\sum_{n=1}^N \varepsilon_n d_n\big\|_2^2 \leq \mathbb{E}\big\|\sum_{n=1}^N d_n\big\|_2^2$ *for all sequences* $(\varepsilon_n)_{n=1}^N \subseteq [-1, 1]$.

Proof. (1) By Lemma A.26, $\big(d_n(u)\big)_{n=1}^N$ is an L^2 martingale difference sequence with respect to $(\mathscr{F}_n)_{n=1}^N$ for all $u \in U$ outside a μ-null set U_0. For each fixed $u \in U \setminus U_0$ and $1 \leq m < n \leq N$, we obtain

$$\mathbb{E} d_n(u) d_m(u) = \mathbb{E}\big(\mathbb{E}[d_n(u) d_m(u)|\mathscr{F}_{n-1}]\big) = \mathbb{E}\big(d_m(u)\mathbb{E}[d_n(u)|\mathscr{F}_{n-1}]\big) = 0,$$

where we used (A.1), (A.4) and the $\mathscr{F}_{n-1}$-measurability of $d_m(u)$, as well as Remark 2.2 (2). Applying Fubini's Theorem now yield

$$\mathbb{E} \int_U d_n d_m \, \mathrm{d}\mu = \int_U \mathbb{E} d_n d_m \, \mathrm{d}\mu = 0.$$

(2) Using **(1)** and Fubini's theorem, we obtain

$$\mathbb{E}\Big\| \sum_{n=1}^{N} d_n \Big\|_2^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} \mathbb{E} \int_U d_n d_m \,\mathrm{d}\mu = \sum_{n=1}^{N} \mathbb{E} \int_U d_n^2 \,\mathrm{d}\mu,$$

which finally leads to

$$\mathbb{E}\Big\| \sum_{n=1}^{N} \varepsilon_n d_n \Big\|_2^2 = \sum_{n=1}^{N} \varepsilon_n^2 \mathbb{E} \int_U d_n^2 \,\mathrm{d}\mu \le \sum_{n=1}^{N} \mathbb{E} \int_U d_n^2 \,\mathrm{d}\mu = \mathbb{E}\Big\| \sum_{n=1}^{N} d_n \Big\|_2^2. \qquad \blacksquare$$

2.2 The Strong Doob and Strong Burkholder-Gundy Inequality

Motivated by Proposition 2.4 in the previous section and Doob's martingale inequality (cf. Theorem A.24), we are going to show the following estimates:

Theorem 2.5 (UMD[1] property). *Let $1 < p < \infty$ and $(\varepsilon_n)_{n=1}^N \subseteq [-1,1]$. Then we have for all $L^r(U)$-valued L^p martingale difference sequences $(d_n)_{n=1}^N$*

$$\mathbb{E}\Big\| \sum_{n=1}^{N} \varepsilon_n d_n \Big\|_r^p \lesssim_{p,r} \mathbb{E}\Big\| \sum_{n=1}^{N} d_n \Big\|_r^p. \tag{UMD}$$

Theorem 2.6 (Strong Doob inequality). *Let $1 < p < \infty$. Then we have for all $L^r(U)$-valued L^p martingales $(M_n)_{n=1}^N$*

$$\mathbb{E}\Big\| \max_{n=1}^{N} |M_n| \Big\|_r^p \lesssim_{p,r} \mathbb{E}\|M_N\|_r^p. \tag{SD}$$

Remark 2.7. Note that

$$\mathbb{E} \max_{n=1}^{N} \|M_n\|_r^p \le \mathbb{E}\Big\| \max_{n=1}^{N} |M_n| \Big\|_r^p,$$

so the result of Theorem 2.6 is stronger than the 'classical' Doob inequality. ■

The proof of each estimate consists of two parts: a reduction of the problem to so called Haar martingales and then proving those estimates for this class of martingales.

[1] The term 'UMD' is an abbreviation for 'unconditional martingale differences'.

2.2.1 The Reduction Process

We consider the reduction process for both estimates simultaneously. There will be 3 steps:

Step 1: Reduction to divisible probability spaces

A probability space $(\Omega, \mathscr{A}, \mathbb{P})$ is said to be *divisible* if for all $A \in \mathscr{A}$ and $0 < s < 1$ we have $A = A_1 \cup A_2$ with $A_1, A_2 \in \mathscr{A}$ and

$$\mathbb{P}(A_1) = s\mathbb{P}(A), \quad \mathbb{P}(A_2) = (1-s)\mathbb{P}(A).$$

Lemma 2.8. *Let $1 < p < \infty$. If (UMD) or (SD) hold on any divisible probability space, then (UMD) or (SD) hold on an arbitrary probability space, respectively.*

Proof. Let $(M_n)_{n=1}^N$ be an $L^r(U)$-valued L^p martingale with respect to the filtration $(\mathscr{F}_n)_{n=1}^N$ and let $(d_n)_{n=1}^N$ be its martingale difference sequence, defined on an arbitrary probability space $(\Omega, \mathscr{A}, \mathbb{P})$. We set

$$\widetilde{\Omega} := \Omega \times [0,1], \quad \widetilde{\mathscr{A}} := \mathscr{A} \otimes \mathscr{B}([0,1]), \quad \widetilde{\mathbb{P}} := \mathbb{P} \otimes \lambda_{[0,1]}.$$

Then $(\widetilde{\Omega}, \widetilde{\mathscr{A}}, \widetilde{\mathbb{P}})$ is a divisible probability space. To see this, we define for an $\widetilde{A} \in \widetilde{\mathscr{A}}$

$$\phi_{\widetilde{A}} \colon [0,1] \to [0,1], \quad \phi_{\widetilde{A}}(t) = \widetilde{\mathbb{P}}\Big(\widetilde{A} \cap \big(\Omega \times [0,t]\big)\Big).$$

Then, by construction, $\phi_{\widetilde{A}}$ is continuous with $\phi_{\widetilde{A}}(1) = \widetilde{\mathbb{P}}(\widetilde{A})$ and $\phi_{\widetilde{A}}(0) = 0$. Now fix $0 < s < 1$. By the intermediate value theorem, there exists a $t_0 \in [0,1]$ with $\phi_{\widetilde{A}}(t_0) = s\widetilde{\mathbb{P}}(\widetilde{A})$. Thus, by taking

$$\widetilde{A}_1 := \widetilde{A} \cap \big(\Omega \times [0, t_0]\big) \quad \text{and} \quad \widetilde{A}_2 := \widetilde{A} \setminus \widetilde{A}_1,$$

we have $\widetilde{A}_1, \widetilde{A}_2 \in \widetilde{\mathscr{A}}$ satisfying $\widetilde{A}_1 \cup \widetilde{A}_2 = \widetilde{A}$, $\widetilde{\mathbb{P}}(\widetilde{A}_1) = s\widetilde{\mathbb{P}}(\widetilde{A})$, and $\widetilde{\mathbb{P}}(\widetilde{A}_2) = \widetilde{\mathbb{P}}(\widetilde{A}) - \widetilde{\mathbb{P}}(\widetilde{A}_1) = (1-s)\widetilde{\mathbb{P}}(\widetilde{A})$.

Now define $\widetilde{M}_n(\omega, t) := \mathbb{1}_{[0,1]}(t) M_n(\omega)$ for $(\omega, t) \in \widetilde{\Omega}$ and $\widetilde{\mathscr{F}}_n := \mathscr{F}_n \otimes \mathscr{B}([0,1])$. Clearly, $(\widetilde{M}_n)_{n=1}^N$ is integrable and adapted to $(\widetilde{\mathscr{F}}_n)_{n=1}^N$. For $1 \le m < n \le N$, using (A.8), we obtain

$$\begin{aligned}\mathbb{E}[\widetilde{M}_n | \widetilde{\mathscr{F}}_m] &= \mathbb{E}\big[\mathbb{1}_{[0,1]} M_n | \mathscr{F}_m \otimes \mathscr{B}([0,1])\big] = \mathbb{E}\big[\mathbb{1}_{[0,1]} | \mathscr{B}([0,1])\big]\, \mathbb{E}[M_n | \mathscr{F}_m] \\ &= \mathbb{1}_{[0,1]} M_m = \widetilde{M}_m.\end{aligned}$$

Therefore, $(\widetilde{M}_n)_{n=1}^N$ is a martingale with respect to $(\widetilde{\mathscr{F}}_n)_{n=1}^N$. Now, let $(\varepsilon_n)_{n=1}^N \subseteq [-1,1]$. Then, by the assumption, we obtain for $\widetilde{d}_n := \widetilde{M}_n - \widetilde{M}_{n-1}$

$$\mathbb{E}\Big\| \sum_{n=1}^N \varepsilon_n d_n \Big\|_r^p = \widetilde{\mathbb{E}}\Big\| \sum_{n=1}^N \varepsilon_n \widetilde{d}_n \Big\|_r^p \lesssim_{p,r} \widetilde{\mathbb{E}}\Big\| \sum_{n=1}^N \widetilde{d}_n \Big\|_r^p = \mathbb{E}\Big\| \sum_{n=1}^N d_n \Big\|_r^p.$$

And similarly, we get

$$\mathbb{E}\big\|\max_{n=1}^N |M_n|\big\|_r^p = \widetilde{\mathbb{E}}\big\|\max_{n=1}^N |\widetilde{M}_n|\big\|_r^p \lesssim_{p,r} \widetilde{\mathbb{E}}\|\widetilde{M}_N\|_r^p = \mathbb{E}\|M_N\|_r^p. \qquad \blacksquare$$

Step 2: Reduction to dyadic filtrations
Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space, and let $\mathscr{G}$ be a sub-σ-algebra of $\mathscr{A}$. $\mathscr{G}$ is called *dyadic* if it is generated by 2^m disjoint sets of measure 2^{-m} for an integer $m \geq 0$. Accordingly, we call a filtration in $(\Omega, \mathscr{A}, \mathbb{P})$ *dyadic* if each of its constituting σ-algebras is dyadic.

We first need a simple approximation result.

Lemma 2.9. *Let $1 \leq p < \infty$, $\varepsilon > 0$, and f be an $L^r(U)$-valued simple random variable on a divisible probability space $(\Omega, \mathscr{A}, \mathbb{P})$. Let $\mathscr{G}$ be a dyadic sub-σ-algebra of $\mathscr{A}$. Then there exists a dyadic sub-σ-algebra $\mathscr{G} \subseteq \mathscr{H} \subseteq \mathscr{A}$ and an $\mathscr{H}$-measurable simple random variable h satisfying $\big(\mathbb{E}\|h - f\|_r^p\big)^{\frac{1}{p}} < \varepsilon$.*

Proof. (1) Suppose that $\mathscr{G}$ is generated by the 2^m sets $(G_i)_{i=1}^{2^m}$ with $\mathbb{P}(G_i) = 2^{-m}$ for all $i \in \{1, \dots, 2^m\}$. We first prove the lemma for an indicator function $f = \mathbb{1}_A$ with $A \in \mathscr{A}$ fixed. Write $\mathbb{1}_A = \sum_{i=1}^{2^m} \mathbb{1}_{A \cap G_i}$, and, for an arbitrary generating set G_k, let $(b_j^{G_k})_{j=1}^\infty$ be the sequence of digits in the binary expansion of the real number $\mathbb{P}(A \cap G_k)$, i.e.

$$\mathbb{P}(A \cap G_k) = \sum_{j=1}^\infty b_j^{G_k} 2^{-j}.$$

Then, set $A_0^{G_k} := A \cap G_k$ and $B_0^{G_k} := \emptyset$, and for $j \geq 1$, if $b_j^{G_k} = 1$, choose $B_j^{G_k} \subseteq A_{j-1}^{G_k}$ satisfying $B_j^{G_k} \in \mathscr{A}$ and $\mathbb{P}(B_j^{G_k}) = b_j^{G_k} 2^{-j}$. If $b_j^{G_k} = 0$, take $B_j^{G_k} = \emptyset$. Then set $A_j^{G_k} := A_{j-1}^{G_k} \setminus B_j^{G_k}$ and continue. The so constructed sets $(B_j^{G_k})_{j=1}^\infty \subseteq \mathscr{A}$ are disjoint, contained in G_k, and satisfy

$$\begin{aligned}\mathbb{P}\Big((A \cap G_k) \setminus \bigcup_{j=1}^\infty B_j^{G_k}\Big) &= \mathbb{P}(A \cap G_k) - \mathbb{P}\Big(\bigcup_{j=1}^\infty B_j^{G_k}\Big) = \mathbb{P}(A \cap G_k) - \sum_{j=1}^\infty \mathbb{P}(B_j^{G_k}) \\ &= \mathbb{P}(A \cap G_k) - \sum_{j=1}^\infty b_j^{G_k} 2^{-j} = 0.\end{aligned}$$

Let $n_k \geq 1$ be the first integer with

$$\mathbb{P}\Big((A \cap G_k) \setminus \bigcup_{j=1}^{n_k} B_j^{G_k}\Big) \leq \Big(\frac{\varepsilon}{2^m}\Big)^p.$$

For every $1 \leq j \leq n_k$ with $b_j^{G_k} = 1$ we have $\mathbb{P}(B_j^{G_k}) = 2^{-j}$. It follows that we can split G_k into disjoint subsets of measure 2^{-n_k} such that every $B_j^{G_k}$ is a finite union of these subsets.

We now repeat this construction for each of the 2^m generating sets $(G_i)_{i=1}^{2^m}$. Then every G_i is divided into sets of measure 2^{-n_i}. Set $N := \max_{i=1}^{2^m} n_i$, and, if necessary, subdivide the generating sets even further such that every G_i is divided into sets of measure 2^{-N}. Let $\mathscr{H}$ be the σ-algebra generated by the 2^N sets of measure 2^{-N} thus obtained. This σ-algebra is dyadic with $\mathscr{G} \subseteq \mathscr{H}$, and the function

$$h := \sum_{i=1}^{2^m} \sum_{j=1, b_j^{G_i}=1}^{N} \mathbb{1}_{B_j^{G_i}}$$

is $\mathscr{H}$-measurable with

$$\begin{aligned}
\big(\mathbb{E}|f-h|^p\big)^{\frac{1}{p}} &= \Big(\mathbb{E}\Big|\sum_{i=1}^{2^m} \mathbb{1}_{A\cap G_i} - \sum_{i=1}^{2^m} \sum_{j=1, b_j^{G_i}=1}^{N} \mathbb{1}_{B_j^{G_i}}\Big|^p\Big)^{\frac{1}{p}} \\
&= \Big(\mathbb{E}\Big|\sum_{i=1}^{2^m} \mathbb{1}_{(A\cap G_i)\setminus \bigcup_{j=1}^{N} B_j^{G_i}}\Big|^p\Big)^{\frac{1}{p}} \\
&\leq \sum_{i=1}^{2^m} \mathbb{P}\Big((A \cap G_i) \setminus \bigcup_{j=1}^{N} B_j^{G_i}\Big)^{\frac{1}{p}} \\
&\leq \sum_{i=1}^{2^m} \frac{\varepsilon}{2^m} = \varepsilon.
\end{aligned}$$

(2) For the general case, we may assume that at least one x_n is not 0 and define $\widetilde{\varepsilon} := \varepsilon\big(\sum_{n=1}^{N} \|x_n\|_r\big)^{-1}$. By **(1)** we can find for every indicator function $\mathbb{1}_{A_n}$ an $\mathscr{H}$-measurable indicator function $\mathbb{1}_{B_n}$ with $\big(\mathbb{E}|\mathbb{1}_{A_n} - \mathbb{1}_{B_n}|^p\big)^{\frac{1}{p}} < \widetilde{\varepsilon}$. Taking $h := \sum_{n=1}^{N} \mathbb{1}_{B_n} x_n$ finally yield

$$\big(\mathbb{E}\|h-f\|_r^p\big)^{\frac{1}{p}} \leq \sum_{n=1}^{N} \big(\mathbb{E}|\mathbb{1}_{A_n} - \mathbb{1}_{B_n}|^p \|x_n\|_r^p\big)^{\frac{1}{p}} < \sum_{n=1}^{N} \|x_n\|_r \widetilde{\varepsilon} = \varepsilon. \qquad \blacksquare$$

With this preparation we can finish step 2.

Lemma 2.10. *Let $1 < p < \infty$. If (UMD) or (SD) hold on a divisible probability space $(\Omega, \mathscr{A}, \mathbb{P})$ for all $L^r(U)$-valued L^p martingales with respect to a dyadic filtration, then (UMD) or (SD) hold for all $L^r(U)$-valued L^p martingales on an arbitrary probability space, respectively.*

Proof. Let $\varepsilon > 0$ and $(M_n)_{n=1}^N$ be an arbitrary $L^r(U)$-valued L^p martingale with respect to the filtration $(\mathscr{F}_n)_{n=1}^N$ on a divisible probability space $(\Omega, \mathscr{A}, \mathbb{P})$. Let $(d_n)_{n=1}^N$ be its martingale difference sequence. Since every M_n is $\mathscr{F}_n$-measurable and separably valued, we can find an $\mathscr{F}_n$-measurable simple function $s_n \colon \Omega \to L^r(U)$ with $\big(\mathbb{E}\,\|M_n - s_n\|_r^p\big)^{\frac{1}{p}} \leq \frac{\varepsilon}{4N}$, using Proposition A.3. By repeated applications of Lemma 2.9 we can find a sequence of dyadic sub-σ-algebras $(\widetilde{\mathscr{F}_n})_{n=1}^N$ with $\widetilde{\mathscr{F}_0} = \{\emptyset, \Omega\}$ and $\widetilde{\mathscr{F}_{n-1}} \subseteq \widetilde{\mathscr{F}_n} \subseteq \mathscr{F}_n$, and a sequence of $\widetilde{\mathscr{F}_n}$-measurable simple functions $(\widetilde{s}_n)_{n=1}^N$ with $\big(\mathbb{E}\,\|s_n - \widetilde{s}_n\|_r^p\big)^{\frac{1}{p}} \leq \frac{\varepsilon}{4N}$.

For $n \geq 1$, define $\widetilde{M_n} := \mathbb{E}[M_n|\widetilde{\mathscr{F}_n}]$. Then $\widetilde{M_n}$ is $\widetilde{\mathscr{F}_n}$-measurable, integrable, and

$$\begin{aligned}
\mathbb{E}[\widetilde{M_n}|\widetilde{\mathscr{F}_{n-1}}] &= \mathbb{E}\big[\mathbb{E}[M_n|\widetilde{\mathscr{F}_n}]\big|\widetilde{\mathscr{F}_{n-1}}\big] = \mathbb{E}[M_n|\widetilde{\mathscr{F}_{n-1}}] \\
&= \mathbb{E}\big[\mathbb{E}[M_n|\mathscr{F}_{n-1}]\big|\widetilde{\mathscr{F}_{n-1}}\big] = \mathbb{E}[M_{n-1}|\widetilde{\mathscr{F}_{n-1}}] \\
&= \widetilde{M}_{n-1}.
\end{aligned}$$

Therefore, $(\widetilde{M_n})_{n=1}^N$ is a martingale with respect to $(\widetilde{\mathscr{F}_n})_{n=1}^N$. Finally, by the L^p-contractivity of the conditional expectation operator (cf. Theorem A.21), we obtain

$$\begin{aligned}
\big(\mathbb{E}\|M_n - \widetilde{M_n}\|_r^p\big)^{\frac{1}{p}} &\leq \big(\mathbb{E}\|M_n - s_n\|_r^p\big)^{\frac{1}{p}} + \big(\mathbb{E}\|s_n - \widetilde{s}_n\|_r^p\big)^{\frac{1}{p}} + \big(\mathbb{E}\|\widetilde{s}_n - \widetilde{M_n}\|_r^p\big)^{\frac{1}{p}} \\
&\leq \frac{\varepsilon}{2N} + \Big(\mathbb{E}\Big\|\mathbb{E}[\widetilde{s}_n|\widetilde{\mathscr{F}_n}] - \mathbb{E}[M_n|\widetilde{\mathscr{F}_n}]\Big\|_r^p\Big)^{\frac{1}{p}} \\
&= \frac{\varepsilon}{2N} + \Big(\mathbb{E}\Big\|\mathbb{E}[\widetilde{s}_n - M_n|\widetilde{\mathscr{F}_n}]\Big\|_r^p\Big)^{\frac{1}{p}} \\
&\leq \frac{\varepsilon}{2N} + \big(\mathbb{E}\|\widetilde{s}_n - M_n\|_r^p\big)^{\frac{1}{p}} \\
&\leq \frac{\varepsilon}{2N} + \frac{\varepsilon}{2N} = \frac{\varepsilon}{N}.
\end{aligned}$$

Moreover, if $(\widetilde{d}_n)_{n=1}^N$ is the martingale difference sequence associated with $(\widetilde{M_n})_{n=1}^N$, then,

$$\begin{aligned}
\big(\mathbb{E}\|d_n - \widetilde{d}_n\|_r^p\big)^{\frac{1}{p}} &\leq \big(\mathbb{E}\|M_n - \widetilde{M_n}\|_r^p\big)^{\frac{1}{p}} + \big(\mathbb{E}\|M_{n-1} - \widetilde{M}_{n-1}\|_r^p\big)^{\frac{1}{p}} \\
&\leq \frac{2\varepsilon}{N}.
\end{aligned}$$

Finally, by the assumptions and the L^p-contractivity, we get

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^N \varepsilon_n \widetilde{d}_n\Big\|_r^p\Big)^{\frac{1}{p}} \lesssim_{p,r} \Big(\mathbb{E}\Big\|\sum_{n=1}^N \widetilde{d}_n\Big\|_r^p\Big)^{\frac{1}{p}} = \big(\mathbb{E}\|\widetilde{M}_N\|_r^p\big)^{\frac{1}{p}} = \big(\mathbb{E}\big\|\mathbb{E}[M_N|\widetilde{\mathscr{F}}_N]\big\|_r^p\big)^{\frac{1}{p}}$$
$$\leq \big(\mathbb{E}\|M_N\|_r^p\big)^{\frac{1}{p}} = \Big(\mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|_r^p\Big)^{\frac{1}{p}},$$

and similarly

$$\big(\mathbb{E}\big\|\max_{n=1}^N |\widetilde{M}_n|\big\|_r^p\big)^{\frac{1}{p}} \lesssim_{p,r} \big(\mathbb{E}\|\widetilde{M}_N\|_r^p\big)^{\frac{1}{p}} = \big(\mathbb{E}\big\|\mathbb{E}[M_N|\widetilde{\mathscr{F}}_N]\big\|_r^p\big)^{\frac{1}{p}}$$
$$\leq \big(\mathbb{E}\|M_N\|_r^p\big)^{\frac{1}{p}}.$$

This leads to

$$\Big(\mathbb{E}\Big\|\sum_{n=1}^N \varepsilon_n d_n\Big\|_r^p\Big)^{\frac{1}{p}} \leq \Big(\mathbb{E}\Big\|\sum_{n=1}^N \varepsilon_n (d_n - \widetilde{d}_n)\Big\|_r^p\Big)^{\frac{1}{p}} + \Big(\mathbb{E}\Big\|\sum_{n=1}^N \varepsilon_n \widetilde{d}_n\Big\|_r^p\Big)^{\frac{1}{p}}$$
$$\lesssim_{p,r} \sum_{n=1}^N |\varepsilon_n| \big(\mathbb{E}\|d_n - \widetilde{d}_n\|_r^p\big)^{\frac{1}{p}} + \Big(\mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|_r^p\Big)^{\frac{1}{p}}$$
$$\leq 2\varepsilon + \Big(\mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|_r^p\Big)^{\frac{1}{p}},$$

and

$$\big(\mathbb{E}\big\|\max_{n=1}^N |M_n|\big\|_r^p\big)^{\frac{1}{p}} \leq \big(\mathbb{E}\big\|\max_{n=1}^N |M_n - \widetilde{M}_n|\big\|_r^p\big)^{\frac{1}{p}} + \big(\mathbb{E}\big\|\max_{n=1}^N |\widetilde{M}_n|\big\|_r^p\big)^{\frac{1}{p}}$$
$$\lesssim_{p,r} \sum_{n=1}^N \big(\mathbb{E}\|M_n - \widetilde{M}_n\|_r^p\big)^{\frac{1}{p}} + \big(\mathbb{E}\|M_N\|_r^p\big)^{\frac{1}{p}}$$
$$\leq \varepsilon + \big(\mathbb{E}\|M_N\|_r^p\big)^{\frac{1}{p}}.$$

Since $\varepsilon > 0$ was arbitrary, this shows the result together with Lemma 2.8. ■

Step 3: Reduction to Haar filtrations

An *atom* of a σ-algebra $\mathscr{G}$ is a nonempty set $G \in \mathscr{G}$ such that $H \subseteq G$ with $H \in \mathscr{G}$ implies $H \in \{\emptyset, G\}$. In the final step we now want to shrink the class of martingales to *Haar martingales*, which are defined as martingales with respect to a *Haar filtration*. This is a filtration $(\mathscr{F}_n)_{n=1}^N$ where $\mathscr{F}_1 = \{\emptyset, \Omega\}$ and for $n \geq 1$ each $\mathscr{F}_n$ 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. By construction, each $\mathscr{F}_n$ is generated by n atoms of measure 2^{-k-1} or 2^{-k}, where k is the unique integer such that $2^{k-1} < n \leq 2^k$.

Lemma 2.11. *Let $1 < p < \infty$. If (UMD) or (SD) hold on a divisible probability space and for all $L^r(U)$-valued L^p martingales with respect to a Haar filtration, then (UMD) or (SD) hold for all $L^r(U)$-valued L^p martingales on an arbitrary probability space, respectively.*

Proof. Let $(M_n)_{n=1}^N$ be an $L^r(U)$-valued L^p martingale with respect to a dyadic filtration $(\mathscr{F}_n)_{n=1}^N$ on a divisible probability space $(\Omega, \mathscr{A}, \mathbb{P})$, and let $(d_n)_{n=1}^N$ be its martingale difference sequence. We now want to construct a Haar martingale $(\widetilde{M}_k)_{k=1}^K$ in which we can 'embed' $(M_n)_{n=1}^N$.

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 $k_1 < \ldots < k_N$. Now set $k_0 := 1$, $\widetilde{\mathscr{F}}_{k_0} := \{\emptyset, \Omega\}$, and $\widetilde{\mathscr{F}}_{k_n} := \mathscr{F}_n$. The σ-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. Completing this process, we get a Haar filtration $(\widetilde{\mathscr{F}}_k)_{k=1}^K$ with $K = k_N = 2^{l_N}$. Now take

$$\widetilde{M}_{k_n} := M_n \quad \text{and} \quad \widetilde{M}_k := \mathbb{E}[\widetilde{M}_{k_n} | \widetilde{\mathscr{F}}_k] = \mathbb{E}[M_n | \widetilde{\mathscr{F}}_k] \quad \text{if } k_{n-1} < k < k_n.$$

Then $\widetilde{M}_k$ is $\widetilde{\mathscr{F}}_k$-measurable, integrable, and satisfies for $1 \le m < k \le K$

$$\mathbb{E}[\widetilde{M}_k | \widetilde{\mathscr{F}}_m] = \mathbb{E}\big[\mathbb{E}[\widetilde{M}_{k_n} | \widetilde{\mathscr{F}}_k] \big| \widetilde{\mathscr{F}}_m\big] = \mathbb{E}[\widetilde{M}_{k_n} | \widetilde{\mathscr{F}}_m] = \widetilde{M}_m.$$

Therefore, $(\widetilde{M}_k)_{k=1}^K$ is a Haar martingale. Let $(\widetilde{d}_k)_{k=1}^K$ be the martingale difference sequence associated with $(\widetilde{M}_k)_{k=1}^K$. Then,

$$\sum_{j=k_{n-1}+1}^{k_n} \widetilde{d}_j = \widetilde{M}_{k_n} - \widetilde{M}_{k_{n-1}} = M_n - M_{n-1} = d_n.$$

Now let $(\varepsilon_n)_{n=1}^N \subseteq [-1, 1]$ and set $\widetilde{\varepsilon}_k := \varepsilon_n$ for $k = k_{n-1}+1, \ldots, k_n$, $n = 1, \ldots, N$. This yield

$$\begin{aligned}
\mathbb{E}\Big\| \sum_{n=1}^N \varepsilon_n d_n \Big\|_r^p &= \mathbb{E}\Big\| \sum_{n=1}^N \sum_{j=k_{n-1}+1}^{k_n} \varepsilon_n \widetilde{d}_j \Big\|_r^p = \mathbb{E}\Big\| \sum_{k=1}^K \widetilde{\varepsilon}_k \widetilde{d}_k \Big\|_r^p \\
&\lesssim_{p,r} \mathbb{E}\Big\| \sum_{k=1}^K \widetilde{d}_k \Big\|_r^p = \mathbb{E}\Big\| \sum_{n=1}^N \sum_{j=k_{n-1}+1}^{k_n} \widetilde{d}_j \Big\|_r^p \\
&= \mathbb{E}\Big\| \sum_{n=1}^N d_n \Big\|_r^p.
\end{aligned}$$

Finally, observe that $\widetilde{M}_K = \widetilde{M}_{k_N} = M_N$, which gives us

$$\mathbb{E}\big\|\max_{n=1}^{N} |M_n|\big\|_r^p \le \mathbb{E}\big\|\max_{k=1}^{K} |\widetilde{M}_k|\big\|_r^p \lesssim_{p,r} \mathbb{E}\|\widetilde{M}_K\|_r^p = \mathbb{E}\,\|M_N\|_r^p\,.$$

Together with Lemma 2.10, this completes the prove. ■

At the end of this subsection we want to show a special property of Haar martingales.

Lemma 2.12. *Let $(M_n)_{n=1}^N$ be an $L^r(U)$-valued Haar martingale and let $(d_n)_{n=1}^N$ be its martingale difference sequence. Then $\|d_{n+1}\|_r$ is $\mathscr{F}_n$-measurable for all $n = 1, \dots, N-1$.*

Proof. Note that for any $m = 1, \dots, N$, we have

$$\mathscr{F}_m = \sigma(A_1, \dots, A_m) = \Big\{ \bigcup_{j \in M} A_j : M \subseteq \{1, \dots, m\} \Big\}.$$

Fix $k \in \{1, \dots, m\}$, choose an arbitrary $\omega \in A_k$, and set $M_m(\omega) = z$. By the $\mathscr{F}_m$-measurability of M_m we get $M_m^{-1}(\{z\}) \in \mathscr{F}_m$. Therefore, there exists an $A \in \mathscr{F}_m$ with $M_m(A) = \{z\}$, and, by the structure of $\mathscr{F}_m$, it holds $A_k \subseteq A$. From this we infer that M_m is constant on every generating atom of $\mathscr{F}_m$. Now fix an $n \in \{1, \dots, N-1\}$. $\mathscr{F}_{n+1}$ is obtained by splitting one generating atom $B \in \mathscr{F}_n$ into two subsets B_1 and B_2 of equal measure. By the foregoing remark, M_n and M_{n+1} only differ on B. In fact, let C be one of the generating atoms of $\mathscr{F}_n$, unequal to B. Then C is also a generating atom of $\mathscr{F}_{n+1}$, and $M_n = z_1$, $M_{n+1} = z_2$ on C for some z_1, $z_2 \in L^r(U)$. Hence, by Theorem A.21,

$$\mathbb{P}(C)(z_2 - z_1) = \int_C d_{n+1}\,\mathrm{d}\mathbb{P} = \int_C \mathbb{E}[d_{n+1}|\mathscr{F}_n]\,\mathrm{d}\mathbb{P} = 0.$$

Since C has positive measure, this implies $z_1 = z_2$, and hence, M_{n+1} and M_n are equal on each generating atom of $\mathscr{F}_n$, unequal to B.

Also, d_{n+1} is constant on B_1 and B_2 with values x_1 and x_2. Then,

$$\mathbb{P}(B_1)x_1 + \mathbb{P}(B_2)x_2 = \int_B d_{n+1}\,\mathrm{d}\mathbb{P} = \int_B \mathbb{E}[d_{n+1}|\mathscr{F}_n]\,\mathrm{d}\mathbb{P} = 0,$$

and from $\mathbb{P}(B_1) = \mathbb{P}(B_2) > 0$ we deduce that $x_1 = -x_2$. Therefore,

$$\|d_{n+1}\|_r = \mathbb{1}_{B_1}\|x_1\|_r + \mathbb{1}_{B_2}\|x_2\|_r = \mathbb{1}_B\|x_1\|_r$$

is $\mathscr{F}_n$-measurable. ■

2.2.2 The UMD Property

By subsection 2.2.1 it suffices to consider Haar martingales in order to prove Theorem 2.5. Thus, for the rest of this subsection, let $(M_n)_{n=1}^N$ be an $L^r(U)$-valued L^p martingale with respect to the Haar filtration $(\mathscr{F}_n)_{n=1}^N$, and let $(d_n)_{n=1}^N$ be its martingale difference sequence. For a fixed sequence $(\varepsilon_n)_{n=1}^N \subseteq [-1,1]$ we denote by $(g_n)_{n=1}^N$ the martingale transform $g_n := \sum_{j=1}^n \varepsilon_j d_j$, and we define

$$M^*(\omega) := \max_{n=1}^N \|M_n(\omega)\|_r, \quad g^*(\omega) := \max_{n=1}^N \|g_n(\omega)\|_r.$$

Remark 2.13. If $(X_n)_{n=1}^N$ is a sequence of $L^r(U)$-valued random variables and $\tau\colon \Omega \to \{1,\dots,N\}$ is another random variable, then we define

$$X_\tau\colon \Omega \to L^r(U), \quad X_\tau(\omega) := X_{\tau(\omega)}(\omega).$$

Note that X_τ is still measurable since $X_\tau = \sum_{n=1}^N \mathbb{1}_{\{\tau=n\}} X_n$. ■

Lemma 2.14. *Suppose that (UMD) holds for some $1 < q < \infty$. Then we have for all $\delta > 0$, $\beta > 2\delta + 1$, and all $\lambda > 0$*

$$\mathbb{P}(g^* > \beta\lambda,\ M^* \le \delta\lambda) \le \alpha(\delta)^q\, \mathbb{P}(g^* > \lambda),$$

where $\alpha(\delta) := \frac{4\delta c(q,r)}{\beta - 2\delta - 1}$ and $c(q,r)$ is the constant from Theorem 2.5.

Proof. We define

$$\begin{aligned}
\mu(\omega) &:= \min\{1 \le n \le N\colon \|g_n(\omega)\|_r > \lambda\},\\
\nu(\omega) &:= \min\{1 \le n \le N\colon \|g_n(\omega)\|_r > \beta\lambda\} \quad \text{and}\\
\sigma(\omega) &:= \min\{1 \le n \le N\colon \|M_n(\omega)\|_r > \delta\lambda \ \text{ or } \|d_{n+1}(\omega)\|_r > 2\delta\lambda\}
\end{aligned}$$

with the convention that $\min\emptyset := N+1$ and $d_{N+1} := 0$.

It holds that $\{\mu = j\}, \{\nu = j\} \in \mathscr{F}_j$ and, by Lemma 2.12, even $\{\sigma = j\} \in \mathscr{F}_j$ for all $j = 1,\dots,N$. Now define for $n = 1,\dots,N$

$$v_n := \mathbb{1}_{\{\mu < n \le \nu\wedge\sigma\}},$$

and note that

$$\begin{aligned}
\{\mu < n \le \nu\wedge\sigma\} &= \{\mu < n\} \cap \{\nu\wedge\sigma \ge n\}\\
&= \{\mu < n\} \cap \{\nu\wedge\sigma < n\}^C \in \mathscr{F}_{n-1}.
\end{aligned}$$

Therefore, $(v_n)_{n=1}^N$ is predictable with respect to $(\mathscr{F}_n)_{n=1}^N$, and so, by Example A.23,

$$V_n := \sum_{j=1}^n v_j d_j$$

defines a martingale $(V_n)_{n=1}^N$, which is adapted to $(\mathscr{F}_n)_{n=1}^N$.

On the set $\{\sigma \le \mu\}$ we have $v_n = 0$ for all $n = 1, \dots, N$, which means that $V_N = 0$ there. Now let $\omega \in \{\sigma > \mu\}$. Then $\|M_\mu(\omega)\|_r \le \delta\lambda$. Also, if $\nu(\omega) \wedge \sigma(\omega) > 1$, then from $\|M_{(\nu\wedge\sigma)-1}(\omega)\|_r \le \delta\lambda$ and $\|d_{\nu\wedge\sigma}(\omega)\|_r \le 2\delta\lambda$ it follows that

$$\|M_{\nu\wedge\sigma}(\omega)\|_r \le \|M_{(\nu\wedge\sigma)-1}(\omega)\|_r + \|d_{\nu\wedge\sigma}(\omega)\|_r \le 3\delta\lambda.$$

If $\nu(\omega) \wedge \sigma(\omega) = 1$, then, since $\mu(\omega) \ge 1$, we again have $V_N(\omega) = 0$. Hence, on the set $\{\sigma > \mu\}$ we obtain

$$\|V_N\|_r = \Big\| \sum_{\mu < j \le \nu\wedge\sigma} d_j \Big\|_r = \|M_{\nu\wedge\sigma} - M_\mu\|_r \le 4\delta\lambda.$$

Note that $\{g^* \le \lambda\} = \{\mu = N+1\} \subseteq \{\sigma \le \mu\}$, and thus $\{\sigma > \mu\} \subseteq \{g^* > \lambda\}$. We infer that

$$\mathbb{E}\|V_N\|_r^q \le (4\delta\lambda)^q\, \mathbb{P}(\sigma > \mu) \le (4\delta\lambda)^q\, \mathbb{P}(g^* > \lambda).$$

Next, consider the martingale transform $(V_n^\varepsilon)_{n=1}^N$ defined by

$$V_n^\varepsilon := \sum_{j=1}^n \varepsilon_j v_j d_j.$$

If $\omega \in \{\nu \le N,\ \sigma = N+1\}$, we obtain $\nu(\omega)\wedge\sigma(\omega) = \nu(\omega)$ as well as $\|g_\nu(\omega)\|_r > \beta\lambda$ and $\|d_n(\omega)\|_r \le 2\delta\lambda$ for all $n = 2, \dots, N+1$ and $\|d_1(\omega)\|_r = \|M_1(\omega)\|_r \le \delta\lambda$. Therefore, if $\mu(\omega) > 1$,

$$\|g_\mu(\omega)\|_r \le \|g_{\mu-1}(\omega)\|_r + \|d_\mu(\omega)\|_r \le \lambda + 2\delta\lambda.$$

and if $\mu(\omega) = 1$,

$$\|g_\mu(\omega)\|_r = \|g_1(\omega)\|_r \le \|M_1(\omega)\|_r \le \delta\lambda.$$

Hence, on the set $\{\nu \le N,\ \sigma = N+1\}$ we get in any case

$$\|V_N^\varepsilon\|_r = \Big\| \sum_{\mu < j \le \nu} \varepsilon_j d_j \Big\|_r = \|g_\nu - g_\mu\|_r > \beta\lambda - 2\delta\lambda - \lambda.$$

Observe that $\{g^* > \beta\lambda\} = \{\nu \le N\}$ and $\{M^* \le \delta\lambda\} = \{\sigma = N+1\}$. Then Chebyshev's inequality yield

$$\begin{aligned}
\mathbb{P}(g^* > \beta\lambda,\ M^* \le \delta\lambda) &= \mathbb{P}(\nu \le N,\ \sigma = N+1) \\
&\le \mathbb{P}\big(\|V_N^\varepsilon\|_r > \beta\lambda - 2\delta\lambda - \lambda\big) \\
&\le \frac{1}{(\beta\lambda - 2\delta\lambda - \lambda)^q}\mathbb{E}\|V_N^\varepsilon\|_r^q \\
&\le \frac{c(q,r)^q}{(\beta\lambda - 2\delta\lambda - \lambda)^q}\mathbb{E}\|V_N\|_r^q \\
&\le \frac{(4\delta)^q c(q,r)^q}{(\beta - 2\delta - 1)^q}\mathbb{P}(g^* > \lambda).
\end{aligned}$$

■

Theorem 2.15. *If (UMD) holds for some $1 < q < \infty$, then it holds for all $1 < p < \infty$.*

Proof. By the previous subsection and the definitions made before Lemma 2.14, we need to prove the estimate

$$\mathbb{E}\|g_N\|_r^p \le b^p\, \mathbb{E}\|M_N\|_r^p$$

with a constant $b \ge 0$ depending only on p, q and r.

Note that for any $A,\ B \in \mathscr{A}$ we have

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^C) \le \mathbb{P}(A \cap B) + \mathbb{P}(B^C).$$

Now fix a $\beta > 1$, and let $\delta > 0$ be so small such that $\beta > 2\delta + 1$. Then, by Lemma 2.14,

$$\begin{aligned}
\mathbb{P}(g^* > \beta\lambda) &\le \mathbb{P}(g^* > \beta\lambda,\ M^* \le \delta\lambda) + \mathbb{P}(M^* > \delta\lambda) \\
&\le \alpha(\delta)^q \mathbb{P}(g^* > \lambda) + \mathbb{P}(M^* > \delta\lambda).
\end{aligned}$$

Therefore, by Proposition 1.2 and Doob's inequality, we obtain

$$\begin{aligned}
\mathbb{E}\|g_N\|_r^p \le \mathbb{E}|g^*|^p &= \int_0^\infty p\lambda^{p-1}\mathbb{P}(g^* > \lambda)\,\mathrm{d}\lambda = \beta^p \int_0^\infty p\lambda^{p-1}\mathbb{P}(g^* > \beta\lambda)\,\mathrm{d}\lambda \\
&\le \alpha(\delta)^q\beta^p \int_0^\infty p\lambda^{p-1}\mathbb{P}(g^* > \lambda)\,\mathrm{d}\lambda + \beta^p \int_0^\infty p\lambda^{p-1}\mathbb{P}(M^* > \delta\lambda)\,\mathrm{d}\lambda \\
&= \alpha(\delta)^q\beta^p\mathbb{E}|g^*|^p + \frac{\beta^p}{\delta^p}\mathbb{E}|M^*|^p \\
&\le \alpha(\delta)^q\beta^p\Big(\frac{p}{p-1}\Big)^p\mathbb{E}\|g_N\|_r^p + \frac{\beta^p}{\delta^p}\Big(\frac{p}{p-1}\Big)^p\mathbb{E}\|M_N\|_r^p.
\end{aligned}$$

Since $\lim_{\delta\to 0}\alpha(\delta) = 0$, we may arrange that $\alpha(\delta)^q\beta^p\big(\frac{p}{p-1}\big)^p < 1$ by taking $\delta > 0$ small enough. Recalling that $(M_n)_{n=1}^N$ is an L^p martingale, we note that $\mathbb{E}\|g_N\|_r^p < \infty$. Hence,

$$\mathbb{E}\|g_N\|_r^p \le \frac{\beta^p\big(\frac{p}{p-1}\big)^p}{\Big(1-\big(\frac{p}{p-1}\big)^p\alpha(\delta)^q\beta^p\Big)\delta^p}\mathbb{E}\|M_N\|_r^p. \qquad \blacksquare$$

Remark 2.16. By combining Proposition 2.4 and Theorem 2.15, we obtain for all $1 < p < \infty$, all $\mathbb{R}$-valued L^p martingale difference sequences $(d_n)_{n=1}^N$, and all sequences $(\varepsilon_n)_{n=1}^N \subseteq [-1,1]$

$$\mathbb{E}\Big|\sum_{n=1}^N \varepsilon_n d_n\Big|^p \lesssim_p \mathbb{E}\Big|\sum_{n=1}^N d_n\Big|^p. \qquad \blacksquare$$

Now we can prove the main result of this subsection.

Proof (of Theorem 2.5). Let $(d_n)_{n=1}^N$ be an arbitrary $L^r(U)$-valued L^p martingale difference sequence and $(\varepsilon_n)_{n=1}^N \subseteq [-1,1]$. Then, by Lemma A.26, $(d_n(u))_{n=1}^N$ is an $\mathbb{R}$-valued martingale difference sequence for μ-almost all $u\in U$. Thus, by the foregoing remark,

$$\mathbb{E}\Big|\sum_{n=1}^N \varepsilon_n d_n(u)\Big|^r \lesssim_r \mathbb{E}\Big|\sum_{n=1}^N d_n(u)\Big|^r$$

for μ-almost all $u\in U$. Applying Fubini's theorem now yield

$$\mathbb{E}\Big\|\sum_{n=1}^N \varepsilon_n d_n\Big\|_r^r = \int_U \mathbb{E}\Big|\sum_{n=1}^N \varepsilon_n d_n\Big|^r \,\mathrm{d}\mu \lesssim_r \int_U \mathbb{E}\Big|\sum_{n=1}^N d_n\Big|^r \,\mathrm{d}\mu = \mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|_r^r.$$

Hence, (UMD) holds for $1 < r < \infty$, and thanks to Theorem 2.15, the UMD property holds for all $1 < p < \infty$. $\blacksquare$

Remark 2.17. If we restrict the choice of the sequence $(\varepsilon_n)_{n=1}^N$ to $\{+1,-1\}$, then we get even more. If $(d_n)_{n=1}^N$ is an $L^r(U)$-valued L^p martingale difference sequence, then the same is true for $(\varepsilon_n d_n)_{n=1}^N$. Hence, by applying Theorem 2.5, we obtain

$$\mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|_r^p = \mathbb{E}\Big\|\sum_{n=1}^N \varepsilon_n(\varepsilon_n d_n)\Big\|_r^p \lesssim_{p,r} \mathbb{E}\Big\|\sum_{n=1}^N \varepsilon_n d_n\Big\|_r^p.$$

So, in this case we get

$$\mathbb{E}\Big\|\sum_{n=1}^N \varepsilon_n d_n\Big\|_r^p \eqsim_{p,r} \mathbb{E}\Big\|\sum_{n=1}^N d_n\Big\|_r^p. \qquad \blacksquare$$

2.2.3 The Strong Doob Inequality

By the reduction process of subsection 2.2.1 it is enough to consider Haar martingales to prove Theorem 2.6. In what follows, we let $(M_n)_{n=1}^N$ be an $L^r(U)$-valued L^p martingale with respect to a Haar filtration $(\mathscr{F}_n)_{n=1}^N$ and with difference sequence $(d_n)_{n=1}^N$. Then we define

$$M^*(\omega) := \max_{n=1}^N \|M_n(\omega)\|_r, \quad \widetilde{M}^*(\omega) := \big\|\max_{n=1}^N |M_n(\omega)|\big\|_r.$$

Lemma 2.18. *Let $(V_n)_{n=1}^N$ be an $L^r(U)$-valued L^r martingale. Then we have for all $\lambda > 0$*

$$\mathbb{P}\Big(\big\|\max_{n=1}^N |V_n|\big\|_r > \lambda\Big) \le \frac{(\frac{r}{r-1})^r}{\lambda^r}\mathbb{E}\|V_N\|_r^r.$$

Proof. By Lemma A.26, $\big(V_n(u)\big)_{n=1}^N$ is an $\mathbb{R}$-valued L^r martingale for μ-almost all $u \in U$. Thus, by Doob's martingale inequality, we obtain

$$\mathbb{E}\max_{n=1}^N |V_n(u)|^r \le \Big(\frac{r}{r-1}\Big)^r \mathbb{E}|V_N(u)|^r$$

for μ-almost all $u \in U$. Define

$$\widetilde{V}^* := \big\|\max_{n=1}^N |V_n|\big\|_r.$$

Then Fubini's theorem yield

$$\begin{aligned}
\lambda^r \mathbb{P}\big(\widetilde{V}^* > \lambda\big) &\le \mathbb{E}\big(\mathbb{1}_{\{\widetilde{V}^*>\lambda\}}\widetilde{V}^{*r}\big) \le \mathbb{E}\widetilde{V}^{*r} \\
&= \mathbb{E}\Big(\int_U \max_{n=1}^N |V_n|^r \,\mathrm{d}\mu\Big) = \int_U \Big(\mathbb{E}\max_{n=1}^N |V_n|^r\Big)\,\mathrm{d}\mu \\
&\le \Big(\frac{r}{r-1}\Big)^r \int_U \mathbb{E}|V_N|^r \,\mathrm{d}\mu = \Big(\frac{r}{r-1}\Big)^r \mathbb{E}\int_U |V_N|^r \,\mathrm{d}\mu \\
&= \Big(\frac{r}{r-1}\Big)^r \mathbb{E}\|V_N\|_r^r.
\end{aligned}$$

■

Lemma 2.19. *For all $\delta > 0$, $\beta > 2\delta + 1$, and all $\lambda > 0$ we have*

$$\mathbb{P}\big(\widetilde{M}^* > \beta\lambda,\ M^* \le \delta\lambda\big) \le \alpha(\delta)^r \mathbb{P}\big(\widetilde{M}^* > \lambda\big),$$

where $\alpha(\delta) := \frac{r}{r-1}\frac{4\delta}{\beta-2\delta-1}$.

Proof. We define

$$\begin{aligned}\mu(\omega) &:= \min\Big\{1 \le n \le N\colon \big\|\max_{j=1}^{n} |M_j(\omega)|\big\|_r > \lambda\Big\},\\ \nu(\omega) &:= \min\Big\{1 \le n \le N\colon \big\|\max_{j=1}^{n} |M_j(\omega)|\big\|_r > \beta\lambda\Big\} \quad \text{and}\\ \sigma(\omega) &:= \min\Big\{1 \le n \le N\colon \|M_n(\omega)\|_r > \delta\lambda \ \text{ or } \|d_{n+1}(\omega)\|_r > 2\delta\lambda\Big\}\end{aligned}$$

with the understanding that $\min\emptyset := N+1$ and $d_{N+1} := 0$. Since $\beta > 1$, we have $\mu \le \nu$. Observe that $\{\mu = j\}$, $\{\nu = j\} \in \mathscr{F}_j$ and, by Lemma 2.12, also $\{\sigma = j\} \in \mathscr{F}_j$ for all $j = 1, \dots, N$. For $n = 1, \dots, N$ we set

$$v_n := \mathbb{1}_{\{\mu < n \le \nu\wedge\sigma\}}.$$

Since

$$\begin{aligned}\{\mu < n \le \nu \wedge \sigma\} &= \{\mu < n\} \cap \{\nu \wedge \sigma \ge n\}\\ &= \{\mu < n\} \cap \{\nu \wedge \sigma < n\}^C \in \mathscr{F}_{n-1},\end{aligned}$$

$(v_n)_{n=1}^N$ is an $(\mathscr{F}_n)_{n=1}^N$-predictable sequence of bounded random variables. Therefore, by Example A.23,

$$V_n := \sum_{j=1}^{n} v_j d_j$$

defines a martingale $(V_n)_{n=1}^N$, which is adapted to $(\mathscr{F}_n)_{n=1}^N$.

On the set $\{\sigma \le \mu\}$ we have $v_n = 0$ for all $n = 1, \dots, N$, which means that $V_N = 0$ there. Now let $\omega \in \{\sigma > \mu\}$. Then $\|M_\mu(\omega)\|_r \le \delta\lambda$. Also, if $\nu(\omega) \wedge \sigma(\omega) > 1$, then from

$$\|M_{(\nu\wedge\sigma)-1}(\omega)\|_r \le \delta\lambda \quad \text{and} \quad \|d_{\nu\wedge\sigma}(\omega)\|_r \le 2\delta\lambda$$

it follows that

$$\|M_{\nu\wedge\sigma}(\omega)\|_r \le \|M_{(\nu\wedge\sigma)-1}(\omega)\|_r + \|d_{\nu\wedge\sigma}(\omega)\|_r \le 3\delta\lambda.$$

If $\nu(\omega) \wedge \sigma(\omega) = 1$, then, since $\mu(\omega) \ge 1$, we again have $V_N(\omega) = 0$. Hence, on the set $\{\sigma > \mu\}$ we obtain

$$\|V_N\|_r = \Big\| \sum_{\mu < j \le \nu\wedge\sigma} d_j \Big\|_r = \|M_{\nu\wedge\sigma} - M_\mu\|_r \le 4\delta\lambda.$$

Note that $\{\widetilde{M}^* \le \lambda\} = \{\mu = N+1\} \subseteq \{\sigma \le \mu\}$, and thus $\{\sigma > \mu\} \subseteq \{\widetilde{M}^* > \lambda\}$. We deduce that

$$\mathbb{E}\|V_N\|_r^r \le (4\delta\lambda)^r\, \mathbb{P}(\sigma > \mu) \le (4\delta\lambda)^r\, \mathbb{P}\big(\widetilde{M}^* > \lambda\big).$$

On the set $\{\nu \le N,\ \sigma = N+1\}$ we have $\nu \wedge \sigma = \nu$, and therefore

$$V_n = \sum_{\mu < j \le \nu \wedge n} M_j - M_{j-1} = \begin{cases} 0, & \text{if } 1 \le n \le \mu, \\ M_n - M_\mu, & \text{if } \mu < n \le \nu, \\ M_\nu - M_\mu, & \text{if } n > \nu. \end{cases}$$

We infer that

$$\big\|\max_{n=1}^{N} |V_n|\big\|_r = \big\|\max_{\mu < n \le \nu} |M_n - M_\mu|\big\|_r.$$

Now fix a $u \in U$. Then

$$\max_{1 \le n \le \nu} |M_n(u)| \le \max_{1 \le n \le \mu} |M_n(u)| + \max_{\mu < n \le \nu} |M_n(u)|,$$

and

$$\max_{1 \le n \le \mu} |M_n(u)| \le \max_{1 \le n < \mu} |M_n(u)| + |M_\mu(u)|.$$

Using these inequalities, we obtain

$$\begin{aligned} \max_{\mu < n \le \nu} |M_n(u) - M_\mu(u)| &\ge \max_{\mu < n \le \nu} |M_n(u)| - |M_\mu(u)| \\ &\ge \max_{1 \le n \le \nu} |M_n(u)| - \max_{1 \le n \le \mu} |M_n(u)| - |M_\mu(u)| \\ &\ge \max_{1 \le n \le \nu} |M_n(u)| - \max_{1 \le n < \mu} |M_n(u)| - 2|M_\mu(u)|. \end{aligned}$$

It follows that

$$\big\|\max_{n=1}^{N} |V_n|\big\|_r \ge \big\|\max_{1 \le n \le \nu} |M_n|\big\|_r - \big\|\max_{1 \le n < \mu} |M_n|\big\|_r - 2\|M_\mu\|_r.$$

Next, observe that $\sigma = N+1$ implies that $\|M_n\|_r \le \delta\lambda$ for all $n = 1, \dots, N$, and from $\nu \le N$ it follows that $\big\|\max_{1 \le n \le \nu} |M_n|\big\|_r > \beta\lambda$ and $\big\|\max_{1 \le n < \mu} |M_n|\big\|_r \le \lambda$. Therefore, we have

$$\big\|\max_{n=1}^{N} |V_n|\big\|_r > \beta\lambda - \lambda - 2\delta\lambda.$$

Observe that $\{\widetilde{M}^* > \beta\lambda\} = \{\nu \le N\}$ and $\{M^* \le \delta\lambda\} = \{\sigma = N+1\}$. Putting all these estimates together and using Lemma 2.18, we get

$$\begin{aligned} \mathbb{P}\big(\widetilde{M}^* > \beta\lambda,\ M^* \le \delta\lambda\big) &= \mathbb{P}(\nu \le N,\ \sigma = N+1) \\ &\le \mathbb{P}\Big(\big\|\max_{n=1}^{N} |V_n|\big\|_r > \beta\lambda - 2\delta\lambda - \lambda\Big) \\ &\le \frac{(\frac{r}{r-1})^r}{(\beta\lambda - 2\delta\lambda - \lambda)^r} \mathbb{E}\|V_N\|_r^r \\ &\le \frac{(4\delta)^r (\frac{r}{r-1})^r}{(\beta - 2\delta - 1)^r} \mathbb{P}\big(\widetilde{M}^* > \lambda\big). \end{aligned}$$

■

Proof (of Theorem 2.6). By subsection 2.2.1 and the definitions we made before Lemma 2.18, we have to show that

$$\mathbb{E}|\widetilde{M}^*|^p \leq c^p\, \mathbb{E}\|M_N\|_r^p$$

with a constant $c \geq 0$ depending only on p and r.

Note that for any A, $B \in \mathscr{A}$ we have

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^C) \leq \mathbb{P}(A \cap B) + \mathbb{P}(B^C).$$

Now fix a $\beta > 1$, and let $\delta > 0$ be so small such that $\beta > 2\delta + 1$. Then, by Lemma 2.19, we have

$$\begin{aligned}\mathbb{P}\big(\widetilde{M}^* > \beta\lambda\big) &\leq \mathbb{P}\big(\widetilde{M}^* > \beta\lambda,\, M^* \leq \delta\lambda\big) + \mathbb{P}(M^* > \delta\lambda)\\ &\leq \alpha(\delta)^r \mathbb{P}\big(\widetilde{M}^* > \lambda\big) + \mathbb{P}(M^* > \delta\lambda).\end{aligned}$$

Thus, by Proposition 1.2 and Doob's martingale inequality, we obtain

$$\begin{aligned}\mathbb{E}|\widetilde{M}^*|^p &= \int_0^\infty p\lambda^{p-1}\mathbb{P}\big(\widetilde{M}^* > \lambda\big)\,\mathrm{d}\lambda\\ &= \beta^p \int_0^\infty p\lambda^{p-1}\mathbb{P}\big(\widetilde{M}^* > \beta\lambda\big)\,\mathrm{d}\lambda\\ &\leq \alpha(\delta)^r\beta^p \int_0^\infty p\lambda^{p-1}\mathbb{P}\big(\widetilde{M}^* > \lambda\big)\,\mathrm{d}\lambda + \beta^p \int_0^\infty p\lambda^{p-1}\mathbb{P}(M^* > \delta\lambda)\,\mathrm{d}\lambda\\ &= \alpha(\delta)^r\beta^p\mathbb{E}|\widetilde{M}^*|^p + \frac{\beta^p}{\delta^p}\mathbb{E}\,|M^*|^p\\ &\leq \alpha(\delta)^r\beta^p\mathbb{E}|\widetilde{M}^*|^p + \frac{\beta^p}{\delta^p}\Big(\frac{p}{p-1}\Big)^p\mathbb{E}\,\|M_N\|_r^p\,.\end{aligned}$$

Since $\lim_{\delta\to 0}\alpha(\delta) = 0$, we may take $\delta > 0$ so small such that $\alpha(\delta)^r\beta^p < 1$. By recalling that $(M_n)_{n=1}^N$ is an L^p martingale, we note that $\mathbb{E}|\widetilde{M}^*|^p < \infty$. Hence we get

$$\mathbb{E}|\widetilde{M}^*|^p \leq \frac{\beta^p\big(\frac{p}{p-1}\big)^p}{(1-\alpha(\delta)^r\beta^p)\,\delta^p}\mathbb{E}\,\|M_N\|_r^p\,. \qquad \blacksquare$$

Remark 2.20. Note that $\|M_N\|_r \leq \big\|\max_{n=1}^N |M_n|\big\|_r$ almost surely. Therefore, we have

$$\mathbb{E}\|M_N\|_r^p \leq \mathbb{E}\big\|\max_{n=1}^N |M_n|\big\|_r^p,$$

and together with Theorem 2.6 this gives

$$\mathbb{E}\big\|\max_{n=1}^N |M_n|\big\|_r^p \eqsim_{p,r} \mathbb{E}\|M_N\|_r^p. \qquad \blacksquare$$

2.2.4 The Strong Burkholder-Gundy Inequality

Theorem 2.21. *Let $1 < p < \infty$ and $(d_n)_{n=1}^N$ be an $L^r(U)$-valued L^p martingale difference sequence. Then we have*

$$\mathbb{E}\Big\| \sum_{n=1}^N d_n \Big\|_r^p \eqsim_{p,r} \mathbb{E}\Big\| \Big(\sum_{n=1}^N |d_n|^2\Big)^{\frac{1}{2}} \Big\|_r^p.$$

Proof. Let $(\Omega', \mathscr{A}', \mathbb{P}')$ be another probability space and $(r_n')_{n=1}^N$ be a sequence of independent Rademacher variables on $(\Omega', \mathscr{A}', \mathbb{P}')$. Then, by the Kahane inequality, we get for each fixed $\omega \in \Omega$

$$\Big\| \Big(\sum_{n=1}^N |d_n(\omega)|^2\Big)^{\frac{1}{2}} \Big\|_r^p \eqsim_{p,r} \mathbb{E}'\Big\| \sum_{n=1}^N r_n' d_n(\omega) \Big\|_r^p.$$

Note that $\big(r_n'(\omega')\big)_{n=1}^N \subseteq \{+1, -1\}$ for all $\omega' \in \Omega'$. Therefore, by Fubini's theorem and Remark 2.17,

$$\begin{aligned}
\mathbb{E}\Big\| \Big(\sum_{n=1}^N |d_n|^2\Big)^{\frac{1}{2}} \Big\|_r^p &\eqsim_{p,r} \mathbb{E}\mathbb{E}'\Big\| \sum_{n=1}^N r_n' d_n \Big\|_r^p = \mathbb{E}'\mathbb{E}\Big\| \sum_{n=1}^N r_n' d_n \Big\|_r^p \\
&\eqsim_{p,r} \mathbb{E}'\mathbb{E}\Big\| \sum_{n=1}^N d_n \Big\|_r^p \\
&= \mathbb{E}\Big\| \sum_{n=1}^N d_n \Big\|_r^p.
\end{aligned}$$

■

Theorem 2.22 (Strong Burkholder-Gundy inequality). *Let $1 < p < \infty$ and $(M_n)_{n=1}^N$ be an $L^r(U)$-valued L^p martingale with difference sequence $(d_n)_{n=1}^N$. Then we have*

$$\mathbb{E}\big\|\max_{n=1}^N |M_n|\big\|_r^p \eqsim_{p,r} \mathbb{E}\Big\| \Big(\sum_{n=1}^N |d_n|^2\Big)^{\frac{1}{2}} \Big\|_r^p.$$

Proof. By combining Remark 2.20 and Theorem 2.21, we get

$$\mathbb{E}\big\|\max_{n=1}^N |M_n|\big\|_r^p \eqsim_{p,r} \mathbb{E}\|M_N\|_r^p = \mathbb{E}\Big\| \sum_{n=1}^N d_n \Big\|_r^p \eqsim_{p,r} \mathbb{E}\Big\| \Big(\sum_{n=1}^N |d_n|^2\Big)^{\frac{1}{2}} \Big\|_r^p.$$

■

2.3 Decoupling

This section is an application of Theorem 2.5 we presented in Subsection 2.2.2 and will play a central role in Chapter 3.

Let $1 < p, r < \infty$ be fixed, $(\mathscr{F}_n)_{n=1}$ be a filtration on $(\Omega, \mathscr{A}, \mathbb{P})$, and $(\xi_n)_{n=1}^N$ be a sequence of centered integrable random variables in $L^p(\Omega)$. Assume that

(1) ξ_n is $\mathscr{F}_n$-measurable for all $1 \le n \le N$, and

(2) ξ_n is independent of $\mathscr{F}_m$ for all $1 \le m < n \le N$.

Note that $\mathbb{E}\left[\xi_n | \mathscr{F}_m\right] = \mathbb{E}\xi_n = 0$. Hence, by Proposition 2.3, $(\xi_n)_{n=1}^N$ is a martingale difference sequence with respect to $(\mathscr{F}_n)_{n=1}^N$.

On the product space $(\Omega \times \Omega, \mathscr{A} \otimes \mathscr{A}, \mathbb{P} \otimes \mathbb{P})$ we define

$$\xi_n(\omega, \widetilde{\omega}) := \xi_n(\omega), \quad \widetilde{\xi}_n(\omega, \widetilde{\omega}) := \xi_n(\widetilde{\omega}).$$

The sequences $(\xi_n)_{n=1}^N$ and $(\widetilde{\xi}_n)_{n=1}^N$ are independent, identically distributed, and 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

$$\mathscr{F}_n := \mathscr{F}_n \otimes \{\emptyset, \Omega\}, \quad \widetilde{\mathscr{F}}_n := \{\emptyset, \Omega\} \otimes \mathscr{F}_n.$$

Let $(v_n)_{n=1}^N$ be an $(\mathscr{F}_n)_{n=1}^N$-predictable sequence of $L^r(U)$-valued random variables. In the same way as above, we identify $(v_n)_{n=1}^N$ with a predictable sequence $(v_n)_{n=1}^N$ on $\Omega \times \Omega$ by taking $v_n(\omega, \widetilde{\omega}) := v_n(\omega)$.

Theorem 2.23 (Decoupling). *Under the assumptions taken above, we have*

$$\mathbb{E}\widetilde{\mathbb{E}} \Big\| \sum_{n=1}^N \xi_n v_n \Big\|_r^p \eqsim_{p,r} \mathbb{E}\widetilde{\mathbb{E}} \Big\| \sum_{n=1}^N \widetilde{\xi}_n v_n \Big\|_r^p.$$

Proof. For $n = 1, \dots, N$, define

$$d_{2n-1} := \tfrac{1}{2}(\xi_n + \widetilde{\xi}_n) v_n \quad \text{and} \quad d_{2n} := \tfrac{1}{2}(\xi_n - \widetilde{\xi}_n) v_n,$$

as well as

$$\mathscr{G}_{2n-1} := \sigma(\mathscr{F}_{n-1}, \widetilde{\mathscr{F}}_{n-1}, \xi_n + \widetilde{\xi}_n) \quad \text{and} \quad \mathscr{G}_{2n} := \sigma(\mathscr{F}_n, \widetilde{\mathscr{F}}_n).$$

Then, $(\mathscr{G}_j)_{j=1}^{2N}$ is a filtration and d_j is $\mathscr{G}_j$-measurable for all $j = 1, \dots, 2N$. Note that the independence of ξ_n and $\mathscr{F}_{n-1}$ implies the independence of $\sigma(\xi_n + \widetilde{\xi}_n)$ and

$\sigma(\mathscr{F}_{n-1}, \widetilde{\mathscr{F}}_{n-1})$. Now, using the $\mathscr{F}_{n-1}$-measurability of v_n, (A.7), and (A.9), we obtain

$$\mathbb{E}[d_{2n}|\mathscr{G}_{2n-1}] = \tfrac{1}{2}v_n\mathbb{E}[\xi_n - \widetilde{\xi}_n|\mathscr{G}_{2n-1}] = \tfrac{1}{2}v_n\mathbb{E}[\xi_n - \widetilde{\xi}_n|\xi_n + \widetilde{\xi}_n] = 0.$$

Moreover,

$$\mathbb{E}[d_{2n-1}|\mathscr{G}_{2n-2}] = \tfrac{1}{2}v_n\mathbb{E}[\xi_n + \widetilde{\xi}_n|\mathscr{F}_{n-1}, \widetilde{\mathscr{F}}_{n-1}] = \tfrac{1}{2}v_n\mathbb{E}(\xi_n + \widetilde{\xi}_n) = 0,$$

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 centered. Therefore, by Proposition 2.3, $(d_j)_{j=1}^{2N}$ is a martingale difference sequence with respect to the filtration $(\mathscr{G}_j)_{j=1}^{2N}$.

Now observe that

$$\sum_{j=1}^{2N} d_j = \sum_{n=1}^{N} d_{2n-1} + d_{2n} = \sum_{n=1}^{N} \xi_n v_n,$$

and

$$\sum_{j=1}^{2N} (-1)^{j+1} d_j = \sum_{n=1}^{N} d_{2n-1} - d_{2n} = \sum_{n=1}^{N} \widetilde{\xi}_n v_n.$$

Finally, by Remark 2.17, we obtain

$$\begin{aligned}\mathbb{E}\widetilde{\mathbb{E}}\Big\| \sum_{n=1}^{N} \xi_n v_n \Big\|_r^p &= \mathbb{E}\widetilde{\mathbb{E}}\Big\| \sum_{j=1}^{2N} d_j \Big\|_r^p \eqsim_{p,r} \mathbb{E}\widetilde{\mathbb{E}}\Big\| \sum_{j=1}^{2N} (-1)^{j+1} d_j \Big\|_r^p \\ &= \mathbb{E}\widetilde{\mathbb{E}}\Big\| \sum_{n=1}^{N} \widetilde{\xi}_n v_n \Big\|_r^p. \end{aligned}$$ ■

As a combination of the Decoupling theorem and the Kahane inequality, we get the following corollary.

Corollary 2.24. *Let $\big(\beta(t)\big)_{t\in[0,T]}$ be an arbitrary Brownian motion, and let $0 = t_0 < \ldots < t_N = T$ be fixed. Let $(\mathscr{F}_n)_{n=1}^N$ be a filtration which satisfies*

(1) $\beta(t_n)$ is $\mathscr{F}_n$-measurable,

(2) $\beta(t_n) - \beta(t_{n-1})$ is independent of $\mathscr{F}_{n-1}$,

and let $(v_n)_{n=1}^N$ be an $(\mathscr{F}_n)_{n=1}^N$-predictable sequence of $L^r(U)$-valued random variables. Then we have

$$\mathbb{E}\Big\| \sum_{n=1}^{N} v_n\big(\beta(t_n) - \beta(t_{n-1})\big) \Big\|_r^p \eqsim_{p,r} \mathbb{E}\Big\| \Big(\sum_{n=1}^{N} |v_n|^2 (t_n - t_{n-1}) \Big)^{\frac{1}{2}} \Big\|_r^p.$$

Proof. For $n = 1, \dots, N$ we define

$$\xi_n := \beta(t_n) - \beta(t_{n-1}) \quad \text{and} \quad \gamma_n := \frac{1}{(t_n - t_{n-1})^{\frac{1}{2}}} \big(\beta(t_n) - \beta(t_{n-1})\big).$$

Then $(\xi_n)_{n=1}^N$ satisfies the assumptions of the Decoupling theorem, and, by the definition of the Brownian motion, $(\gamma_n)_{n=1}^N$ is a sequence of independent standard Gaussian variables. Using the Kahane inequality, the Decoupling theorem and the notions given there, we obtain

$$\begin{aligned}
\mathbb{E}\Big\| \sum_{n=1}^N v_n \big(\beta(t_n) - \beta(t_{n-1})\big) \Big\|_r^p &= \mathbb{E}\widetilde{\mathbb{E}}\Big\| \sum_{n=1}^N v_n \xi_n \Big\|_r^p \\
&\eqsim_{p,r} \mathbb{E}\widetilde{\mathbb{E}}\Big\| \sum_{n=1}^N v_n \widetilde{\xi}_n \Big\|_r^p \\
&= \mathbb{E}\widetilde{\mathbb{E}}\Big\| \sum_{n=1}^N v_n \, (t_n - t_{n-1})^{\frac{1}{2}} \, \widetilde{\gamma}_n \Big\|_r^p \\
&\eqsim_{p,r} \mathbb{E}\Big\| \Big(\sum_{n=1}^N |v_n|^2 \, (t_n - t_{n-1}) \Big)^{\frac{1}{2}} \Big\|_r^p . \qquad \blacksquare
\end{aligned}$$

Remark 2.25. It follows from the definition of the Brownian motion that the filtration $(\mathscr{F}_n^\beta)_{n=1}^N$, given by

$$\mathscr{F}_n^\beta := \sigma\big(\beta(t_m), \ 1 \le m \le n\big),$$

satisfies the properties needed in the previous lemma. See also Remark 1.19. ■

Chapter 3

Stochastic Integration

In this chapter we finally take up the study of stochastic integration in L^r spaces. We first present a stochastic integral for functions $f\colon [0,T] \to L^r(U)$, where we assume that $1 \le r < \infty$. While showing some of its properties, we will observe that these integrals are Gaussian random variables. Using the results of Chapter 2, we extend the first integral to $L^r(U)$-valued stochastic processes, now assuming that $1 < r < \infty$. In section 3.3 we will then see that this integral is a time-continuous martingale. Unfortunately, due to the construction process, we are bound to a strong integrability condition. To get rid of this condition, we will apply a localization argument which allows the stochastic integral to be defined for a much larger class of integrands.

3.1 The Wiener Integral

In this section we are going to define a stochastic integral for suitable functions $f\colon [0,T] \to L^r(U)$, $1 \le r < \infty$, with respect to a Brownian motion $\big(\beta(t)\big)_{t\in[0,T]}$. Such integrals are sometimes called Wiener integrals.

Definition 3.1. *For an $L^r(U)$-valued* step function *of the form*

$$f = \sum_{n=1}^{N} x_n \mathbb{1}_{[t_{n-1},t_n)}\colon [0,T] \to L^r(U),$$

with $x_1, \dots, x_N \in L^r(U)$ and $0 = t_0 < \dots < t_N = T$, we define the random variable $\int_0^T f \,\mathrm{d}\beta\colon \Omega \to L^r(U)$ by

$$\int_0^T f \,\mathrm{d}\beta := \sum_{n=1}^{N} x_n \big(\beta(t_n) - \beta(t_{n-1})\big).$$

Remark 3.2. By this definition, we have $\int_0^T f \,\mathrm{d}\beta \in L^p(\Omega; L^r(U))$ for all $1 \leq p < \infty$, satisfying

$$\mathbb{E}\int_0^T f \,\mathrm{d}\beta = 0$$

for each step function f. ■

In order to extend the stochastic integral to a broader class of functions, we make the following observation.

Lemma 3.3 (Itô isomorphism for step functions). *For all $1 \leq p < \infty$ and all step functions $f\colon [0,T] \to L^r(U)$ we have*

$$\Big(\mathbb{E}\Big\| \int_0^T f \,\mathrm{d}\beta \Big\|_r^p\Big)^{\frac{1}{p}} \eqsim_{p,r} \Big\| \Big(\int_0^T |f|^2 \,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_r. \tag{3.1}$$

Proof. Let $x_1, \ldots, x_N \in L^r(U)$ and $0 = t_0 < \ldots < t_N = T$. Then, for $f := \sum_{n=1}^N x_n \mathbb{1}_{[t_{n-1},t_n)}$, we have

$$\begin{aligned}\int_0^T |f|^2 \,\mathrm{d}t &= \int_0^T \Big| \sum_{n=1}^N x_n \mathbb{1}_{[t_{n-1},t_n)} \Big|^2 \mathrm{d}t = \int_0^T \sum_{n=1}^N |x_n|^2 \mathbb{1}_{[t_{n-1},t_n)} \,\mathrm{d}t \\ &= \sum_{n=1}^N |x_n|^2 (t_n - t_{n-1}),\end{aligned}$$

where we used that the sets of the indicator functions are disjoint.

Now set $\gamma_n := \frac{1}{\sqrt{t_n - t_{n-1}}}\big(\beta(t_n) - \beta(t_{n-1})\big)$, $1 \leq n \leq N$. Then, by the definition of the Brownian motion, $(\gamma_n)_{n=1}^N$ is a sequence of independent standard Gaussian variables. Hence, the Kahane inequality leads to

$$\begin{aligned}\Big(\mathbb{E}\Big\| \int_0^T f \,\mathrm{d}\beta \Big\|_r^p\Big)^{\frac{1}{p}} &= \Big(\mathbb{E}\Big\| \sum_{n=1}^N x_n \big(\beta(t_n) - \beta(t_{n-1})\big) \Big\|_r^p\Big)^{\frac{1}{p}} \\ &= \Big(\mathbb{E}\Big\| \sum_{n=1}^N x_n \sqrt{t_n - t_{n-1}} \gamma_n \Big\|_r^p\Big)^{\frac{1}{p}} \\ &\eqsim_{p,r} \Big\| \Big(\sum_{n=1}^N |x_n|^2 (t_n - t_{n-1})\Big)^{\frac{1}{2}} \Big\|_r \\ &= \Big\| \Big(\int_0^T |f|^2 \,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_r.\end{aligned}$$

■

By computation, the norm of the space $L^r(U; L^2([0,T]))$ emerged in the previous lemma, and in what follows, we will see that this space is the correct space of integrands. It is for this reason that we will use the abbreviation

$$\boxed{L^r_{\gamma,T} := L^r(U; L^2([0,T]))}$$

for the rest of this thesis.

Lemma 3.4. *The space of all step functions, given by*

$$D_{r,T} := \Big\{ f = \sum_{n=1}^{N} x_n \mathbb{1}_{[t_{n-1},t_n)} \colon (x_n)_{n=1}^N \subseteq L^r(U), 0 = t_0 < \ldots < t_N = T \Big\},$$

is dense in $L^r_{\gamma,T}$.

Proof. (1) We first show an auxiliary result. For this purpose we define

$$D_1 := \Big\{ f = \sum_{n=1}^{N} g_n \mathbb{1}_{A_n} \colon (g_n)_{n=1}^N \subseteq L^2([0,T]), (A_n)_{n=1}^N \subseteq \Sigma \text{ with } \mu(A_n) < \infty \Big\}.$$

Now let $\varepsilon > 0$, and choose an arbitrary $f_1 = \sum_{n=1}^N g_n \mathbb{1}_{A_n} \in D_1$, where we may assume that $\mu(A_n) > 0$ for every $n = 1, \ldots, N$. Since $g_n \in L^2([0,T])$, there exists a function

$$h_n = \sum_{k=1}^{K_n} \alpha_k^{(n)} \mathbb{1}_{[t_{k-1}^{(n)}, t_k^{(n)})}$$

with $\big(\alpha_k^{(n)}\big)_{k=1}^{K_n} \subseteq \mathbb{R}$, $0 = t_0^{(n)} < \ldots < t_{K_n}^{(n)} = T$, and $\|g_n - h_n\|_{L^2([0,T])} < \frac{\varepsilon}{N\mu(A_n)^{\frac{1}{r}}}$ for each $n = 1, \ldots, N$. We next define

$$f_2 := \sum_{n=1}^{N} h_n \mathbb{1}_{A_n} = \sum_{n=1}^{N} \sum_{k=1}^{K_n} \alpha_k^{(n)} \mathbb{1}_{[t_{k-1}^{(n)}, t_k^{(n)})} \mathbb{1}_{A_n},$$

which is an element of $D_{r,T}$ and satisfies $\|f_1 - f_2\|_{L^r_{\gamma,T}} < \varepsilon$. Let us prove this. Since $\big(t_k^{(n)}\big)_{k=0,\ldots,K_n}^{n=1,\ldots,N} \subseteq \mathbb{R}$, we can arrange the sequence in ascending order. Let $(t_i)_{i=0}^M$ be the sequence sorted this way, with $M := \sum_{n=1}^N K_n + N - 1$. For every set $[t_{i-1}, t_i)$ and all $n = 1, \ldots, N$, there exists exactly one $k_{i,n} \in \{1, \ldots, K_n\}$ with $[t_{i-1}, t_i) \subseteq [t_{k_{i,n}-1}^{(n)}, t_{k_{i,n}}^{(n)})$. For $i = 1, \ldots, M$ we define

$$x_i := \sum_{n=1}^{N} \alpha_{k_{i,n}}^{(n)} \mathbb{1}_{A_n}.$$

By recalling that $\mu(A_n) < \infty$ for all $n = 1, \ldots, N$, we infer that $x_i \in L^r(U)$ for each $i = 1, \ldots, M$, and we have

$$f_2 = \sum_{n=1}^{N} \sum_{k=1}^{K_n} \alpha_k^{(n)} \mathbb{1}_{[t_{k-1}^{(n)}, t_k^{(n)})} \mathbb{1}_{A_n} = \sum_{i=1}^{M} x_i \mathbb{1}_{[t_{i-1}, t_i)},$$

which means that $f_2 \in D_{r,T}$. Moreover,

$$\begin{aligned}
\|f_1 - f_2\|_{L^r_{\gamma,T}} &= \bigg(\int_U \Big\| \sum_{n=1}^{N} (g_n - h_n) \mathbb{1}_{A_n} \Big\|^r_{L^2([0,T])} \,\mathrm{d}\mu \bigg)^{\frac{1}{r}} \\
&\leq \sum_{n=1}^{N} \Big(\int_U \mathbb{1}_{A_n} \|g_n - h_n\|^r_{L^2([0,T])} \,\mathrm{d}\mu \Big)^{\frac{1}{r}} \\
&< \sum_{n=1}^{N} \frac{\varepsilon}{N \mu(A_n)^{\frac{1}{r}}} \Big(\int_U \mathbb{1}_{A_n} \,\mathrm{d}\mu \Big)^{\frac{1}{r}} \\
&= \sum_{n=1}^{N} \frac{\varepsilon}{N} = \varepsilon.
\end{aligned}$$

(2) Now we return to the actual proof. Note that $D_{r,T} \subseteq L^r_{\gamma,T}$. Let again $\varepsilon > 0$, and choose an arbitrary $f \in L^r_{\gamma,T}$. Observe that $\overline{D_1}^{L^r_{\gamma,T}} = L^r_{\gamma,T}$ (cf. Chapter A.1). Therefore, there exists an $h \in D_1$ with $\|f - h\|_{L^r_{\gamma,T}} < \frac{\varepsilon}{2}$. By **(1)** we can find a $g \in D_{r,T}$ with $\|h - g\|_{L^r_{\gamma,T}} < \frac{\varepsilon}{2}$, and this leads to

$$\|f - g\|_{L^r_{\gamma,T}} \leq \|f - h\|_{L^r_{\gamma,T}} + \|h - g\|_{L^r_{\gamma,T}} < \varepsilon. \qquad \blacksquare$$

By Lemma 3.3, Lemma 3.4, and the open mapping theorem, we get the following result.

Theorem 3.5 (Itô isomorphism). *For every $1 \leq p < \infty$, the linear mapping*

$$I^\beta \colon D_{r,T} \to L^p(\Omega; L^r(U)), \quad f \mapsto \int_0^T f \,\mathrm{d}\beta$$

has a unique extension to a bounded operator

$$I^\beta \colon L^r_{\gamma,T} \to L^p(\Omega; L^r(U)),$$

which is an isomorphism onto its range and satisfies (3.1).

This motivates the next definition.

Definition 3.6. *Let $f \in L^r_{\gamma,T}$ and $(f_n)_{n=1}^\infty \subseteq D_{r,T}$ be an approximating sequence of f. Then we define the* stochastic integral *of f with respect to $\big(\beta(t)\big)_{t\in[0,T]}$ by*

$$\int_0^T f \,\mathrm{d}\beta := I^\beta(f) = \lim_{n\to\infty} \int_0^T f_n \,\mathrm{d}\beta,$$

where the convergence holds in $L^p(\Omega; L^r(U))$ for each $1 \le p < \infty$.

Remark 3.7. (1) Lemma 3.4 guarantees that such an approximating sequence always exists, and Theorem 3.5 ensures that $\int_0^T f \,\mathrm{d}\beta \in L^p(\Omega; L^r(U))$ for each $1 \le p < \infty$ and every $f \in L^r_{\gamma,T}$.

(2) The stochastic integral is well-defined in the sense that it is independent of the approximating sequence.

(3) In general, the space $L^r_{\gamma,T}$ does not consist of functions $f\colon [0,T] \to L^r(U)$, even though the functions of $D_{r,T}$ are dense in this space. ■

Remark 3.8. Taking a closer look on the case $r = 2$ and $p = 2$, we note that we have equality in (3.1). This follows from Lemma 1.1 (observe that the constant is 1 in this case) and Fubini's theorem. Moreover, we have

$$\langle h_1, h_2\rangle = \frac{1}{4}\big(\|h_1 + h_2\|_2^2 - \|h_1 - h_2\|_2^2\big)$$

for all $h_1, h_2 \in L^2(U)$. Using this, Theorem 3.5, and Fubini's theorem, we obtain for all step functions f and g the formula

$$\begin{aligned}
\mathbb{E}\Big\langle \int_0^T f \,\mathrm{d}\beta, \int_0^T g \,\mathrm{d}\beta \Big\rangle &= \frac{1}{4}\Big(\mathbb{E}\Big\| \int_0^T f + g \;\mathrm{d}\beta \Big\|_2^2 - \mathbb{E}\Big\| \int_0^T f - g \;\mathrm{d}\beta \Big\|_2^2\Big) \\
&= \frac{1}{4}\Big(\Big\| \Big(\int_0^T |f+g|^2 \,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_2^2 - \Big\| \Big(\int_0^T |f-g|^2 \,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_2^2\Big) \\
&= \frac{1}{4}\Big(\int_U \int_0^T |f+g|^2 \,\mathrm{d}t \,\mathrm{d}\mu - \int_U \int_0^T |f-g|^2 \,\mathrm{d}t \,\mathrm{d}\mu\Big) \\
&= \frac{1}{4}\Big(\int_0^T \|f+g\|_2^2 \,\mathrm{d}t - \int_0^T \|f-g\|_2^2 \,\mathrm{d}t\Big) \\
&= \int_0^T \langle f, g\rangle \,\mathrm{d}t.
\end{aligned}$$

■

Next, we collect some properties of this stochastic integral.

Corollary 3.9. *For every $f \in L^r_{\gamma,T}$ and $0 \leq s \leq t \leq T$ we almost surely have*

$$\int_s^t f \,\mathrm{d}\beta := \int_s^t f|_{[s,t)} \,\mathrm{d}\beta = \int_0^T \mathbb{1}_{[s,t)} f \,\mathrm{d}\beta.$$

Proof. **(1)** Let $f \in D_{r,T}$. Then there exists an $N \in \mathbb{N}, (x_n)_{n=1}^N \subseteq L^r(U)$, and $0 = t_0 < \ldots < t_N = T$ with $f = \sum_{n=1}^N x_n \mathbb{1}_{[t_{n-1},t_n)}$. Moreover, there exists a $k \in \{1, \ldots, N\}$ satisfying $t_{k-1} \leq t \leq t_k$. Therefore,

$$\mathbb{1}_{[0,t)} f = \sum_{n=1}^{k-1} x_n \mathbb{1}_{[t_{n-1},t_n)} + x_k \mathbb{1}_{[t_{k-1},t)} + 0 \cdot \mathbb{1}_{[t,T)}.$$

Clearly, $\mathbb{1}_{[0,t)} f \in D_{r,t} \cap D_{r,T}$ and $\mathbb{1}_{[0,t)} f = f|_{[0,t)}$ on $[0,t)$. We compute

$$\int_0^T \mathbb{1}_{[0,t)} f \,\mathrm{d}\beta = \sum_{n=1}^{k-1} x_n \big(\beta(t_n) - \beta(t_{n-1})\big) + x_k \big(\beta(t) - \beta(t_{k-1})\big) = \int_0^t f \,\mathrm{d}\beta.$$

Now let $f \in L^r_{\gamma,T}$ and $(f_n)_{n=1}^\infty \subseteq D_{r,T}$ be an approximating sequence of f. Then we have

$$\lim_{n\to\infty} \|\mathbb{1}_{[0,t)} f_n - \mathbb{1}_{[0,t)} f\|_{L^r_{\gamma,T}} = \lim_{n\to\infty} \|f_n - f\|_{L^r_{\gamma,t}} \leq \lim_{n\to\infty} \|f_n - f\|_{L^r_{\gamma,T}} = 0,$$

which leads to

$$\int_0^T \mathbb{1}_{[0,t)} f \,\mathrm{d}\beta = \lim_{n\to\infty} \int_0^T \mathbb{1}_{[0,t)} f_n \,\mathrm{d}\beta = \lim_{n\to\infty} \int_0^t f_n \,\mathrm{d}\beta = \int_0^t f \,\mathrm{d}\beta \quad \text{in } L^p(\Omega; L^r(U)).$$

(2) Similarly as in **(1)**, we obtain

$$\int_s^T f \,\mathrm{d}\beta = \int_0^T \mathbb{1}_{[s,T)} f \,\mathrm{d}\beta.$$

Therefore, we finally get

$$\int_s^t f \,\mathrm{d}\beta = \int_0^t \mathbb{1}_{[s,t)} f \,\mathrm{d}\beta = \int_0^T \mathbb{1}_{[0,t)} \mathbb{1}_{[s,t)} f \,\mathrm{d}\beta = \int_0^T \mathbb{1}_{[s,t)} f \,\mathrm{d}\beta. \qquad \blacksquare$$

Corollary 3.10. *For every $f \in L^r_{\gamma,T}$, the random variable $\int_0^T f \,\mathrm{d}\beta$ is Gaussian, especially $\mathbb{E} \int_0^T f \,\mathrm{d}\beta = 0$.*

Proof. **(1)** For a step function $f = \sum_{n=1}^{N} x_n \mathbb{1}_{[t_{n-1},t_n)}$ with $x_1, \ldots, x_N \in L^r(U)$ and $0 = t_0 < \ldots < t_N = T$, we set

$$\gamma_n := \frac{1}{\sqrt{t_n - t_{n-1}}} \big(\beta(t_n) - \beta(t_{n-1})\big)$$

for each $n = 1, \ldots, N$. Then, by the definition of the Brownian motion, $(\gamma_n)_{n=1}^{N}$ is a sequence of independent standard Gaussian variables, and therefore

$$\int_0^T f \,\mathrm{d}\beta = \sum_{n=1}^{N} x_n \big(\beta(t_n) - \beta(t_{n-1})\big) = \sum_{n=1}^{N} x_n \sqrt{t_n - t_{n-1}}\gamma_n$$

is a Gaussian random variable.

(2) Now, let $f \in L^r_{\gamma,T}$ and $(f_n)_{n=1}^{\infty} \subseteq D_{r,T}$ be an approximating sequence of f. Then, for every $1 \leq p < \infty$, we have

$$\lim_{n\to\infty} \int_0^T f_n \,\mathrm{d}\beta = \int_0^T f \,\mathrm{d}\beta \quad \text{in } L^p(\Omega; L^r(U)).$$

Hence, Theorem 1.16 and Corollary 1.9 conclude the proof. ■

Remark 3.11. By the previous corollary and Theorem 1.16, the L^p convergence in Definition 3.6 is equivalent to convergence in probability. ■

In what follows, it may be of interest under which conditions a function $f\colon [0,T] \to L^r(U)$, which satisfies $\langle f, g\rangle \in L^2([0,T])$ for all $g \in L^{r'}(U)$, is already an element of $L^r_{\gamma,T}$, i.e., has a well-defined stochastic integral. For this purpose we need the following proposition.

Proposition 3.12. *Fix $1 \leq p < \infty$ and define, for $n \in \mathbb{N}_0$, the linear operator $A_n : L^p([0,T]; L^r(U)) \to L^p([0,T]; L^r(U))$ by*

$$A_n f := \sum_{j=1}^{2^n} \mathbb{1}_{[\frac{(j-1)T}{2^n}, \frac{jT}{2^n})} x_{j,n},$$

where

$$x_{j,n} := \frac{2^n}{T} \int_{\frac{(j-1)T}{2^n}}^{\frac{jT}{2^n}} f \,\mathrm{d}t \in L^r(U).$$

Then we have $\lim_{n\to\infty} A_n f = f$ in $L^p([0,T]; L^r(U))$ for all $f \in L^p([0,T]; L^r(U))$.

Proof. Let $n \in \mathbb{N}_0$ and $f \in L^p([0,T];L^r(U))$. Since the dyadic intervals are disjoint, we get

$$\|A_n f\|_r^p = \sum_{j=1}^{2^n} \mathbb{1}_{[\frac{(j-1)T}{2^n},\frac{jT}{2^n})} \|x_{j,n}\|_r^p \leq \sum_{j=1}^{2^n} \mathbb{1}_{[\frac{(j-1)T}{2^n},\frac{jT}{2^n})} \left(\frac{2^n}{T} \int_{\frac{(j-1)T}{2^n}}^{\frac{jT}{2^n}} \|f\|_r \, \mathrm{d}s \right)^p.$$

Since $\frac{2^n}{T}\,\mathrm{d}s$ is a probability measure, and since $x \mapsto |x|^p$ is a convex function, Jensen's inequality yield

$$\|A_n f\|_r^p \leq \sum_{j=1}^{2^n} \mathbb{1}_{[\frac{(j-1)T}{2^n},\frac{jT}{2^n})} \frac{2^n}{T} \int_{\frac{(j-1)T}{2^n}}^{\frac{jT}{2^n}} \|f\|_r^p \, \mathrm{d}s.$$

Therefore, we obtain

$$\begin{aligned}\int_0^T \|A_n f\|_r^p \, \mathrm{d}t &\leq \sum_{j=1}^{2^n} \int_0^T \frac{2^n}{T} \mathbb{1}_{[\frac{(j-1)T}{2^n},\frac{jT}{2^n})} \, \mathrm{d}t \int_{\frac{(j-1)T}{2^n}}^{\frac{jT}{2^n}} \|f\|_r^p \, \mathrm{d}s \\ &= \sum_{j=1}^{2^n} \int_{\frac{(j-1)T}{2^n}}^{\frac{jT}{2^n}} \|f\|_r^p \, \mathrm{d}s = \int_0^T \|f\|_r^p \, \mathrm{d}s.\end{aligned}$$

Hence, A_n is bounded and $\|A_n\| \leq 1$.

We next want to show that dyadic step functions are dense in $L^p([0,T];L^r(U))$. Since continuous functions are dense in this space, it suffices to approximate them by dyadic step functions. For this purpose, let $f \in C([0,T];L^r(U))$. Then f is uniformly continuous. Let $\varepsilon > 0$ be arbitrary. Then there exists a $\delta > 0$ such that $|s-t| < \delta$ implies $\|f(s) - f(t)\|_r < \varepsilon$ for all $s,t \in [0,T]$. For $k \in \mathbb{N}$ we define a dyadic step function by $f_k := A_k f$. By choosing k so large such that $\frac{T}{2^k} < \delta$, we obtain similar as above

$$\begin{aligned}\int_0^T \|f_k - f\|_r^p \, \mathrm{d}t &= \int_0^T \sum_{j=1}^{2^k} \mathbb{1}_{[\frac{(j-1)T}{2^k},\frac{jT}{2^k})}(t) \left\| \frac{2^k}{T} \int_{\frac{(j-1)T}{2^k}}^{\frac{jT}{2^k}} f(s) \, \mathrm{d}s - f(t) \right\|_r^p \mathrm{d}t \\ &\leq \sum_{j=1}^{2^k} \int_{\frac{(j-1)T}{2^k}}^{\frac{jT}{2^k}} \frac{2^k}{T} \int_{\frac{(j-1)T}{2^k}}^{\frac{jT}{2^k}} \|f(s) - f(t)\|_r^p \, \mathrm{d}s \, \mathrm{d}t \\ &\leq \sum_{j=1}^{2^k} \int_{\frac{(j-1)T}{2^k}}^{\frac{jT}{2^k}} \varepsilon^p \, \mathrm{d}t = T\varepsilon^p.\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, this shows the desired density result.

Finally, for a dyadic step function f we have $A_n f = f$ for n large enough, and the result now follows by the uniform boundedness of the operators $(A_n)_{n=1}^\infty$. ■

Theorem 3.13. *Fix $1 \leq p < \infty$. For a measurable function $f\colon [0,T] \to L^r(U)$, which satisfies $\langle f, g\rangle \in L^2([0,T])$ for all $g \in L^{r'}(U)$, the following assertions are equivalent:*

(1) $f \in L^r_{\gamma,T}$;

(2) there exists a random variable $X\colon \Omega \to L^r(U)$ such that for all $g \in L^{r'}(U)$

$$\langle X, g\rangle = \int_0^T \langle f, g\rangle \,\mathrm{d}\beta \quad \textit{almost surely.}$$

In this situation, we have $X = \int_0^T f \,\mathrm{d}\beta$ almost surely.

Proof. (1) $\Rightarrow$ (2): Let $g \in L^{r'}(U)$ be arbitrary, and let $(f_n)_{n=1}^\infty \subseteq D_{r,T}$ be an approximating sequence for f. Take $X := \int_0^T f \,\mathrm{d}\beta \in L^p(\Omega; L^r(U))$ as well as $X_n := \int_0^T f_n \,\mathrm{d}\beta$. Then $\lim_{n\to\infty} X_n = X$ in $L^p(\Omega; L^r(U))$, and for each $n \in \mathbb{N}$ and appropriate sequences $0 = t_1^{(n)} < \ldots < t_{N_n}^{(n)} = T$ and $(x_k^{(n)})_{k=1}^{N_n} \subseteq L^r(U)$, almost surely we have

$$\langle X_n, g\rangle = \sum_{k=1}^{N_n} \langle x_k^{(n)}, g\rangle \big(\beta(t_k^{(n)}) - \beta(t_{k-1}^{(n)})\big) = \int_0^T \langle f_n, g\rangle \,\mathrm{d}\beta.$$

This implies

$$\lim_{n\to\infty} \int_0^T \langle f_n, g\rangle \,\mathrm{d}\beta = \lim_{n\to\infty} \langle X_n, g\rangle = \langle X, g\rangle \quad \text{in } L^p(\Omega). \tag{$\triangle$}$$

Since (U, Σ, μ) is a σ-finite measure space, U is a countable union of disjoint sets $U_k \in \Sigma$, $k \in \mathbb{N}$, of finite measure. Take $U^{(K)} := \bigcup_{k=1}^K U_k$. By Hölder's inequality and Fubini's theorem we obtain for each fixed $K \in \mathbb{N}$

$$\lim_{n\to\infty} \mathbb{1}_{U^{(K)}} f_n = \mathbb{1}_{U^{(K)}} f \quad \text{in } L^1([0,T] \times U^{(K)}),$$

since $\lim_{n\to\infty} f_n = f$ in $L^r_{\gamma,T}$. Moreover, there exists an appropriate subsequence and a set $N^{(K)} \subseteq [0,T] \times U^{(K)}$ of measure 0, such that

$$\lim_{j\to\infty} \mathbb{1}_{U^{(K)}}(u) f_{n_j}(t,u) = \mathbb{1}_{U^{(K)}}(u) f(t,u)$$

and hence

$$\lim_{j\to\infty} \mathbb{1}_{U^{(K)}}(u) f_{n_j}(t,u) g(u) = \mathbb{1}_{U^{(K)}}(u) f(t,u) g(u)$$

for all $(t,u) \in \big([0,T] \times U^{(K)}\big) \setminus N^{(K)}$.

We note that $\mathbb{1}_{U(K)}f(t) \in L^r(U)$, and thus $\mathbb{1}_{U(K)}f(t)g \in L^1(U)$ for each $t \in [0,T]$. Therefore, by the dominated convergence theorem, we get for almost all $t \in [0,T]$

$$\begin{aligned}\lim_{j\to\infty} \langle \mathbb{1}_{U(K)} f_{n_j}(t), g\rangle &= \lim_{j\to\infty} \int_U \mathbb{1}_{U(K)} f_{n_j}(t) g \,\mathrm{d}\mu = \int_U \mathbb{1}_{U(K)} f(t) g \,\mathrm{d}\mu \\ &= \langle \mathbb{1}_{U(K)} f(t), g\rangle.\end{aligned}$$

Since $\mathbb{1}_{U(K)}g \in L^{r'}(U)$, we have $\langle \mathbb{1}_{U(K)}f, g\rangle = \langle f, \mathbb{1}_{U(K)}g\rangle \in L^2([0,T])$, by assumption. Now, again the dominated convergence theorem yield

$$\lim_{j\to\infty} \langle f_{n_j}, \mathbb{1}_{U(K)}g\rangle = \lim_{j\to\infty} \langle \mathbb{1}_{U(K)} f_{n_j}, g\rangle = \langle \mathbb{1}_{U(K)} f, g\rangle = \langle f, \mathbb{1}_{U(K)}g\rangle \quad \text{in } L^2([0,T]).$$

And together with ($\triangle$) we infer that

$$\int_0^T \langle f, \mathbb{1}_{U(K)}g\rangle \,\mathrm{d}\beta = \lim_{j\to\infty} \int_0^T \langle f_{n_j}, \mathbb{1}_{U(K)}g\rangle \,\mathrm{d}\beta = \langle X, \mathbb{1}_{U(K)}g\rangle \quad \text{in } L^p(\Omega).$$

Once more by the dominated convergence theorem we have

$$\lim_{K\to\infty} \langle X, \mathbb{1}_{U(K)}g\rangle = \langle X, g\rangle \quad \text{in } L^p(\Omega)$$

and

$$\lim_{K\to\infty} \langle f, \mathbb{1}_{U(K)}g\rangle = \langle f, g\rangle \quad \text{in } L^2([0,T]),$$

and this finally leads to

$$\langle X, g\rangle = \lim_{K\to\infty} \langle X, \mathbb{1}_{U(K)}g\rangle = \lim_{K\to\infty} \int_0^T \langle f, \mathbb{1}_{U(K)}g\rangle \,\mathrm{d}\beta = \int_0^T \langle f, g\rangle \,\mathrm{d}\beta \quad \text{in } L^p(\Omega).$$

(2) $\Rightarrow$ (1): For each $k \in \mathbb{N}$ we set $f^{(k)}(t) := \mathbb{1}_{\{\|f(t)\|_r \le k\}}(t) f(t)$, $t \in [0,T]$. Note that $f^{(k)} \in L^2([0,T]; L^r(U))$. Moreover, since $\lim_{n\to\infty} \langle f^{(n)}(t), g\rangle = \langle f(t), g\rangle$ for all $t \in [0,T]$ and $\langle f, g\rangle \in L^2([0,T])$ for all $g \in L^{r'}(U)$, we obtain

$$\lim_{n\to\infty} \langle f^{(n)}, g\rangle = \langle f, g\rangle \quad \text{in } L^2([0,T])$$

by the dominated convergence theorem. Next, define $f_n := A_n f^{(n)}$ for $n \in \mathbb{N}$, and with A_n being the operator from Proposition 3.12. Since $f^{(n)} \in L^2([0,T]; L^r(U))$, we have for each $g \in L^{r'}(U)$ and all $j = 1, \dots, 2^n$

$$\left\langle \int_{\frac{(j-1)T}{2^n}}^{\frac{jT}{2^n}} f^{(n)}(t) \,\mathrm{d}t, g \right\rangle = \int_{\frac{(j-1)T}{2^n}}^{\frac{jT}{2^n}} \langle f^{(n)}, g\rangle(t) \,\mathrm{d}t,$$

and this leads to

$$\langle f_n, g\rangle = \sum_{j=1}^{2^n} \mathbb{1}_{[\frac{(j-1)T}{2^n}, \frac{jT}{2^n})} \frac{2^n}{T} \int_{\frac{(j-1)T}{2^n}}^{\frac{jT}{2^n}} \langle f^{(n)}, g\rangle(t) \,\mathrm{d}t = A_n \langle f^{(n)}, g\rangle.$$

Therefore, by the properties of A_n, we have

$$\begin{aligned}
&\lim_{n\to\infty} \|\langle f_n, g\rangle - \langle f, g\rangle\|_{L^2([0,T])} \\
&\qquad \le \lim_{n\to\infty} \big(\|\langle f_n, g\rangle - A_n\langle f, g\rangle\|_{L^2([0,T])} + \|A_n\langle f, g\rangle - \langle f, g\rangle\|_{L^2([0,T])}\big) \qquad (\star) \\
&\qquad \le \lim_{n\to\infty} \big(\|\langle f^{(n)}, g\rangle - \langle f, g\rangle\|_{L^2([0,T])} + \|A_n\langle f, g\rangle - \langle f, g\rangle\|_{L^2([0,T])}\big) = 0.
\end{aligned}$$

Next, we define $X_n := \int_0^T f_n - f_{n-1}\,\mathrm{d}\beta$ for $n \in \mathbb{N}$ and with $f_0 := 0$. Then we get

$$\langle X_n, h\rangle = \int_0^T \langle f_n - f_{n-1}, h\rangle\,\mathrm{d}\beta$$

for each $n \in \mathbb{N}$ and every $h \in L^{r'}(U)$. We now want to show that the random variables X_n are independent. By the linearity of the stochastic integral, the random variables X_n are jointly Gaussian. Therefore, by Proposition A.16, it suffices to check that $\mathbb{E}\langle X_m, h_1\rangle\langle X_n, h_2\rangle = 0$ for $m \neq n$ and $h_1, h_2 \in L^{r'}(U)$. Together with Remark 3.8 we obtain

$$\mathbb{E}\langle X_m, h_1\rangle\langle X_n, h_2\rangle = \int_0^T \langle f_m - f_{m-1}, h_1\rangle\langle f_n - f_{n-1}, h_2\rangle\,\mathrm{d}t.$$

For $j = 1, \dots, 2^m$ and $k = 1, \dots, 2^n$ we define

$$x_{j,m} := \frac{2^m}{T}\int_{\frac{(j-1)T}{2^m}}^{\frac{jT}{2^m}} \langle f^{(m)}, h_1\rangle\,\mathrm{d}t, \quad y_{k,n} := \frac{2^n}{T}\int_{\frac{(k-1)T}{2^n}}^{\frac{kT}{2^n}} \langle f^{(n)}, h_2\rangle\,\mathrm{d}t.$$

Let $m > n$. Then we obtain

$$\begin{aligned}
\mathbb{E}\langle X_m, h_1\rangle\langle X_n, h_2\rangle &= \int_0^T \Big(\sum_{j=1}^{2^m} \mathbb{1}_{[\frac{(j-1)T}{2^m}, \frac{jT}{2^m})} x_{j,m} - \sum_{j=1}^{2^{m-1}} \mathbb{1}_{[\frac{(j-1)T}{2^{m-1}}, \frac{jT}{2^{m-1}})} x_{j,m-1}\Big) \\
&\qquad\qquad \cdot \Big(\sum_{k=1}^{2^n} \mathbb{1}_{[\frac{(k-1)T}{2^n}, \frac{kT}{2^n})} y_{k,n} - \sum_{k=1}^{2^{n-1}} \mathbb{1}_{[\frac{(k-1)T}{2^{n-1}}, \frac{kT}{2^{n-1}})} y_{k,n-1}\Big)\,\mathrm{d}t \\
= \frac{T}{2^m}\sum_{k=1}^{2^n}\Big(&\sum_{l=1+(k-1)2^{m-n}}^{k2^{m-n}} x_{l,m}\Big) y_{k,n} - \frac{T}{2^{m-1}}\sum_{k=1}^{2^n}\Big(\sum_{l=1+(k-1)2^{m-n-1}}^{k2^{m-n-1}} x_{l,m-1}\Big) y_{k,n} \\
- \frac{T}{2^m}\sum_{k=1}^{2^{n-1}}\Big(&\sum_{l=1+(k-1)2^{m-n+1}}^{k2^{m-n+1}} x_{l,m}\Big) y_{k,n-1} + \frac{T}{2^{m-1}}\sum_{k=1}^{2^{n-1}}\Big(\sum_{l=1+(k-1)2^{m-n}}^{k2^{m-n}} x_{l,m-1}\Big) y_{k,n-1} \\
= \sum_{k=1}^{2^n}\Big(&\frac{T}{2^m}\sum_{l=1+(k-1)2^{m-n}}^{k2^{m-n}} x_{l,m} - \frac{T}{2^{m-1}}\sum_{l=1+(k-1)2^{m-n-1}}^{k2^{m-n-1}} x_{l,m-1}\Big) y_{k,n} \\
- \sum_{k=1}^{2^{n-1}}\Big(&\frac{T}{2^m}\sum_{l=1+(k-1)2^{m-n+1}}^{k2^{m-n+1}} x_{l,m} - \frac{T}{2^{m-1}}\sum_{l=1+(k-1)2^{m-n}}^{k2^{m-n}} x_{l,m-1}\Big) y_{k,n-1} = 0,
\end{aligned}$$

since for each k we have

$$\frac{T}{2^m}\sum_{l=1+(k-1)2^{m-n}}^{k2^{m-n}} x_{l,m} = \int_{\frac{(k-1)T}{2^n}}^{\frac{kT}{2^n}} \langle f^{(m)}, h_1\rangle \,\mathrm{d}t = \frac{T}{2^{m-1}}\sum_{l=1+(k-1)2^{m-n-1}}^{k2^{m-n-1}} x_{l,m-1},$$
$$\frac{T}{2^m}\sum_{l=1+(k-1)2^{m-n+1}}^{k2^{m-n+1}} x_{l,m} = \int_{\frac{(k-1)T}{2^{n-1}}}^{\frac{kT}{2^{n-1}}} \langle f^{(m)}, h_1\rangle \,\mathrm{d}t = \frac{T}{2^{m-1}}\sum_{l=1+(k-1)2^{m-n}}^{k2^{m-n}} x_{l,m-1}.$$

Now take $S_N := \sum_{n=1}^N X_n = \int_0^T f_N \,\mathrm{d}\beta$. Then, by (2), Theorem 3.5, and ($\star$), we obtain

$$\begin{aligned}\lim_{N\to\infty} \mathbb{E}\big|\langle X, g\rangle - \langle S_N, g\rangle\big|^2 &= \lim_{N\to\infty} \mathbb{E}\Big(\int_0^T \langle f, g\rangle \,\mathrm{d}\beta - \int_0^T \langle f_N, g\rangle \,\mathrm{d}\beta\Big)^2 \\ &= \lim_{N\to\infty} \mathbb{E}\Big(\int_0^T \langle f - f_N, g\rangle \,\mathrm{d}\beta\Big)^2 \\ &= \lim_{N\to\infty} \int_0^T \langle f - f_N, g\rangle^2 \,\mathrm{d}t = 0\end{aligned}$$

for all $g \in L^{r'}(U)$. Then, by Proposition A.17, X is a Gaussian random variable, and by Fernique's theorem we deduce that $X \in L^p(\Omega; L^r(U))$. Thus, applying Proposition 1.13 leads to

$$\lim_{N\to\infty} \int_0^T f_N \,\mathrm{d}\beta = X \quad \text{in } L^p(\Omega; L^r(U)).$$

By this convergence and Theorem 3.5, $(f_n)_{n=1}^\infty$ is Cauchy in $L^r_{\gamma,T}$ with limit $\widetilde{f}$. As in the first part of this proof, we can show that

$$\lim_{j\to\infty} \mathbb{1}_{U^{(N)}}(u) f_{n_j}(t,u) = \mathbb{1}_{U^{(N)}}(u)\widetilde{f}(t,u)$$

for almost all $(t,u) \in [0,T] \times U^{(N)}$ and any fixed $N \in \mathbb{N}$, with $U^{(N)} \in \Sigma$ defined as above. And similarly, we can show that

$$\lim_{k\to\infty} \mathbb{1}_{U^{(N)}}(u) f_{n_{j_k}}(t,u) = \mathbb{1}_{U^{(N)}}(u) f(t,u)$$

for almost all $(t,u) \in [0,T] \times U^{(N)}$ and all $N \in \mathbb{N}$, since $\lim_{j\to\infty} f_{n_j} = f$ in $L^2([0,T]; L^r(U))$. Therefore, we have $f = \widetilde{f} \in L^r_{\gamma,T}$. This also shows that

$$X = \int_0^T \widetilde{f} \,\mathrm{d}\beta = \int_0^T f \,\mathrm{d}\beta \quad \text{almost surely.}$$

■

3.2 The Itô Integral

In the previous section we were introduced to a stochastic integral of functions $f\colon [0,T] \to L^r(U)$. Next, we are going to define a stochastic integral for processes $f\colon [0,T] \times \Omega \to L^r(U)$, where we now assume that $1 < r < \infty$. As it turns out, the Decoupling theorem will play a central role.

Definition 3.14. *A function $f\colon [0,T] \times \Omega \to L^r(U)$ is said to be an* adapted step process *with respect to a filtration $\mathbb{F} = (\mathscr{F}_t)_{t\in[0,T]}$ if it is of the form*

$$f(t,\omega) = \sum_{n=1}^{N} \mathbb{1}_{[t_{n-1},t_n)}(t) \sum_{k=1}^{K_n} x_k^{(n)} \mathbb{1}_{A_k^{(n)}}(\omega), \tag{3.2}$$

where $0 = t_0 < \ldots < t_N = T$ and, for all $n = 1,\ldots,N$, $x_1^{(n)},\ldots,x_{K_n}^{(n)} \in L^r(U)$ and $A_1^{(n)},\ldots,A_{K_n}^{(n)} \in \mathscr{F}_{t_{n-1}}$.

For the rest of this section we assume that $\big(\beta(t)\big)_{t\in[0,T]}$ is a Brownian motion which is *adapted* to the filtration $\mathbb{F}$ in the sense that for all $0 \le s < t \le T$

(1) $\beta(t)$ is $\mathscr{F}_t$-measurable and

(2) $\beta(t) - \beta(s)$ is independent of $\mathscr{F}_s$.

Definition 3.15. *The* stochastic integral *with respect to $\big(\beta(t)\big)_{t\in[0,T]}$ of an adapted step process f of the form (3.2) is defined by*

$$\int_0^T f \,\mathrm{d}\beta := \sum_{n=1}^{N} \sum_{k=1}^{K_n} x_k^{(n)} \mathbb{1}_{A_k^{(n)}} \big(\beta(t_n) - \beta(t_{n-1})\big).$$

Remark 3.16. It is straight forward to check that this definition does not depend on the particular representation of f. Also, if we choose $A_k^{(n)} = \Omega$ for all $n = 1,\ldots,N$ and each $k = 1,\ldots,K_n$, this integral coincides with the Wiener integral. From this point of view, the Itô integral is an extension of the integral given in section 3.1. ■

Additionally we have the following properties.

Proposition 3.17. *For every adapted step process f we have $\int_0^T f \,\mathrm{d}\beta \in L^p(\Omega, \mathscr{F}_T; L^r(U))$ for all $1 \le p < \infty$ satisfying $\mathbb{E} \int_0^T f \,\mathrm{d}\beta = 0$.*

Proof. Clearly, the random variable $\int_0^T f \,\mathrm{d}\beta$ is $\mathscr{F}_T$-measurable. By Remark 1.8 we have $\big(\mathbb{E}|\beta(t)-\beta(s)|^p\big)^{\frac{1}{p}} < \infty$ for every fixed $0 \le s < t \le T$ and all $1 \le p < \infty$. Hence,

$$\Big(\mathbb{E}\Big\| \int_0^T f \,\mathrm{d}\beta \Big\|_r^p\Big)^{\frac{1}{p}} = \Big(\mathbb{E}\Big\| \sum_{n=1}^N \sum_{k=1}^{K_n} x_k^{(n)} \mathbb{1}_{A_k^{(n)}} \big(\beta(t_n) - \beta(t_{n-1})\big) \Big\|_r^p\Big)^{\frac{1}{p}}$$
$$\le \sum_{n=1}^N \sum_{k=1}^{K_n} \|x_k^{(n)}\|_r \Big(\mathbb{E}\big|\beta(t_n) - \beta(t_{n-1})\big|^p\Big)^{\frac{1}{p}} < \infty.$$

In addition,

$$\mathbb{E}\Big(\mathbb{1}_{A_k^{(n)}} \big(\beta(t_n) - \beta(t_{n-1})\big)\Big) = \mathbb{E}\Big(\mathbb{E}\big[\mathbb{1}_{A_k^{(n)}} \big(\beta(t_n) - \beta(t_{n-1})\big)\big|\mathscr{F}_{t_{n-1}}\big]\Big)$$
$$= \mathbb{E}\Big(\mathbb{1}_{A_k^{(n)}} \mathbb{E}\big[\big(\beta(t_n) - \beta(t_{n-1})\big)\big|\mathscr{F}_{t_{n-1}}\big]\Big)$$
$$= \mathbb{E}\Big(\mathbb{1}_{A_k^{(n)}} \mathbb{E}\big(\beta(t_n) - \beta(t_{n-1})\big)\Big) = 0,$$

using that $\mathbb{1}_{A_k^{(n)}}$ is $\mathscr{F}_{t_{n-1}}$-measurable and the fact that $\beta(t_n) - \beta(t_{n-1})$ is independent of $\mathscr{F}_{t_{n-1}}$. Therefore, by linearity,

$$\mathbb{E} \int_0^T f \,\mathrm{d}\beta = \sum_{n=1}^N \sum_{k=1}^{K_n} x_k^{(n)} \mathbb{E}\Big(\mathbb{1}_{A_k^{(n)}} \big(\beta(t_n) - \beta(t_{n-1})\big)\Big) = 0. \qquad \blacksquare$$

Similar to Section 3.1, we want to give an estimate for the $L^p(\Omega; L^r(U))$-norm of $\int_0^T f \,\mathrm{d}\beta$. Here we are going to use results of Section 2.3, and therefore, we have to confine ourselves to the cases $1 < p < \infty$.

Lemma 3.18 (Itô isomorphism for adapted step processes). *For* $1 < p < \infty$ *and every adapted step process* $f\colon [0,T] \times \Omega \to L^r(U)$, *we have*

$$\mathbb{E}\Big\| \int_0^T f \,\mathrm{d}\beta \Big\|_r^p \eqsim_{p,r} \mathbb{E}\Big\| \Big(\int_0^T |f|^2 \,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_r^p. \tag{3.3}$$

Proof. Let f be an adapted step process of the form (3.2). Then we define for $n = 1, \dots, N$

$$v_n := \sum_{k=1}^{K_n} x_k^{(n)} \mathbb{1}_{A_k^{(n)}} \quad \text{and} \quad \mathscr{F}_n := \mathscr{F}_{t_n}.$$

So, $(v_n)_{n=1}^N$ is predictable with respect to the filtration $(\mathscr{F}_n)_{n=1}^N$ since every $A_k^{(n)}$ is $\mathscr{F}_{n-1}$-measurable. Additionally, by the general assumption in this section, $(\mathscr{F}_n)_{n=1}^N$ fulfills the conditions of Corollary 2.24.

Moreover,

$$\int_0^T |f|^2 \,\mathrm{d}t = \int_0^T \Big|\sum_{n=1}^N v_n \mathbb{1}_{[t_{n-1},t_n)}\Big|^2 \,\mathrm{d}t = \sum_{n=1}^N \int_0^T |v_n|^2 \mathbb{1}_{[t_{n-1},t_n)} \,\mathrm{d}t$$
$$= \sum_{n=1}^N |v_n|^2 (t_n - t_{n-1}).$$

Finally, by applying Corollary 2.24, we obtain

$$\mathbb{E}\Big\| \int_0^T f \,\mathrm{d}\beta \Big\|_r^p = \mathbb{E}\Big\| \sum_{n=1}^N v_n \big(\beta(t_n) - \beta(t_{n-1})\big) \Big\|_r^p$$
$$\eqsim_{p,r} \mathbb{E}\Big\| \Big(\sum_{n=1}^N |v_n|^2 (t_n - t_{n-1})\Big)^{\frac{1}{2}} \Big\|_r^p$$
$$= \mathbb{E}\Big\| \Big(\int_0^T |f|^2 \,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_r^p. \qquad \blacksquare$$

Definition 3.19. *A random variable* $f \in L^p(\Omega; L^r_{\gamma,T})$ *is called an* adapted L^p process *with respect to a filtration* $\mathbb{F}$ *if*

$$\mathbb{1}_{[0,t)} f \in L^p(\Omega, \mathscr{F}_t; L^r_{\gamma,t}) \quad \textit{for all } t \in [0,T],$$

which means that f *has a representative* $\widetilde{f} \colon [0,T] \times \Omega \times U \to \mathbb{R}$ *such that* $\mathbb{1}_{[0,t)} \widetilde{f}$ *is* $\mathscr{B}([0,t)) \otimes \mathscr{F}_t \otimes \Sigma$*-measurable.*

We denote by $L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ *the closed subspace in* $L^p(\Omega; L^r_{\gamma,T})$ *of all* $\mathbb{F}$*-adapted elements.*

Remark 3.20. Functions with the property given in the definition above are often called *progressively measurable*, whereas functions $f \colon [0,T] \times \Omega \times U \to \mathbb{R}$ are usually called *adapted* if $f(t)$ is $\mathscr{F}_t \otimes \Sigma$-measurable. But since the first implies the latter, the above definition is still meaningful. ■

By Definition 3.19, we get the following density result.

Lemma 3.21. *For every* $1 \le p < \infty$*, the space of all adapted step processes with respect to a filtration* $\mathbb{F}$ *is dense in* $L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$*.*

Proof. Note that the space of adapted step processes is a subset of $L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$.

(1) We first assume that $\mu(U) < \infty$. Taking $q := \max\{r, p, 2\}$, we observe that $L^q_{\mathbb{F}}(\Omega; L^q(U; L^q([0,T])))$ is dense in $L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$. Therefore it suffices to approximate functions in the first space. By Fubini's theorem we have

$$L^q(\Omega; L^q(U; L^q([0,T]))) \simeq L^q([0,T]; L^q(\Omega; L^q(U))) =: E.$$

Let $f \in E$ be an adapted process and $\varepsilon > 0$. Choose $\delta > 0$ such that $\|f_\delta - f\|_E < \varepsilon$ for $f_\delta := \mathbb{1}_{[0,T-\delta]} f(\cdot - \delta) \in E$, and define for $n \in \mathbb{N}$

$$f_n := A_n f_\delta = \sum_{j=1}^{2^n} \mathbb{1}_{[t_{j-1,n}, t_{j,n})} v_{j,n},$$

where

$$t_{j,n} := \frac{jT}{2^n}, \quad v_{j,n} := \frac{2^n}{T} \int_{t_{j-1,n}}^{t_{j,n}} f_\delta \,\mathrm{d}t \;\in L^q(\Omega; L^q(U)),$$

and A_n is the operator from Proposition 3.12. Now choose $N \in \mathbb{N}$ so large such that $\|f_N - f_\delta\|_E < \varepsilon$ and $\frac{T}{2^N} < \delta$. Then we have $\|f - f_N\|_E < 2\varepsilon$ and

$$v_{j,N} = \frac{2^N}{T} \int_{t_{j-1,N}}^{t_{j,N}} f_\delta \,\mathrm{d}t = \frac{2^N}{T} \int_{t_{j-1,N}-\delta}^{t_{j,N}-\delta} f \,\mathrm{d}t.$$

Since $t_{j,N} - \delta < t_{j,N} - \frac{T}{2^N} = t_{j-1,N}$ and f is an adapted process, we infer that f is $\mathscr{F}_{t_{j-1,N}}$-measurable on $[0, t_{j,N} - \delta)$. This implies that $v_{j,N} \in L^q(\Omega, \mathscr{F}_{t_{j-1,N}}; L^q(U))$. Finally, we can find for each $j = 1, \ldots, 2^N$ a simple function satisfying

$$\Big\| v_{j,N} - \sum_{k=1}^{N_j} x_k^{(j)} \mathbb{1}_{A_k^{(j)}} \Big\|_{L^q(\Omega; L^q(U))} < \frac{\varepsilon}{2^N} \frac{2^{\frac{N}{q}}}{T^{\frac{1}{q}}}$$

with $A_k^{(j)} \in \mathscr{F}_{t_{j-1,N}}$. Then,

$$g := \sum_{j=1}^{2^N} \mathbb{1}_{[t_{j-1,N}, t_{j,N})} \sum_{k=1}^{N_j} x_k^{(j)} \mathbb{1}_{A_k^{(j)}}$$

is an adapted step process with respect to $\mathbb{F}$, and

$$\begin{aligned}
\|f_N - g\|_E &= \left(\int_0^T \|f_N - g\|^q_{L^q(\Omega; L^q(U))} \,\mathrm{d}t \right)^{\frac{1}{q}} \\
&\le \sum_{j=1}^{2^N} \left(\int_0^T \mathbb{1}_{[t_{j-1,N}, t_{j,N})} \Big\| v_{j,N} - \sum_{k=1}^{N_j} x_k^{(j)} \mathbb{1}_{A_k^{(j)}} \Big\|^q_{L^q(\Omega; L^q(U))} \,\mathrm{d}t \right)^{\frac{1}{q}} \\
&< \sum_{j=1}^{2^N} \frac{\varepsilon}{2^N} \left(\int_0^T \mathbb{1}_{[t_{j-1,N}, t_{j,N})} \frac{2^N}{T} \,\mathrm{d}t \right)^{\frac{1}{q}} = \varepsilon.
\end{aligned}$$

Hence, $\|f-g\|_E < 3\varepsilon$. By Hölder's inequality we have

$$\|\cdot\|_{L^p(\Omega;L^r_{\gamma,T})} \lesssim \|\cdot\|_{L^q(\Omega;L^q(U;L^q([0,T])))}.$$

Therefore,

$$\|f-g\|_{L^p(\Omega;L^r_{\gamma,T})} \lesssim \|f-g\|_{L^q(\Omega;L^q(U;L^q([0,T])))} = \|f-g\|_E < 3\varepsilon.$$

(2) Now assume that (U,Σ,μ) is a σ-finite measure space, and let $\varepsilon > 0$. Then $U = \bigcup_{n=1}^\infty U_n$ for sets $U_n \in \Sigma$ with $\mu(U_n) < \infty$. Given $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ and $N \in \mathbb{N}$, we define

$$f_N := \mathbb{1}_{\bigcup_{n=1}^N U_n} f.$$

Since $f \in L^p(\Omega; L^r_{\gamma,T})$ and $\lim_{N\to\infty} f_N(\omega,u) = f(\omega,u)$ in $L^2([0,T])$ for all $\omega \in \Omega$ and $u \in U$, the dominated convergence theorem yield $\lim_{N\to\infty} f_N = f$ in $L^p(\Omega; L^r_{\gamma,T})$. Thus, we can choose an $N \in \mathbb{N}$ such that $\|f-f_N\|_{L^p(\Omega;L^r_{\gamma,T})} < \varepsilon$. Since f_N has support of finite measure, step **(1)** donates an adapted step process g which satisfies $\|f_N - g\|_{L^p(\Omega;L^r_{\gamma,T})} < \varepsilon$. This finally yields

$$\|f-g\|_{L^p(\Omega;L^r_{\gamma,T})} < 2\varepsilon.$$ ■

By Lemma 3.18 and Lemma 3.21 we get the next result.

Theorem 3.22 (Itô isomorphism for adapted L^p processes). *For every $1 < p < \infty$ the stochastic integral extends uniquely to a bounded operator*

$$I^\beta_{\mathbb{F}} : L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T}) \to L^p(\Omega, \mathscr{F}_T; L^r(U)),$$

which is an isomorphism onto its range and satisfies (3.3).

As in Section 3.1, we can now define a stochastic integral for adapted L^p processes.

Definition 3.23. *Fix $1 < p < \infty$. Let $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ and $(f_n)_{n=1}^\infty$ be an approximating sequence of adapted step processes. Then we define the* stochastic integral *of f by*

$$\int_0^T f \,\mathrm{d}\beta := I^\beta_{\mathbb{F}}(f) = \lim_{n\to\infty} \int_0^T f_n \,\mathrm{d}\beta,$$

where convergence holds in $L^p(\Omega, \mathscr{F}_T; L^r(U))$.

Remark 3.24. Similar to Remark 3.7, we have $\int_0^T f\,\mathrm{d}\beta \in L^p(\Omega; L^r(U))$ for all $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, and it is straight forward to check that the stochastic integral does not depend on the approximating sequence of f. Further, note that the space $L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ does generally not consist of functions $f\colon [0,T] \times \Omega \to L^r(U)$. ■

We now have the following equivalent to Corollary 3.9.

Corollary 3.25. *For $1 < p < \infty$, every $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, and $s, t \in [0,T]$, we have*

$$\int_s^t f\,\mathrm{d}\beta := \int_s^t f|_{[s,t)}\,\mathrm{d}\beta = \int_0^T \mathbb{1}_{[s,t)} f\,\mathrm{d}\beta.$$

Proof. This assertion is proved in a very similar way as Corollary 3.9. By replacing the sequence $(x_n)_{n=1}^N \subseteq L^r(U)$ with a sequence $(\sum_{k=1}^{K_n} x_k^{(n)} \mathbb{1}_{A_k^{(n)}})_{n=1}^N \subseteq L^p(\Omega; L^r(U))$, the statement follows in the same way for adapted step processes as for step functions. And once again it follows by an approximation argument that the result holds for any adapted L^p process. ■

If we choose the *augmented* Brownian filtration $\mathbb{F}^\beta := (\mathscr{F}_t^\beta)_{t\in[0,T]}$ (i.e., $\mathscr{F}_0^\beta$ contains all $\mathbb{P}$-null sets), we can get some interesting results. Recall that $\mathbb{F}^\beta$ fulfills the conditions we assumed above. See also Remark 1.19 and Remark 2.25.

Lemma 3.26. *Let $1 < p < \infty$ and $\xi \in L^p(\Omega, \mathscr{F}_T^\beta)$. Then there exists a unique $\phi \in L^p_{\mathbb{F}^\beta}(\Omega; L^2([0,T]))$ such that*

$$\xi = \mathbb{E}\xi + \int_0^T \phi\,\mathrm{d}\beta.$$

Proof. Let $(\xi_n)_{n=1}^\infty$ be a sequence of simple functions converging to ξ in $L^p(\Omega, \mathscr{F}_T^\beta)$. For each $n \in \mathbb{N}$ we have $\xi_n \in L^2(\Omega, \mathscr{F}_T^\beta)$. Thus, by the classical representation theorem (cf. [14, Theorem 12.2]), we obtain a $\phi_n \in L^2_{\mathbb{F}^\beta}(\Omega; L^2([0,T]))$ which satisfies

$$\xi_n = \mathbb{E}\xi_n + \int_0^T \phi_n\,\mathrm{d}\beta.$$

Since $\lim_{n\to\infty} \xi_n = \xi$ in $L^p(\Omega, \mathscr{F}_T^\beta)$, we have $\lim_{n\to\infty} \mathbb{E}\xi_n = \mathbb{E}\xi$. By Theorem 3.22 we infer that the sequence $(\phi_n)_{n=1}^\infty$ is Cauchy in $L^p_{\mathbb{F}^\beta}(\Omega; L^2([0,T]))$. Hence, there exists a unique limit $\phi \in L^p_{\mathbb{F}^\beta}(\Omega; L^2([0,T]))$ with the property

$$\xi - \mathbb{E}\xi = \lim_{n\to\infty} \xi_n - \mathbb{E}\xi_n = \lim_{n\to\infty} \int_0^T \phi_n\,\mathrm{d}\beta = \int_0^T \phi\,\mathrm{d}\beta.$$ ■

Theorem 3.27 (Representation theorem). *Let* $1 < p < \infty$ *and* $X \in L^p(\Omega, \mathscr{F}_T^\beta; L^r(U))$. *Then there exists a unique* $f \in L^p_{\mathbb{F}^\beta}(\Omega; L^r_{\gamma,T})$ *such that*

$$X = \mathbb{E}X + \int_0^T f \,\mathrm{d}\beta.$$

Proof. There exists a sequence $(X_n)_{n=1}^\infty$ of simple $\mathscr{F}_T^\beta$-measurable random variables with $\lim_{n\to\infty} X_n = X$ in $L^p(\Omega; L^r(U))$ (cf. Chapter A.1). For each n we assume that $X_n = \sum_{k=1}^{K_n} \mathbb{1}_{A_k^{(n)}} x_k^{(n)}$, where $A_k^{(n)} \in \mathscr{F}_T^\beta$ and $x_k^{(n)} \in L^r(U)$ for all $k = 1, \ldots, K_n$. By Lemma 3.26, there exist unique processes $\phi_k^{(n)} \in L^p_{\mathbb{F}^\beta}(\Omega; L^2([0,T]))$ such that

$$\mathbb{1}_{A_k^{(n)}} = \mathbb{E}\mathbb{1}_{A_k^{(n)}} + \int_0^T \phi_k^{(n)} \,\mathrm{d}\beta.$$

Now define $f_n := \sum_{k=1}^{K_n} \phi_k^{(n)} x_k^{(n)}$. Since

$$\begin{aligned}
\Big(\mathbb{E}\Big\| \Big(\int_0^T |f_n|^2 \,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_r^p\Big)^{\frac{1}{p}} &= \Big(\mathbb{E}\Big\| \Big(\int_0^T \Big|\sum_{k=1}^{K_n} \phi_k^{(n)} x_k^{(n)}\Big|^2 \,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_r^p\Big)^{\frac{1}{p}} \\
&\le \sum_{k=1}^{K_n} \Big(\mathbb{E}\Big(\int_0^T |\phi_k^{(n)}|^2 \,\mathrm{d}t\Big)^{\frac{p}{2}} \|x_k^{(n)}\|_r^p\Big)^{\frac{1}{p}} \\
&= \sum_{k=1}^{K_n} \|\phi_k^{(n)}\|_{L^p(\Omega; L^2([0,T]))} \|x_k^{(n)}\|_r < \infty
\end{aligned}$$

and every $\phi_k^{(n)}$ belongs to $L^p_{\mathbb{F}^\beta}(\Omega; L^2([0,T]))$, we infer that $f_n \in L^p_{\mathbb{F}^\beta}(\Omega; L^r_{\gamma,T})$. Moreover,

$$\begin{aligned}
X_n = \sum_{k=1}^{K_n} \mathbb{1}_{A_k^{(n)}} x_k^{(n)} &= \sum_{k=1}^{K_n} \Big(\mathbb{E}\mathbb{1}_{A_k^{(n)}} + \int_0^T \phi_k^{(n)} \,\mathrm{d}\beta\Big) x_k^{(n)} \\
&= \mathbb{E} \sum_{k=1}^{K_n} \mathbb{1}_{A_k^{(n)}} x_k^{(n)} + \int_0^T \sum_{k=1}^{K_n} \phi_k^{(n)} x_k^{(n)} \,\mathrm{d}\beta \\
&= \mathbb{E}X_n + \int_0^T f_n \,\mathrm{d}\beta.
\end{aligned}$$

Since $\lim_{n\to\infty} X_n = X$ in $L^p(\Omega; L^r(U))$, it holds that the sequence $\big(\int_0^T f_n \,\mathrm{d}\beta\big)_{n=1}^\infty$ converges in $L^p(\Omega; L^r(U))$, and therefore, the isomorphism $I^\beta_{\mathbb{F}^\beta}$ of Theorem 3.22 implies that the sequence $(f_n)_{n=1}^\infty$ is Cauchy in $L^p_{\mathbb{F}^\beta}(\Omega; L^r_{\gamma,T})$. Hence, there exists a unique limit $f \in L^p_{\mathbb{F}^\beta}(\Omega; L^r_{\gamma,T})$, which satisfies

$$\int_0^T f \,\mathrm{d}\beta = \lim_{n\to\infty} \int_0^T f_n \,\mathrm{d}\beta = \lim_{n\to\infty} X_n - \mathbb{E}X_n = X - \mathbb{E}X. \qquad \blacksquare$$

As a consequence of Theorem 3.22 and Theorem 3.27, we get the following corollary.

Corollary 3.28. *For $1 < p < \infty$, the stochastic integral defines an isomorphism of Banach spaces*

$$I^{\beta}_{\mathbb{F}^\beta} : L^p_{\mathbb{F}^\beta}(\Omega; L^r_{\gamma,T}) \simeq L^p_0(\Omega, \mathscr{F}^\beta_T; L^r(U)),$$

where $L^p_0(\Omega, \mathscr{F}^\beta_T; L^r(U))$ is the closed subspace of $L^p(\Omega, \mathscr{F}^\beta_T; L^r(U))$ consisting of all elements with mean 0.

Returning to a general filtration $\mathbb{F}$, which satisfies the properties stated after Definition 3.14, we finally extend Theorem 3.13.

Theorem 3.29. *Fix $1 < p < \infty$. For a measurable and adapted function $f\colon [0,T] \times \Omega \to L^r(U)$, which satisfies $\langle f, g\rangle \in L^p(\Omega; L^2([0,T]))$ for all $g \in L^{r'}(U)$, the following assertions are equivalent:*

(1) $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$;

(2) there exists a random variable $X \in L^p(\Omega; L^r(U))$ such that for all $g \in L^{r'}(U)$

$$\langle X, g\rangle = \int_0^T \langle f, g\rangle \,\mathrm{d}\beta \quad \text{almost surely.}$$

In this situation, we have $X = \int_0^T f \,\mathrm{d}\beta$ almost surely.

Proof. (1) $\Rightarrow$ (2): This can be proved in the same way as the corresponding implication of Theorem 3.13.

(2) $\Rightarrow$ (1): For $n \in \mathbb{N}$, let $P_n\colon L^r(U) \to L^r(U)$ be the operator from Lemma A.26. Let $h \in L^r(U)$, and recall that

$$\lim_{n\to\infty} P_n h = h \quad \text{in } L^r(U),$$

and that P_n has the representation

$$(P_n h)(u) = \sum_{j=1}^{N_n} \frac{1}{\mu(U_j^{(n)})} \int_{U_j^{(n)}} h \,\mathrm{d}\mu \; \mathbb{1}_{U_j^{(n)}}(u)$$

for suitable disjoint sets $U_j^{(n)} \subseteq U$, $j = 1, \ldots, N_n$.

Hence, by Fubini's theorem, we get for all $g \in L^{r'}(U)$

$$\begin{aligned}\langle P_n h, g\rangle &= \int_U \Big(\sum_{j=1}^{N_n} \frac{1}{\mu(U_j^{(n)})} \int_{U_j^{(n)}} h \,\mathrm{d}\mu \,\mathbb{1}_{U_j^{(n)}}\Big) g \,\mathrm{d}\mu \\ &= \int_U h \Big(\sum_{j=1}^{N_n} \frac{1}{\mu(U_j^{(n)})} \int_{U_j^{(n)}} g \,\mathrm{d}\mu \,\mathbb{1}_{U_j^{(n)}}\Big) \,\mathrm{d}\mu = \langle h, P_n g\rangle.\end{aligned}$$

Now take $f_n(t,\omega) := P_n f(t,\omega)$ for $(t,\omega) \in [0,T]\times\Omega$. Since $\langle f_n, g\rangle = \langle f, P_n g\rangle \in L^p_{\mathbb{F}}(\Omega; L^2([0,T]))$ for all $g \in L^{r'}(U)$, we obtain for each $k = 1,\dots,N_n$

$$(t,\omega) \mapsto \int_{U_k^{(n)}} f(t,\omega)\,\mathrm{d}\mu = \big\langle f_n(t,\omega), \mu(U_k^{(n)})\mathbb{1}_{U_k^{(n)}}\big\rangle \in L^p_{\mathbb{F}}(\Omega; L^2([0,T])),$$

where the equality follows from the fact that the sets $U_1^{(n)},\dots,U_{N_n}^{(n)}$ are disjoint. Thus, we get

$$\mathbb{E}\Big\| \int_{U_k^{(n)}} f \,\mathrm{d}\mu \,\mathbb{1}_{U_k^{(n)}} \Big\|^p_{L^r_{\gamma,T}} = \mu(U_k^{(n)})^{\frac{p}{r}} \mathbb{E}\Big\| \int_{U_k^{(n)}} f\,\mathrm{d}\mu \Big\|^p_{L^2([0,T])} < \infty.$$

And therefore, we have

$$f_n = \sum_{j=1}^{N_n} \frac{1}{\mu(U_j^{(n)})} \int_{U_j^{(n)}} f\,\mathrm{d}\mu \,\mathbb{1}_{U_j^{(n)}} \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$$

for each $n \in \mathbb{N}$. Hence, almost surely and for an arbitrary $g \in L^{r'}(U)$ we have the well-defined equalities

$$\begin{aligned}\int_0^T \langle f_n, g\rangle \,\mathrm{d}\beta &= \sum_{j=1}^{N_n} \frac{1}{\mu(U_j^{(n)})} \int_0^T \int_{U_j^{(n)}} f\,\mathrm{d}\mu\,\mathrm{d}\beta\,\langle \mathbb{1}_{U_j^{(n)}}, g\rangle \\ &= \Big\langle \int_0^T \sum_{j=1}^{N_n} \frac{1}{\mu(U_j^{(n)})} \int_{U_j^{(n)}} f\,\mathrm{d}\mu\,\mathbb{1}_{U_j^{(n)}}\,\mathrm{d}\beta, g\Big\rangle \\ &= \Big\langle \int_0^T f_n\,\mathrm{d}\beta, g\Big\rangle.\end{aligned}$$

By this estimate and (2), we obtain

$$\langle P_n X, g\rangle = \langle X, P_n g\rangle = \int_0^T \langle f, P_n g\rangle\,\mathrm{d}\beta = \int_0^T \langle f_n, g\rangle\,\mathrm{d}\beta = \Big\langle \int_0^T f_n\,\mathrm{d}\beta, g\Big\rangle$$

almost surely and for all $g \in L^{r'}(U)$. By Corollary A.8, this leads to

$$P_n X = \int_0^T f_n\,\mathrm{d}\beta \quad \text{almost surely.} \tag{$\triangle$}$$

Next, by the properties of the operators $(P_n)_{n=1}^{\infty}$ we mentioned above, almost surely we have $\lim_{n\to\infty} P_n X = X$ in $L^r(U)$, and since $X \in L^p(\Omega; L^r(U))$, the dominated convergence theorem yield

$$\lim_{n\to\infty} P_n X = X \quad \text{in } L^p(\Omega; L^r(U)).$$

Looking at (△), this convergence and Theorem 3.22 imply that $(f_n)_{n=1}^{\infty}$ is Cauchy in $L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, and thus has a limit $\widetilde{f}$. Similar as in the proof of Theorem 3.13, we can show that

$$\lim_{j\to\infty} \mathbb{1}_{U^{(N)}}(u) f_{n_j}(t,\omega,u) = \mathbb{1}_{U^{(N)}}(u) \widetilde{f}(t,\omega,u)$$

for almost all $(t,\omega,u) \in [0,T] \times \Omega \times U$, and moreover we have

$$\lim_{k\to\infty} \mathbb{1}_{U^{(N)}}(u) f_{n_{j_k}}(t,\omega,u) = \mathbb{1}_{U^{(N)}}(u) f(t,\omega,u)$$

for almost all $(t,\omega,u) \in [0,T] \times \Omega \times U$, since for all (t,ω) we have $\lim_{n\to\infty} f_n = f$ in $L^r(U)$. From this we infer that $f = \widetilde{f} \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ and this also shows that $X = \int_0^T f \, \mathrm{d}\beta$ almost surely. ■

3.3 The Integral Process

In this short section we want to investigate the properties of the integral process

$$t \mapsto \int_0^t f \, \mathrm{d}\beta, \quad t \in [0,T],$$

where $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ and $\mathbb{F} = (\mathscr{F}_t)_{t\in[0,T]}$ is a filtration satisfying the conditions we posted in the previous section. Here, we again assume that $1 < r < \infty$ is fixed.

Proposition 3.30. *For each $1 < p < \infty$ and every $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, the integral process $\left(\int_0^t f \, \mathrm{d}\beta\right)_{t\in[0,T]}$ is an L^p martingale with respect to the filtration $\mathbb{F}$.*

Proof. We first assume that f is an adapted step process of the form (3.2). Then, by Proposition 3.17, $\int_0^t f \, \mathrm{d}\beta$ is an $\mathscr{F}_t$-measurable and integrable random variable for every $t \in [0,T]$. Now define $v_n := \sum_{k=1}^{K_n} x_k^{(n)} \mathbb{1}_{A_k^{(n)}}$, and fix $0 \le s < t \le T$. Then there exists a $k \in \{1,\dots,N\}$ such that $t_k \le s < t_{k+1}$, and we can write

$$f = \sum_{n=1}^{k} v_n \mathbb{1}_{[t_{n-1},t_n)} + v_{k+1}(\mathbb{1}_{[t_k,s)} + \mathbb{1}_{[s,t_{k+1})}) + \sum_{n=k+2}^{N} v_n \mathbb{1}_{[t_{n-1},t_n)}.$$

For each $0 \le \tau_1 < \tau_2 \le T$, we obtain

$$\mathbb{E}[\beta(\tau_2) - \beta(\tau_1)|\mathscr{F}_{\tau_1}] = \mathbb{E}\big(\beta(\tau_2) - \beta(\tau_1)\big) = 0,$$

since $\beta(\tau_2) - \beta(\tau_1)$ is independent of $\mathscr{F}_{\tau_1}$ by the properties of the filtration $\mathbb{F}$ we assumed above. Additionally, v_{k+1} is $\mathscr{F}_s$-measurable since $t_k \le s$. Therefore, we have

$$\begin{aligned}
\mathbb{E}\Big[\int_0^t f \,\mathrm{d}\beta \big| \mathscr{F}_s\Big] &= \sum_{n=1}^{k} \mathbb{E}\big[v_n\big(\beta(t_n) - \beta(t_{n-1})\big)\big|\mathscr{F}_s\big] + \mathbb{E}\big[v_{k+1}\big(\beta(s) - \beta(t_k)\big)\big|\mathscr{F}_s\big] \\
&\quad + \mathbb{E}\big[v_{k+1}\big(\beta(t_{k+1}) - \beta(s)\big)\big|\mathscr{F}_s\big] + \sum_{n=k+2}^{N} \mathbb{E}\big[v_n\big(\beta(t_n) - \beta(t_{n-1})\big)\big|\mathscr{F}_s\big] \\
&= \sum_{n=1}^{k} v_n\big(\beta(t_n) - \beta(t_{n-1})\big) + v_{k+1}\big(\beta(s) - \beta(t_k)\big) + v_{k+1}\mathbb{E}[\beta(t_{k+1}) - \beta(s)|\mathscr{F}_s] \\
&\quad + \sum_{n=k+2}^{N} \mathbb{E}\Big[\mathbb{E}\big[v_n\big(\beta(t_n) - \beta(t_{n-1})\big)\big|\mathscr{F}_{t_{n-1}}\big]\Big|\mathscr{F}_s\Big] \\
&= \int_0^s f \,\mathrm{d}\beta + \sum_{n=k+2}^{N} \mathbb{E}\Big[v_n \mathbb{E}\big[\big(\beta(t_n) - \beta(t_{n-1})\big)\big|\mathscr{F}_{t_{n-1}}\big]\Big|\mathscr{F}_s\Big] = \int_0^s f \,\mathrm{d}\beta,
\end{aligned}$$

using that the sequence $(v_n)_{n=1}^N$ is predictable with respect to the filtration $(\mathscr{F}_{t_n})_{n=1}^N$.

Now, let $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, and let $(f_n)_{n=1}^\infty$ be an approximating sequence of adapted step processes for f. Since

$$\lim_{n\to\infty} \mathbb{1}_{[0,t)} f_n = \mathbb{1}_{[0,t)} f \quad \text{in } L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,t}),$$

Theorem 3.22 and Corollary 3.25 imply that

$$\lim_{n\to\infty} \int_0^t f_n \,\mathrm{d}\beta = \int_0^t f \,\mathrm{d}\beta \quad \text{in } L^p(\Omega, \mathscr{F}_t; L^r(U))$$

for each $t \in [0, T]$, which also yield that $\int_0^t f \,\mathrm{d}\beta$ is an $\mathscr{F}_t$-measurable and integrable random variable. Observe next that for each sub-σ-Algebra $\mathscr{G}$ of $\mathscr{A}$ the operator $\mathbb{E}[\,\cdot\,|\mathscr{G}] \colon L^p(\Omega; L^r(U)) \to L^p(\Omega, \mathscr{G}; L^r(U))$ is continuous. Thus, we obtain for $0 \le s < t \le T$

$$\mathbb{E}\Big[\int_0^t f \,\mathrm{d}\beta \big| \mathscr{F}_s\Big] = \lim_{n\to\infty} \mathbb{E}\Big[\int_0^t f_n \,\mathrm{d}\beta \big| \mathscr{F}_s\Big] = \lim_{n\to\infty} \int_0^s f_n \,\mathrm{d}\beta = \int_0^s f \,\mathrm{d}\beta$$

in $L^p(\Omega, \mathscr{F}_s; L^r(U))$. ■

Theorem 3.31. *For $1 < p < \infty$ and $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, the integral process $\left(\int_0^t f \, \mathrm{d}\beta\right)_{t\in[0,T]}$ has a continuous version satisfying the maximal inequality*

$$\mathbb{E}\Big\| \sup_{t\in[0,T]} \Big| \int_0^t f \, \mathrm{d}\beta \Big| \Big\|_r^p \lesssim_{p,r} \mathbb{E}\Big\| \int_0^T f \, \mathrm{d}\beta \Big\|_r^p.$$

Proof. Let $(f_n)_{n=1}^\infty$ be an approximating sequence of adapted step processes for f, and set $X_n(t) := \int_0^t f_n \, \mathrm{d}\beta$. Then, by the definition of the stochastic integral for adapted step processes and the properties of the Brownian motion (cf. Remark 1.23), X_n has a continuous version $\widetilde{X}_n \in L^p(\Omega; L^r(U; C([0,T])))$. By Proposition 3.30 and the Strong Doob inequality, we obtain for every choice of $0 = t_1 < \ldots < t_N = T$

$$\mathbb{E}\Big\|\sup_{j=1}^N |\widetilde{X}_n(t_j) - \widetilde{X}_m(t_j)|\Big\|_r^p \lesssim_{p,r} \mathbb{E}\|X_n(T) - X_m(T)\|_r^p.$$

Thus, by continuity and Fatou's lemma, we obtain

$$\mathbb{E}\Big\| \sup_{t\in[0,T]} |\widetilde{X}_n(t) - \widetilde{X}_m(t)|\Big\|_r^p \lesssim_{p,r} \mathbb{E}\|X_n(T) - X_m(T)\|_r^p,$$

which shows that the sequence $(\widetilde{X}_n)_{n=1}^\infty$ is Cauchy in $L^p(\Omega; L^r(U; C([0,T])))$, and therefore has a limit $\widetilde{X} \in L^p(\Omega; L^r(U; C([0,T]))) \subseteq C([0,T]; L^p(\Omega; L^r(U)))$. Since for all $t \in [0,T]$ we have $\lim_{n\to\infty} X_n(t) = \int_0^t f \, \mathrm{d}\beta$ and $\lim_{n\to\infty} \widetilde{X}_n(t) = \widetilde{X}(t)$ both in $L^p(\Omega; L^r(U))$, $\widetilde{X}$ defines a continuous version of the integral process.

Finally, the maximal inequality follows from the Strong Doob inequality in the same way as above by replacing $\widetilde{X}_n - \widetilde{X}_m$ with $\widetilde{X}$. ■

With this in mind, we get the following result.

Corollary 3.32 (Burkholder-Gundy inequality). *Let $1 < p < \infty$ and $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$. Then we have*

$$\mathbb{E}\Big\| \sup_{t\in[0,T]} \Big| \int_0^t f \, \mathrm{d}\beta \Big| \Big\|_r^p \eqsim_{p,r} \mathbb{E}\Big\| \Big(\int_0^T |f|^2 \, \mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r^p.$$

Proof. Since

$$\mathbb{E}\Big\| \sup_{t\in[0,T]} \Big| \int_0^t f \, \mathrm{d}\beta \Big| \Big\|_r^p \geq \mathbb{E}\Big\| \int_0^T f \, \mathrm{d}\beta \Big\|_r^p$$

is obvious, the assertion follows from Theorem 3.22 and Theorem 3.31. ■

Combining some of the previous results, we get the following representation theorem for a special class of martingales.

Theorem 3.33 (Martingale representation theorem). *Fix $1 < p < \infty$, and let $(M_t)_{t\in[0,T]}$ be an $L^r(U)$-valued L^p martingale with respect to the Brownian filtration $\mathbb{F}^\beta$. Then there exists a unique $f \in L^p_{\mathbb{F}^\beta}(\Omega; L^r_{\gamma,T})$ such that for all $t \in [0,T]$ we have*

$$M_t = M_0 + \int_0^t f \,\mathrm{d}\beta.$$

Especially, the function $t \mapsto M_t$ has a continuous version.

Proof. We have $M_T - M_0 \in L^p_0(\Omega, \mathscr{F}^\beta_T; L^r(U))$. Thus, Theorem 3.27 or Corollary 3.28, respectively, yield a unique $f \in L^p_{\mathbb{F}^\beta}(\Omega; L^r_{\gamma,T})$ such that

$$M_T = M_0 + \int_0^T f \,\mathrm{d}\beta.$$

Applying Proposition 3.30 leads to

$$M_t = \mathbb{E}[M_T|\mathscr{F}^\beta_t] = M_0 + \mathbb{E}\Big[\int_0^T f \,\mathrm{d}\beta \Big| \mathscr{F}^\beta_t\Big] = M_0 + \int_0^t f \,\mathrm{d}\beta.$$

And the last claim finally follows from Theorem 3.31. ■

3.4 Localization

So far the stochastic integral has only been defined for integrands that satisfy the integrability condition $\mathbb{E}\big\|\big(\int_0^T |f|^2 \,\mathrm{d}t\big)^{\frac{1}{2}}\big\|_r^p < \infty$. Unfortunately, most functions, even continuous ones, do not fulfill this property. So, to extend the stochastic integral in a most natural way, we are going to do a localization argument. As in the previous sections, we assume that $1 < r < \infty$ is fixed and that $\mathbb{F} = (\mathscr{F}_t)_{t\in[0,T]}$ is a filtration with the same properties we assumed above.

Definition 3.34. *A random variable $\tau\colon \Omega \to I \cup \{\infty\}$ with $I \in \{\mathbb{N}, \mathbb{R}_+\}$ is called a* stopping time *with respect to a filtration $(\mathscr{G}_i)_{i\in I}$ if*

$$\{\tau \le i\} \in \mathscr{G}_i \quad \textit{for all } i \in I.$$

Proposition 3.35. *For every* $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ *and every stopping time* τ *with values in* $[0,T]$ *we have* $\mathbb{1}_{[0,\tau)} f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ *and*

$$\int_0^\tau f \, \mathrm{d}\beta = \int_0^T \mathbb{1}_{[0,\tau)} f \, \mathrm{d}\beta \quad \text{almost surely.}$$

Proof. We first show that $\mathbb{1}_{[0,\tau)} f$ is adapted with respect to $\mathbb{F} = (\mathscr{F}_t)_{t \in [0,T]}$. For this purpose, we first observe that the function

$$\Phi_f \colon [0,T] \to L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T}), \quad \Phi_f(t) = \mathbb{1}_{[0,t)} f,$$

is continuous by the dominated convergence theorem. For $n \in \mathbb{N}$ we set

$$S(n) := \left\{ \tfrac{k}{2^n} T \colon k = 0, \dots, 2^n \right\} \quad \text{and} \quad \tau_n(\omega) := \min\{ t \in S(n) \colon t \geq \tau(\omega) \}.$$

Since for each $k = 0, \dots, 2^n$ and every $\frac{k}{2^n} T \leq t < \frac{k+1}{2^n} T$ we have

$$\{\tau_n \leq t\} = \{\tau_n \leq \tfrac{k}{2^n} T\} = \{\tau \leq \tfrac{k}{2^n} T\} \in \mathscr{F}_{\frac{k}{2^n} T} \subseteq \mathscr{F}_t,$$

we infer that τ_n is a stopping time with respect to the filtration $\mathbb{F}$. Note that $\tau_n(\omega) \geq \tau(\omega)$ and $\lim_{n\to\infty} \tau_n(\omega) = \tau(\omega)$ for all $\omega \in \Omega$. Moreover, we have for each $n \in \mathbb{N}$ and all $t \in [0,T]$

$$\mathbb{1}_{\{\tau_n \leq t\}} \mathbb{1}_{[0,\tau_n)} f = \sum_{k=0}^{2^n} \mathbb{1}_{\{\tau_n \leq t\}} \mathbb{1}_{\{\tau_n = \frac{k}{2^n} T\}} \mathbb{1}_{[0, \frac{k}{2^n} T)} f.$$

Thus, since f is an adapted process, $\mathbb{1}_{\{\tau_n \leq t\}} \mathbb{1}_{[0,\tau_n)} f$ is $\mathscr{F}_t$-measurable for all $t \in [0,T]$. This implies that

$$\mathbb{1}_{[0, t\wedge\tau_n)} f = \mathbb{1}_{\{\tau_n \leq t\}} \mathbb{1}_{[0,\tau_n)} f + \mathbb{1}_{\{\tau_n > t\}} \mathbb{1}_{[0,t)} f$$

is $\mathscr{F}_t$-measurable. By the continuity of the function Φ_f and the pointwise convergence of the stopping times, we deduce that

$$\lim_{n\to\infty} \mathbb{1}_{[0,t\wedge\tau_n(\omega))} f(\omega) = \mathbb{1}_{[0,t\wedge\tau(\omega))} f(\omega) = \mathbb{1}_{[0,t)} \mathbb{1}_{[0,\tau(\omega))} f(\omega)$$

for almost all $\omega \in \Omega$. Hence, $\mathbb{1}_{[0,t)} \mathbb{1}_{[0,\tau)} f$ is $\mathscr{F}_t$-measurable for all $t \in [0,T]$. So far we have shown that $\mathbb{1}_{[0,\tau)} f$ is adapted to the filtration $\mathbb{F}$. The first assertion then follows by

$$\mathbb{E} \Big\| \Big(\int_0^T \mathbb{1}_{[0,\tau)} |f|^2 \, \mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r^p \leq \mathbb{E} \Big\| \Big(\int_0^T |f|^2 \, \mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r^p < \infty.$$

For the integral identity, we first assume that f is an adapted step process of the form (3.2) with the additional assumption that $t_j := \frac{j}{2^n}T$ for $j = 1, \dots, 2^n$ and $n \in \mathbb{N}$ fixed. By taking $v_j := \sum_{m=1}^{M_j} x_m^{(j)} \mathbb{1}_{A_m^{(j)}}$ we thus can write

$$f = \sum_{j=1}^{2^n} \mathbb{1}_{[\frac{j-1}{2^n}T, \frac{j}{2^n}T)} \sum_{m=1}^{M_j} x_m^{(j)} \mathbb{1}_{A_m^{(j)}} = \sum_{j=1}^{2^n} \mathbb{1}_{[t_{j-1}, t_j)} v_j.$$

Now observe that

$$\begin{aligned}
\mathbb{1}_{[0,\tau_n)} f &= \sum_{j=1}^{2^n} \mathbb{1}_{[0,\tau_n)} \mathbb{1}_{[t_{j-1},t_j)} v_j = \sum_{j=1}^{2^n} \sum_{k=1}^{2^n} \mathbb{1}_{\{\tau_n = t_k\}} \mathbb{1}_{[0,t_k)} \mathbb{1}_{[t_{j-1},t_j)} v_j \\
&= \sum_{j=1}^{2^n} \sum_{k=j}^{2^n} \mathbb{1}_{\{\tau_n = t_k\}} \mathbb{1}_{[t_{j-1},t_j)} v_j = \sum_{j=1}^{2^n} \mathbb{1}_{[t_{j-1},t_j)} \mathbb{1}_{\{\tau_n \ge t_j\}} v_j.
\end{aligned}$$

Hence, since $\{\tau_n \ge t_j\} = \{\tau_n \le t_{j-1}\}^C \in \mathscr{F}_{t_{j-1}}$, the process $\mathbb{1}_{[0,\tau_n)} f$ is still an adapted step process. With this in mind, we obtain

$$\begin{aligned}
\int_0^T \mathbb{1}_{[0,\tau_n)} f \, \mathrm{d}\beta &= \sum_{j=1}^{2^n} \mathbb{1}_{\{\tau_n \ge t_j\}} v_j \big(\beta(t_j) - \beta(t_{j-1})\big) \\
&= \sum_{j=1}^{2^n} \sum_{k=j}^{2^n} \mathbb{1}_{\{\tau_n = t_k\}} v_j \big(\beta(t_j) - \beta(t_{j-1})\big) \\
&= \sum_{k=1}^{2^n} \mathbb{1}_{\{\tau_n = t_k\}} \sum_{j=1}^{k} v_j \big(\beta(t_j) - \beta(t_{j-1})\big) \\
&= \sum_{k=1}^{2^n} \mathbb{1}_{\{\tau_n = t_k\}} \int_0^{t_k} f \, \mathrm{d}\beta \\
&= \int_0^{\tau_n} f \, \mathrm{d}\beta.
\end{aligned}$$

Now let $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ and $(f_n)_{n=1}^\infty$ be an approximating sequence of adapted step processes for f. As seen in the proof of Lemma 3.21, we can choose these step processes with dyadic time steps, as assumed above. Then

$$\lim_{n\to\infty} \big\| \mathbb{1}_{[0,\tau_n)} (f_n - f) \big\|_{L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})} \le \lim_{n\to\infty} \big\| f_n - f \big\|_{L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})} = 0,$$

and, using the continuity of Φ_f, this leads to

$$\begin{aligned}
&\lim_{n\to\infty} \big\| \mathbb{1}_{[0,\tau_n)} f_n - \mathbb{1}_{[0,\tau)} f \big\|_{L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})} \\
&\quad \le \lim_{n\to\infty} \big(\| \mathbb{1}_{[0,\tau_n)} f_n - \mathbb{1}_{[0,\tau_n)} f \|_{L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})} + \| \mathbb{1}_{[0,\tau_n)} f - \mathbb{1}_{[0,\tau)} f \|_{L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})} \big) = 0.
\end{aligned}$$

Therefore,

$$\lim_{n\to\infty}\int_0^T \mathbb{1}_{[0,\tau_n)} f_n \,\mathrm{d}\beta = \int_0^T \mathbb{1}_{[0,\tau)} f \,\mathrm{d}\beta \quad \text{in } L^p(\Omega; L^r(U)).$$

By Theorem 3.31 we furthermore deduce that

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\Big\|\int_0^{\tau_n} f_n\,\mathrm{d}\beta - \int_0^{\tau_n} f\,\mathrm{d}\beta\Big\|_r^p &\le \lim_{n\to\infty}\mathbb{E}\Big\|\sup_{t\in[0,T]}\Big|\int_0^t f_n\,\mathrm{d}\beta - \int_0^t f\,\mathrm{d}\beta\Big|\Big\|_r^p \\ &\lesssim_{p,r} \lim_{n\to\infty}\mathbb{E}\Big\|\int_0^T f_n\,\mathrm{d}\beta - \int_0^T f\,\mathrm{d}\beta\Big\|_r^p = 0.\end{aligned}$$

Thus, we have

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{E}\Big\|\int_0^{\tau_n} f_n\,\mathrm{d}\beta - \int_0^{\tau} f\,\mathrm{d}\beta\Big\|_r^p \\ &\quad\le \lim_{n\to\infty}\Big(\mathbb{E}\Big\|\int_0^{\tau_n} f_n\,\mathrm{d}\beta - \int_0^{\tau_n} f\,\mathrm{d}\beta\Big\|_r^p + \mathbb{E}\Big\|\int_0^{\tau_n} f\,\mathrm{d}\beta - \int_0^{\tau} f\,\mathrm{d}\beta\Big\|_r^p\Big) = 0,\end{aligned}$$

using the continuity of the stochastic integral process and the pointwise convergence of the stopping times for the second term. We finally conclude that

$$\int_0^\tau f\,\mathrm{d}\beta = \lim_{n\to\infty}\int_0^{\tau_n} f_n\,\mathrm{d}\beta = \lim_{n\to\infty}\int_0^T \mathbb{1}_{[0,\tau_n)} f_n\,\mathrm{d}\beta = \int_0^T \mathbb{1}_{[0,\tau)} f\,\mathrm{d}\beta$$

in $L^p(\Omega; L^r(U))$. ■

Lemma 3.36. *If $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, then for all $\delta > 0$ and all $\varepsilon > 0$ we have*

$$\mathbb{P}\Big(\Big\|\sup_{t\in[0,T]}\Big|\int_0^t f\,\mathrm{d}\beta\Big|\Big\|_r > \varepsilon\Big) \le \frac{c_{p,r}\delta^p}{\varepsilon^p} + \mathbb{P}\big(\|f\|_{L^r_{\gamma,T}} \ge \delta\big)$$

and

$$\mathbb{P}\big(\|f\|_{L^r_{\gamma,T}} > \varepsilon\big) \le \frac{c_{p,r}\delta^p}{\varepsilon^p} + \mathbb{P}\Big(\Big\|\sup_{t\in[0,T]}\Big|\int_0^t f\,\mathrm{d}\beta\Big|\Big\|_r \ge \delta\Big),$$

where $c_{p,r}$ is the constant from Theorem 3.22.

Proof. By Proposition 3.35 and Theorem 3.22 we have

$$\mathbb{E}\Big\|\sup_{t\in[0,\tau]}\Big|\int_0^t f\,\mathrm{d}\beta\Big|\Big\|_r^p = \mathbb{E}\Big\|\sup_{t\in[0,T]}\Big|\int_0^{t\wedge\tau} f\,\mathrm{d}\beta\Big|\Big\|_r^p \le c_{p,r}\,\mathbb{E}\|\mathbb{1}_{[0,\tau)} f\|^p_{L^r_{\gamma,T}}$$

for all stopping times τ.

Let δ, $\varepsilon > 0$, and define

$$\mu(\omega) := T \wedge \inf\Big\{t \in [0,T]\colon \Big\| \sup_{s\in[0,t]}\Big|\Big(\int_0^s f\,\mathrm{d}\beta\Big)(\omega)\Big|\Big\|_r \geq \varepsilon\Big\},$$
$$\nu(\omega) := T \wedge \inf\big\{t \in [0,T]\colon \|\mathbb{1}_{[0,t)} f(\omega)\|_{L^r_{\gamma,T}} \geq \delta\big\}.$$

Now take $\tau := \mu \wedge \nu$. Then τ is a stopping time and

$$\mathbb{E}\Big\| \sup_{t\in[0,T]}\Big|\int_0^{t\wedge\tau} f\,\mathrm{d}\beta\Big|\Big\|_r^p \leq \varepsilon^p \quad \text{and} \quad \mathbb{E}\|\mathbb{1}_{[0,\tau)} f\|^p_{L^r_{\gamma,T}} \leq \delta^p,$$

since $t \mapsto \mathbb{1}_{[0,t)} f$ and $t \mapsto \sup_{s\in[0,t]}\big|\int_0^s f\,\mathrm{d}\beta\big|$ have continuous paths starting at zero. Chebyshev's inequality now leads to

$$\begin{aligned}
&\mathbb{P}\Big(\Big\|\sup_{t\in[0,T]}\Big|\int_0^t f\,\mathrm{d}\beta\Big|\Big\|_r > \varepsilon, \Big\|\sup_{t\in[0,T]} \mathbb{1}_{[0,t)}|f|\Big\|_{L^r_{\gamma,T}} < \delta\Big)\\
&\quad\leq \mathbb{P}\Big(\Big\|\sup_{t\in[0,\tau]}\Big|\int_0^t f\,\mathrm{d}\beta\Big|\Big\|_r \geq \varepsilon\Big) \leq \frac{1}{\varepsilon^p}\mathbb{E}\Big\|\sup_{t\in[0,\tau]}\Big|\int_0^t f\,\mathrm{d}\beta\Big|\Big\|_r^p\\
&\quad\leq \frac{c_{p,r}}{\varepsilon^p}\mathbb{E}\|\mathbb{1}_{[0,\tau)} f\|^p_{L^r_{\gamma,T}} \leq \frac{c_{p,r}\delta^p}{\varepsilon^p}.
\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
&\mathbb{P}\Big(\Big\|\sup_{t\in[0,T]}\Big|\int_0^t f\,\mathrm{d}\beta\Big|\Big\|_r > \varepsilon\Big)\\
&\quad= \mathbb{P}\Big(\Big\|\sup_{t\in[0,T]}\Big|\int_0^t f\,\mathrm{d}\beta\Big|\Big\|_r > \varepsilon, \Big\|\sup_{t\in[0,T]} \mathbb{1}_{[0,t)}|f|\Big\|_{L^r_{\gamma,T}} < \delta\Big)\\
&\qquad+ \mathbb{P}\Big(\Big\|\sup_{t\in[0,T]}\Big|\int_0^t f\,\mathrm{d}\beta\Big|\Big\|_r > \varepsilon, \Big\|\sup_{t\in[0,T]} \mathbb{1}_{[0,t)}|f|\Big\|_{L^r_{\gamma,T}} \geq \delta\Big)\\
&\quad\leq \frac{c_{p,r}\delta^p}{\varepsilon^p} + \mathbb{P}\big(\big\|\sup_{t\in[0,T]} \mathbb{1}_{[0,t)}|f|\big\|_{L^r_{\gamma,T}} \geq \delta\big)\\
&\quad= \frac{c_{p,r}\delta^p}{\varepsilon^p} + \mathbb{P}\big(\|f\|_{L^r_{\gamma,T}} \geq \delta\big),
\end{aligned}$$

where the last inequality follows from the fact that for each $\omega \in \Omega$ and all $t \in [0,T]$ we have $\mathbb{1}_{[0,t)}|f(\omega)| \leq |f(\omega)|$ with equality for $t = T$, and therefore, $\|f(\omega)\|_{L^r_{\gamma,T}} = \|\sup_{t\in[0,T]} \mathbb{1}_{[0,t)}|f(\omega)|\|_{L^r_{\gamma,T}}$.

The second inequality follows in the exact same way by interchanging the two processes. Observe that in this case we still have

$$\mathbb{E}\|\mathbb{1}_{[0,\tau)} f\|^p_{L^r_{\gamma,T}} \leq c_{p,r}\,\mathbb{E}\Big\|\sup_{t\in[0,\tau]}\Big|\int_0^t f\,\mathrm{d}\beta\Big|\Big\|_r^p$$

by Proposition 3.35 and Theorem 3.22. ■

For a separable Banach space E, we denote by $L^0(\Omega; E)$ the vector space of all equivalence classes of measurable functions on Ω with values in the Banach space E which are identical almost surely. Moreover, we define the map

$$d_{\mathbb{P}}\colon L^0(\Omega;E)\times L^0(\Omega;E)\to[0,\infty),\quad d_{\mathbb{P}}(X,Y)=\mathbb{E}(\|X-Y\|_E\wedge 1).$$

Proposition 3.37. *The pairing $(L^0(\Omega;E), d_{\mathbb{P}})$ is a complete metric space, and convergence with respect to this metric coincides with convergence in probability.*

Proof. **(1)** Let $X, Y \in L^0(\Omega; E)$. Then, of course, $d_{\mathbb{P}}$ is well-defined and symmetric. Next, if $d_{\mathbb{P}}(X,Y) = 0$, then $\|X-Y\|_E \wedge 1 = 0$ almost surely, which implies that $X = Y$ almost surely. The converse direction is trivial. Finally, using that

$$(a+b)\wedge 1 \le a\wedge 1 + b\wedge 1$$

for all $a, b \geq 0$, we deduce that

$$\begin{aligned}
d_{\mathbb{P}}(X,Y) &\le \mathbb{E}\big((\|X-Z\|_E+\|Z-Y\|_E)\wedge 1\big)\\
&\le \mathbb{E}\big((\|X-Z\|_E\wedge 1)+(\|Z-Y\|_E\wedge 1)\big)\\
&= d_{\mathbb{P}}(X,Z)+d_{\mathbb{P}}(Z,Y)
\end{aligned}$$

for all $Z \in L^0(\Omega; E)$. Therefore, $(L^0(\Omega;E), d_{\mathbb{P}})$ is a metric space.

(2) Next, assume that $(X_n)_{n=1}^\infty \subseteq L^0(\Omega;E)$ with $\lim_{n\to\infty} d_{\mathbb{P}}(X_n, X) = 0$. Then for every $\varepsilon \in (0,1)$ we obtain

$$\begin{aligned}
\lim_{n\to\infty}\mathbb{P}\big(\|X_n-X\|_E>\varepsilon\big) &= \lim_{n\to\infty}\frac{1}{\varepsilon}\mathbb{E}\mathbb{1}_{\{\|X_n-X\|_E>\varepsilon\}}(\varepsilon\wedge 1)\\
&\le \lim_{n\to\infty}\frac{1}{\varepsilon}\mathbb{E}(\|X_n-X\|_E\wedge 1)=\lim_{n\to\infty}\frac{1}{\varepsilon}d_{\mathbb{P}}(X_n,X)=0.
\end{aligned}$$

Now, if $\lim_{n\to\infty} X_n = X$ in probability, we deduce that for all $\varepsilon \in (0,1)$

$$\begin{aligned}
&\lim_{n\to\infty} d_{\mathbb{P}}(X_n,X)\\
&\quad=\lim_{n\to\infty}\int_{\{\|X_n-X\|_E>\varepsilon\}}\|X_n-X\|_E\wedge 1\,\mathrm{d}\mathbb{P}+\int_{\{\|X_n-X\|_E\le\varepsilon\}}\|X_n-X\|_E\wedge 1\,\mathrm{d}\mathbb{P}\\
&\quad\le\lim_{n\to\infty}\int_{\{\|X_n-X\|_E>\varepsilon\}}1\,\mathrm{d}\mathbb{P}+\int_{\{\|X_n-X\|_E\le\varepsilon\}}\|X_n-X\|_E\,\mathrm{d}\mathbb{P}\\
&\quad\le\lim_{n\to\infty}\mathbb{P}\big(\|X_n-X\|_E>\varepsilon\big)+\varepsilon=\varepsilon.
\end{aligned}$$

Since ε was arbitrary, we get $\lim_{n\to\infty} d_{\mathbb{P}}(X_n, X) = 0$. And this shows the desired convergence result.

(3) At last we show the completeness of the metric space. Let $(X_n)_{n=1}^\infty \subseteq L^0(\Omega; E)$ be Cauchy with respect to $d_\mathbb{P}$. Then, by the foregoing result, the sequence $(X_n)_{n=1}^\infty$ is 'Cauchy in probability', that is

$$\lim_{n,m\to\infty} \mathbb{P}\big(\|X_n - X_m\|_E > \varepsilon\big) = 0 \quad \text{for all } \varepsilon > 0.$$

Therefore, we can find a sequence of uprising integers $(k_n)_{n=1}^\infty$ with $\lim_{n\to\infty} k_n = \infty$ such that

$$\mathbb{P}\big(\|X_{k_n} - X_{k_n+m}\|_E > 2^{-n}\big) < 2^{-n} \quad \text{for all } n, m \geq 1,$$

which leads to

$$\sum_{n=1}^\infty \mathbb{P}\big(\|X_{k_n} - X_{k_{n+1}}\|_E > 2^{-n}\big) < \infty.$$

By the Borel-Cantelli lemma we get $\|X_{k_n}(\omega) - X_{k_{n+1}}(\omega)\|_E \leq 2^{-n}$ for all $\omega \in \Omega_0$, where $\Omega_0 \subseteq \Omega$ with $\mathbb{P}(\Omega_0) = 1$. Hence, for each fixed $\omega \in \Omega_0$, the sequence $\big(X_{k_n}(\omega)\big)_{n=1}^\infty$ is Cauchy in E and therefore converges to a $X(\omega) \in E$. Thus, the sequence $(X_{k_n})_{n=1}^\infty$ converges almost surely to the random variable X. Then it does so also in probability and, by the foregoing result, also in the metric $d_\mathbb{P}$. But, since $(X_n)_{n=1}^\infty$ is Cauchy with respect to $d_\mathbb{P}$, we infer that

$$\lim_{n\to\infty} d_\mathbb{P}(X_n, X) \leq \lim_{n\to\infty} d_\mathbb{P}(X_n, X_{k_n}) + d_\mathbb{P}(X_{k_n}, X) = 0,$$

which means that $d_\mathbb{P}$ is complete. ∎

As in Definition 3.19, we denote by $L^0_\mathbb{F}(\Omega; L^r_{\gamma,T})$ the space of all *adapted processes*, i.e., functions $f \in L^0(\Omega; L^r_{\gamma,T})$ with

$$\mathbb{1}_{[0,t)} f \in L^0(\Omega, \mathscr{F}_t; L^r_{\gamma,t}) \quad \text{for all } t \in [0,T],$$

which means that f has a representative $\widetilde{f}\colon [0,T] \times \Omega \times U \to \mathbb{R}$ such that $\mathbb{1}_{[0,t)}\widetilde{f}$ is $\mathscr{B}([0,t)) \otimes \mathscr{F}_t \otimes \Sigma$-measurable.

Definition 3.38. *Let $(\tau_n)_{n=1}^\infty$ be a sequence of stopping times with respect to $\mathbb{F}$ and with values in $[0,T]$, and let $f \in L^0_\mathbb{F}(\Omega; L^r_{\gamma,T})$. Then we call the sequence $(\tau_n)_{n=1}^\infty$* a localizing sequence *for f if*

(1) for all $\omega \in \Omega$ there exists an index $N(\omega) \in \mathbb{N}$ such that $\tau_n(\omega) = T$ for all $n \geq N(\omega)$, and

(2) $\mathbb{1}_{[0,\tau_n)} f \in L^p_\mathbb{F}(\Omega; L^r_{\gamma,T})$ for all $n \in \mathbb{N}$.

Remark 3.39. **(1)** For each $f \in L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ and every stopping time τ with respect to $\mathbb{F}$, the random variable $\mathbb{1}_{[0,\tau)} f$ is adapted to $\mathbb{F}$. This was already shown in the proof of Proposition 3.35.

(2) For every $f \in L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, a localizing sequence $(\tau_n)_{n=1}^\infty$ is given by

$$\tau_n(\omega) := \inf\big\{s \in [0,T] \colon \|\mathbb{1}_{[0,s)} f(\omega)\|_{L^r_{\gamma,T}} \geq n\big\},$$

where we take $\tau_n(\omega) := T$ if the infimum is taken over the empty set. Let us prove this assertion. If we fix an $\omega \in \Omega$ we have

$$\|\mathbb{1}_{[0,s)} f(\omega)\|_{L^r_{\gamma,T}} \leq \|f(\omega)\|_{L^r_{\gamma,T}} < \infty.$$

Thus, there exists an index $N(\omega)$ such that $\|\mathbb{1}_{[0,s)} f(\omega)\|_{L^r_{\gamma,T}} \leq N(\omega)$ for all $0 \leq s \leq T$. Hence, for each $n \geq N(\omega)$ we have $\tau_n(\omega) = T$. Moreover, we have $\|\mathbb{1}_{[0,\tau_n(\omega))} f(\omega)\|_{L^r_{\gamma,T}} \leq n$, and therefore

$$\mathbb{E}\|\mathbb{1}_{[0,\tau_n)} f\|^p_{L^r_{\gamma,T}} \leq n^p < \infty,$$

which, combined with the first remark, gives the second property. ■

For any $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ we recall from Theorem 3.31 that

$$t \mapsto I^\beta_{\mathbb{F}}(\mathbb{1}_{[0,t)} f) = \int_0^t f \,\mathrm{d}\beta$$

has a continuous version. With this in mind, we can give the following localized version of Theorem 3.22.

Theorem 3.40 (Itô homeomorphism). *The mapping $I^\beta_{\mathbb{F}}$ has a unique extension to a linear mapping*

$$I^\beta_{\mathbb{F},loc} \colon L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T}) \to L^0(\Omega; L^r(U; C([0,T]))),$$

which is a homeomorphism onto its closed range. Moreover, the estimates from Lemma 3.36 extend to arbitrary elements $f \in L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$.

Proof. Let $f \in L^0(\Omega; L^r_{\gamma,T})$, and let $(\tau_n)_{n=1}^\infty$ be the localizing sequence for f from Remark 3.39 (2). Then,

$$f_n := \mathbb{1}_{[0,\tau_n)} f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T}) \quad \text{for all } 1 < p < \infty,$$

and by Theorem 3.31,

$$I_{\mathbb{F}}^{\beta}(f_n) \in L^p(\Omega; L^r(U; C([0,T]))) \subseteq L^0(\Omega; L^r(U; C([0,T])))$$

for each $n \in \mathbb{N}$. Since $\lim_{n\to\infty} f_n = f$ almost surely, the sequence $(f_n)_{n=1}^{\infty}$ is 'Cauchy in probability' and, by Lemma 3.36 and Proposition 3.37, we deduce that $\big(I_{\mathbb{F}}^{\beta}(f_n)\big)_{n=1}^{\infty}$ is Cauchy in $L^0(\Omega; L^r(U; C([0,T])))$. Hence, there exists a limit $X \in L^0(\Omega; L^r(U; C([0,T])))$. Now define

$$I_{\mathbb{F},\mathrm{loc}}^{\beta}(f) := X = \lim_{n\to\infty} I_{\mathbb{F}}^{\beta}(f_n) \quad \text{in } L^0(\Omega; L^r(U; C([0,T]))).$$

Then $I_{\mathbb{F},\mathrm{loc}}^{\beta}$ is well-defined. Additionally, by Proposition 3.37, we have

$$\lim_{n\to\infty} I_{\mathbb{F}}^{\beta}(f_n) = I_{\mathbb{F},\mathrm{loc}}^{\beta}(f) \quad \text{in probability,}$$

and thus, by passing to a subsequence, we have

$$\lim_{k\to\infty} I_{\mathbb{F}}^{\beta}(f_{n_k}) = I_{\mathbb{F},\mathrm{loc}}^{\beta}(f) \quad \text{almost surely.}$$

Now let $\varepsilon > 0$. Then, by the σ-continuity of the probability measure $\mathbb{P}$, we have

$$\lim_{k\to\infty} \mathbb{P}\Big(\big\| \sup_{t\in[0,T]} |I_{\mathbb{F}}^{\beta}(\mathbb{1}_{[0,t)} f_{n_k})| \big\|_r > \varepsilon\Big) = \mathbb{P}\Big(\big\| \sup_{t\in[0,T]} |I_{\mathbb{F},\mathrm{loc}}^{\beta}(f)(t)| \big\|_r > \varepsilon\Big),$$

and similarly, since $\lim_{k\to\infty} f_{n_k} = f$ almost surely, we obtain for any $\delta > 0$

$$\lim_{k\to\infty} \mathbb{P}\big(\|f_{n_k}\|_{L^r_{\gamma,T}} \geq \delta\big) = \mathbb{P}\big(\|f\|_{L^r_{\gamma,T}} \geq \delta\big).$$

Next, applying Lemma 3.36 to each f_{n_k}, we get

$$\begin{aligned}
\mathbb{P}\Big(\big\| \sup_{t\in[0,T]} |I_{\mathbb{F},\mathrm{loc}}^{\beta}(f)(t)| \big\|_r > \varepsilon\Big) &= \lim_{k\to\infty} \mathbb{P}\Big(\big\| \sup_{t\in[0,T]} |I_{\mathbb{F}}^{\beta}(\mathbb{1}_{[0,t)} f_{n_k})| \big\|_r > \varepsilon\Big) \\
&\leq \frac{c_{p,r}\delta^p}{\varepsilon^p} + \lim_{k\to\infty} \mathbb{P}\big(\|f_{n_k}\|_{L^r_{\gamma,T}} \geq \delta\big) \\
&= \frac{c_{p,r}\delta^p}{\varepsilon^p} + \mathbb{P}\big(\|f\|_{L^r_{\gamma,T}} \geq \delta\big).
\end{aligned}$$

The other inequality from Lemma 3.36 can be extended in the exact same way. From this, combined with Proposition 3.37, we infer that $I_{\mathbb{F},\mathrm{loc}}^{\beta}$ is continuous and has a continuous inverse. This also shows, that the mapping $I_{\mathbb{F},\mathrm{loc}}^{\beta}$ has a closed range in $L^0(\Omega; L^r(U; C([0,T])))$. ∎

Now we are prepared to define the stochastic integral for adapted processes.

Definition 3.41. *Let $f \in L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ and $(f_n)_{n=1}^\infty$ be an approximating sequence of adapted L^p processes. Then we define the* stochastic integral *of f by*

$$\int_0^{\cdot} f \,\mathrm{d}\beta := I^\beta_{\mathbb{F},loc}(f) = \lim_{n\to\infty} \int_0^{\cdot} f_n \,\mathrm{d}\beta,$$

where convergence holds in $L^0(\Omega; L^r(U; C([0,T])))$.

Remark 3.42. (1) The stochastic integral is well-defined in the sense that it is independent of the approximating sequence.

(2) If $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ for some $1 < p < \infty$, then the integral process which arises from the Itô integral coincides almost surely with the 'localized' stochastic integral process.

(3) This new stochastic integral process is no longer a martingale, but a so called *local martingale* (cf. [10] for more information). ■

Proposition 3.43. *For every $f \in L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, each $t \in [0,T]$, and any stopping time τ with respect to $\mathbb{F}$ and with values in $[0,T]$, we have*

$$\int_0^{t\wedge\tau} f \,\mathrm{d}\beta = \int_0^t \mathbb{1}_{[0,\tau)} f \,\mathrm{d}\beta \quad \textit{almost surely.}$$

Proof. Let $(\tau_n)_{n=1}^\infty$ be the localizing sequence of Remark 3.39 (2), and define

$$f_n := \mathbb{1}_{[0,\tau_n)} f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T}).$$

Then $\lim_{n\to\infty} f_n = f$ almost surely and in $L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$. Therefore, by Theorem 3.40, we obtain

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\Big(\Big\| \int_0^{t\wedge\tau} f_n \,\mathrm{d}\beta - \int_0^{t\wedge\tau} f \,\mathrm{d}\beta \Big\|_r \wedge 1\Big) \\
&\quad \le \lim_{n\to\infty} \mathbb{E}\Big(\Big\| \sup_{s\in[0,T]} \Big| \int_0^s f_n \,\mathrm{d}\beta - \int_0^s f \,\mathrm{d}\beta \Big| \Big\|_r \wedge 1\Big) = 0.
\end{aligned}$$

And since $\lim_{n\to\infty} \mathbb{1}_{[0,\tau)} f_n = \mathbb{1}_{[0,\tau)} f$ in $L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, we also have for all $t \in [0,T]$

$$\lim_{n\to\infty} \int_0^t \mathbb{1}_{[0,\tau)} f_n \,\mathrm{d}\beta = \int_0^t \mathbb{1}_{[0,\tau)} f \,\mathrm{d}\beta \quad \text{in } L^0(\Omega; L^r(U)).$$

Finally, applying Proposition 3.35 and Corollary 3.25 to each f_n, we get

$$\begin{aligned}\int_0^{t\wedge\tau} f\,\mathrm{d}\beta &= \lim_{n\to\infty}\int_0^{t\wedge\tau} f_n\,\mathrm{d}\beta = \lim_{n\to\infty}\int_0^T \mathbb{1}_{[0,t\wedge\tau)} f_n\,\mathrm{d}\beta \\ &= \lim_{n\to\infty}\int_0^T \mathbb{1}_{[0,t)}\mathbb{1}_{[0,\tau)} f_n\,\mathrm{d}\beta \\ &= \lim_{n\to\infty}\int_0^t \mathbb{1}_{[0,\tau)} f_n\,\mathrm{d}\beta \\ &= \int_0^t \mathbb{1}_{[0,\tau)} f\,\mathrm{d}\beta\end{aligned}$$

in $L^0(\Omega; L^r(U))$. ■

Theorem 3.44 (Burkholder-Gundy inequality). *Let $1 < p < \infty$ and $f \in L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$. Then we have*

$$\mathbb{E}\Big\| \sup_{t\in[0,T]} \Big| \int_0^t f\,\mathrm{d}\beta \Big| \Big\|_r^p \eqsim_{p,r} \mathbb{E}\Big\| \Big(\int_0^T |f|^2\,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_r^p.$$

This is understood in the sense that the left-hand side is finite if and only if the right-hand side is finite.

Proof. If the right-hand side is finite, then $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, and therefore the inequalities follow from Corollary 3.32. Next, assume the left-hand side is finite. Let $(\tau_n)_{n=1}^\infty$ be the localizing sequence from Remark 3.39 (2), and define

$$f_n := \mathbb{1}_{[0,\tau_n)} f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T}).$$

Then $\lim_{n\to\infty} f_n = f$ almost surely. And moreover, by passing to a subsequence, $\lim_{k\to\infty} \int_0^{\tau_{n_k}} f\,\mathrm{d}\beta = \int_0^T f\,\mathrm{d}\beta$ almost surely. Thus, the dominated convergence theorem (here we use the assumption), Proposition 3.43, Theorem 3.22, and Fatou's lemma yield

$$\begin{aligned}\mathbb{E}\Big\|\int_0^T f\,\mathrm{d}\beta\Big\|_r^p &= \lim_{k\to\infty}\mathbb{E}\Big\|\int_0^{\tau_{n_k}} f\,\mathrm{d}\beta\Big\|_r^p = \lim_{k\to\infty}\mathbb{E}\Big\|\int_0^T \mathbb{1}_{[0,\tau_{n_k})} f\,\mathrm{d}\beta\Big\|_r^p \\ &\eqsim_{p,r} \liminf_{k\to\infty}\mathbb{E}\Big\|\Big(\int_0^T |f_{n_k}|^2\,\mathrm{d}t\Big)^{\frac{1}{2}}\Big\|_r^p \\ &\geq \mathbb{E}\Big\|\Big(\int_0^T |f|^2\,\mathrm{d}t\Big)^{\frac{1}{2}}\Big\|_r^p.\end{aligned}$$

This shows that $f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, and the result again follows from Corollary 3.32. ■

Finally, we extend Theorem 3.29 to the localized case.

Theorem 3.45. *For a measurable and adapted process $f\colon [0,T]\times\Omega\to L^r(U)$, which satisfies $\langle f,g\rangle\in L^0(\Omega;L^2([0,T]))$ for all $g\in L^{r'}(U)$, the following assertions are equivalent:*

(1) $f\in L^0_{\mathbb{F}}(\Omega;L^r_{\gamma,T})$;

(2) there exists a process $\zeta\in L^0(\Omega;L^r(U;C([0,T])))$ *such that for all* $g\in L^{r'}(U)$ *we have*

$$\langle\zeta,g\rangle=\int_0^{\cdot}\langle f,g\rangle\,\mathrm{d}\beta\quad\text{in } L^0(\Omega;C([0,T])).$$

In this situation, we have $\zeta=\int_0^{\cdot} f\,\mathrm{d}\beta$ *almost surely.*

Proof. (1) $\Rightarrow$ (2): By Remark 3.39 and Lemma 3.21, there exists a sequence $(f_n)_{n=1}^{\infty}$ of adapted step processes converging to f in $L^0_{\mathbb{F}}(\Omega;L^r_{\gamma,T})$. Moreover, we have $f_n\in L^p_{\mathbb{F}}(\Omega;L^r_{\gamma,T})$ and $\langle f_n,g\rangle\in L^p_{\mathbb{F}}(\Omega;L^2([0,T]))$ for all $g\in L^{r'}(U)$ and some $1<p<\infty$. Therefore, by Theorem 3.29 and Corollary 3.25, we obtain for each $t\in[0,T]$

$$\int_0^t\langle f_n,g\rangle\,\mathrm{d}\beta=\Big\langle\int_0^t f_n\,\mathrm{d}\beta,g\Big\rangle\quad\text{almost surely.}$$

Next we define $\zeta:=I^{\beta}_{\mathbb{F},\mathrm{loc}}(f)$. Then, by definition, we have $\lim_{n\to\infty}\int_0^{\cdot} f_n\,\mathrm{d}\beta=\zeta$ in $L^0(\Omega;L^r(U;C([0,T])))$, and thus

$$\lim_{n\to\infty}\int_0^{\cdot}\langle f_n,g\rangle\,\mathrm{d}\beta=\langle\zeta,g\rangle\quad\text{in } L^0(\Omega;C([0,T])).$$

Since U is a σ-finite measure space we can find a sequence $(U_k)_{k=1}^{\infty}$ of disjoint sets of finite measure such that $U=\bigcup_{k=1}^{\infty}U_k$. Similarly as in the proof of Theorem 3.13, we take $U^{(K)}:=\bigcup_{k=1}^{K}U_k$, $K\in\mathbb{N}$, and we can find an appropriate subsequence $(f_{n_j})_{j=1}^{\infty}$ such that

$$\lim_{j\to\infty}\langle f_{n_j},\mathbb{1}_{U^{(K)}}g\rangle=\langle f,\mathbb{1}_{U^{(K)}}g\rangle\quad\text{in } L^0_{\mathbb{F}}(\Omega;L^2([0,T])).$$

This leads to

$$\int_0^{\cdot}\langle f,\mathbb{1}_{U^{(K)}}g\rangle\,\mathrm{d}\beta=\lim_{j\to\infty}\int_0^{\cdot}\langle f_{n_j},\mathbb{1}_{U^{(K)}}g\rangle\,\mathrm{d}\beta=\langle\zeta,\mathbb{1}_{U^{(K)}}g\rangle$$

in $L^0(\Omega;C([0,T]))$.

By the dominated convergence theorem we both have $\lim_{K\to\infty}\langle f, \mathbb{1}_{U(K)}g\rangle = \langle f, g\rangle$ in $L^0_{\mathbb{F}}(\Omega; L^2([0,T]))$ and $\lim_{K\to\infty}\langle \zeta, \mathbb{1}_{U(K)}g\rangle = \langle \zeta, g\rangle$ in $L^0(\Omega; C([0,T]))$, which finally yield

$$\int_0^{\cdot} \langle f, g\rangle \,\mathrm{d}\beta = \lim_{K\to\infty} \int_0^{\cdot} \langle f, \mathbb{1}_{U(K)}g\rangle \,\mathrm{d}\beta = \lim_{K\to\infty} \langle \zeta, \mathbb{1}_{U(K)}g\rangle = \langle \zeta, g\rangle$$

in $L^0(\Omega; C([0,T]))$.

(2) $\Rightarrow$ (1): We define for each $n \in \mathbb{N}$

$$\tau_n(\omega) := T \wedge \inf\big\{t \in [0,T] : \big\| \sup_{s\in[0,t]} |\zeta(\omega, s)|\big\|_r \geq n\big\},$$

which is a stopping time since the map $t \mapsto \big\|\sup_{s\in[0,t]} |\zeta(\omega, s)|\big\|_r$ is continuous for each $\omega \in \Omega$ by the dominated convergence theorem. Then, $\zeta_n := \zeta(\cdot \wedge \tau_n) \in L^p(\Omega; L^r(U; C([0,T])))$ for some $1 < p < \infty$ and $\lim_{n\to\infty} \zeta_n = \zeta$ in $L^0(\Omega; L^r(U; C([0,T])))$. Now fix a $g \in L^{r'}(U)$. By (2) and Proposition 3.43 we have

$$\langle \zeta_n, g\rangle = \int_0^{\cdot} \mathbb{1}_{[0,\tau_n)} \langle f, g\rangle \,\mathrm{d}\beta \quad \text{almost surely.}$$

Since $\zeta_n \in L^p(\Omega; L^r(U; C([0,T])))$, Theorem 3.44 implies that $\mathbb{1}_{[0,\tau_n)} \langle f, g\rangle \in L^p(\Omega; L^2([0,T]))$, in particular

$$\langle \zeta_n(T), g\rangle = \int_0^T \mathbb{1}_{[0,\tau_n)} \langle f, g\rangle \,\mathrm{d}\beta \quad \text{in } L^p(\Omega).$$

Since $g \in L^{r'}(U)$ was arbitrary, Theorem 3.29 yield $f_n := \mathbb{1}_{[0,\tau_n)} f \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ with $\zeta_n(T) = \int_0^T f_n \,\mathrm{d}\beta$. By the same argument applied to $\zeta_n(t)$, combined with Corollary 3.25, we get

$$\zeta_n(t) = \int_0^t f_n \,\mathrm{d}\beta \quad \text{for every } t \in [0,T].$$

Thus, $\big(\int_0^{\cdot} f_n \,\mathrm{d}\beta\big)_{n=1}^\infty$ converges in $L^0(\Omega; L^r(U; C([0,T])))$, as we mentioned above. By Theorem 3.40, the sequence $(f_n)_{n=1}^\infty$ is Cauchy in $L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ with limit $\widetilde{f}$, and hence, a subsequence converges almost surely to $\widetilde{f}$ in $L^r_{\gamma,T}$. By the definition of the stopping times $(\tau_n)_{n=1}^\infty$ we also have $\lim_{n\to\infty} f_n = f$ almost surely in $L^r_{\gamma,T}$. But this means that $f = \widetilde{f} \in L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ and that $\zeta = \int_0^{\cdot} f \,\mathrm{d}\beta$ almost surely. ■

Chapter 4

An Application to Stochastic Evolution Equations

In this chapter we apply the theory developed in Chapter 3 to construct new processes from old ones. On the one hand, we will then be able to establish a connection to the operator-valued integration theory, and on the other hand, they are used in the notation of the Itô formula, which will be proved in Section 4.2. In the final section we want to apply Itô's formula to prove the existence of solutions for an abstract stochastic evolution equation.

In this chapter we will again assume that $1 < r < \infty$.

4.1 Itô Processes

Let $(\beta_n)_{n=1}^{\infty}$ be a sequence of independent Brownian motions $\beta_n = \big(\beta_n(t)\big)_{t\in[0,T]}$ on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$, and let $\mathbb{F} = (\mathscr{F}_t)_{t\in[0,T]}$ be a filtration such that for all $0 \leq s < t \leq T$ and all $n \in \mathbb{N}$,

(1) $\beta_n(t)$ is $\mathscr{F}_t$-measurable and

(2) $\beta_n(t) - \beta_n(s)$ is independent of $\mathscr{F}_s$.

Moreover, we assume that $x_0 \colon \Omega \to L^r(U)$ is an $\mathscr{F}_0$-measurable random variable and that the functions $f \colon [0,T]\times\Omega \to L^r(U)$ and $b_n \colon [0,T]\times\Omega \to L^r(U)$ are adapted with respect to $\mathbb{F}$ for each $n \in \mathbb{N}$ satisfying

$$\int_0^T \|f\|_r \,\mathrm{d}t < \infty \quad \text{and} \quad \Big\| \Big(\int_0^T \sum_{n=1}^{\infty} |b_n|^2 \,\mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r < \infty \quad \text{almost surely.}$$

With these notions we have the following definition.

Definition 4.1. *Under the assumptions taken above, we call an $L^r(U)$-valued process $X = \big(X(t)\big)_{t\in[0,T]}$ an* Itô process *(or* standard process*) if it has the integral representation*

$$X(t) = x_0 + \int_0^t f \,\mathrm{d}s + \sum_{n=1}^{\infty} \int_0^t b_n \,\mathrm{d}\beta_n \quad \textit{for } 0 \le t \le T. \tag{4.1}$$

Remark 4.2. (1) Instead of (4.1), we also formally write

$$\mathrm{d}X = f\,\mathrm{d}t + \sum_{n=1}^{\infty} b_n \,\mathrm{d}\beta_n.$$

(2) The function f is almost surely Bochner-integrable, and therefore the paths of f belong to $L^1([0,T], L^r(U))$. Moreover

$$\Big\| \int_0^T |f| \,\mathrm{d}t \Big\|_r \le \int_0^T \|f\|_r \,\mathrm{d}t < \infty,$$

which means that f also belongs to $L^0_{\mathbb{F}}(\Omega; L^r(U; L^1([0,T])))$.

(3) By the assumptions we made on the filtration, each stochastic integral in (4.1) is well-defined. ■

Next, we give an answer to the question why the series in (4.1) is well-defined. To prove this, we first need some preliminary results where we always assume that $x_0 = 0$ almost surely.

Lemma 4.3. *In addition to the above properties, let $b_n \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ for each $n \in \mathbb{N}$. Then we have for any fixed $N \in \mathbb{N}$*

$$\mathbb{E}\Big\| \sum_{n=1}^{N} \int_0^T b_n \,\mathrm{d}\beta_n \Big\|_r^p \eqsim_{p,r} \mathbb{E}\Big\| \Big(\int_0^T \sum_{n=1}^{N} |b_n|^2 \,\mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r^p.$$

Proof. First consider adapted step processes

$$b_n = \sum_{m=1}^{M} v_m^{(n)} \mathbb{1}_{[t_{m-1}, t_m)},$$

where $0 = t_0 < \ldots < t_M = T$ and $v_m^{(n)} : \Omega \to L^r(U)$ is a simple function, which is $\mathscr{F}_{t_{m-1}}$-measurable for each $m = 1, \ldots, M$ and $n = 1, \ldots, N$.

Now we define for $m = 1, \dots, M$ and $n = 1, \dots, N$

$$\xi_{(m-1)N+n} := \beta_n(t_m) - \beta_n(t_{m-1}), \qquad \mathscr{F}_{(m-1)N+n} := \sigma(\xi_\ell \colon \ell \leq (m-1)N+n),$$
$$\gamma_{(m-1)N+n} := (t_m - t_{m-1})^{-\frac{1}{2}}\, \xi_{(m-1)N+n}, \qquad v_{(m-1)N+n} := v_m^{(n)}.$$

Then, for all $k = 1, \dots, M \cdot N$, ξ_k is a centered, $\mathscr{F}_k$-measurable random variable, which is independent of $\mathscr{F}_{k-1}$. Additionally, v_k is $\mathscr{F}_{k-1}$-measurable, which means that the sequence $(v_k)_{k=1}^{MN}$ is predictable with respect to the filtration $(\mathscr{F}_k)_{k=1}^{MN}$. And furthermore, by the independence of the Brownian motions, $(\gamma_k)_{k=1}^{MN}$ is a sequence of independent standard Gaussian variables.

By applying the Decoupling theorem and the Kahane inequality as in Corollary 2.24, we obtain

$$\begin{aligned}
\mathbb{E}\Big\| \sum_{n=1}^{N} \int_0^T b_n \,\mathrm{d}\beta_n \Big\|_r^p &= \mathbb{E}\Big\| \sum_{n=1}^{N} \sum_{m=1}^{M} v_m^{(n)} \big(\beta_n(t_m) - \beta_n(t_{m-1})\big) \Big\|_r^p \\
&= \mathbb{E}\Big\| \sum_{k=1}^{MN} v_k \xi_k \Big\|_r^p \\
&\eqsim_{p,r} \mathbb{E}\Big\| \Big(\sum_{k=1}^{MN} |v_k|^2 (t_{\lceil \frac{k}{N} \rceil} - t_{\lceil \frac{k}{N} \rceil - 1}) \Big)^{\frac{1}{2}} \Big\|_r^p \\
&= \mathbb{E}\Big\| \Big(\sum_{n=1}^{N} \sum_{m=1}^{M} |v_m^{(n)}|^2 (t_m - t_{m-1}) \Big)^{\frac{1}{2}} \Big\|_r^p \\
&= \mathbb{E}\Big\| \Big(\int_0^T \sum_{n=1}^{N} |b_n|^2 \,\mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r^p.
\end{aligned}$$

The general estimate finally follows by approximating each adapted L^p process b_n with adapted step processes. ■

Proposition 4.4. *Let $(b_n)_{n=1}^\infty \subseteq L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ such that*

$$\mathbb{E}\Big\| \Big(\int_0^T \sum_{n=1}^{\infty} |b_n|^2 \,\mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r^p < \infty$$

for some $1 < p < \infty$. Then for each $n \in \mathbb{N}$, the integral $X_n(t) := \int_0^t b_n \,\mathrm{d}\beta_n$ exists in $L^p(\Omega, \mathscr{F}_t; L^r(U))$, and the sum $X(t) := \sum_{n=1}^\infty X_n(t)$ converges in $L^p(\Omega, \mathscr{F}_t; L^r(U))$ for each fixed $t \in [0, T]$. Moreover, we have

$$\mathbb{E}\|X(t)\|_r^p \eqsim_{p,r} \mathbb{E}\Big\| \Big(\int_0^t \sum_{n=1}^{\infty} |b_n|^2 \,\mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r^p.$$

Proof. By the assumption we infer that $b_n \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$, and hence $X_n(t)$ is an element of $L^p(\Omega, \mathscr{F}_t; L^r(U))$ for all $n \in \mathbb{N}$ and each $t \in [0,T]$. Furthermore, by Lemma 4.3, we obtain

$$\lim_{M,N\to\infty} \mathbb{E}\Big\| \sum_{n=M}^{N} X_n(t) \Big\|_r^p \eqsim_{p,r} \lim_{M,N\to\infty} \mathbb{E}\Big\| \Big(\int_0^t \sum_{n=M}^{N} |b_n|^2 \,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_r^p = 0.$$

So, $\big(\sum_{n=1}^N X_n(t)\big)_{N=1}^\infty$ is Cauchy in $L^p(\Omega, \mathscr{F}_t; L^r(U))$, which gives the desired convergence result. Finally, by Lemma 4.3 and the dominated convergence theorem, we obtain

$$\begin{aligned}\mathbb{E}\|X(t)\|_r^p = \lim_{N\to\infty} \mathbb{E}\Big\| \sum_{n=1}^{N} X_n(t) \Big\|_r^p &\eqsim_{p,r} \lim_{N\to\infty} \mathbb{E}\Big\| \Big(\int_0^t \sum_{n=1}^{N} |b_n|^2 \,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_r^p \\ &= \mathbb{E}\Big\| \Big(\int_0^t \sum_{n=1}^{\infty} |b_n|^2 \,\mathrm{d}t\Big)^{\frac{1}{2}} \Big\|_r^p.\end{aligned}$$

■

Proposition 4.5. *Under the assumptions of Proposition 4.4, the process $\big(X(t)\big)_{t\in[0,T]}$ is an L^p martingale with respect to $\mathbb{F}$ and has a continuous version satisfying the maximal inequality*

$$\mathbb{E}\Big\| \sup_{t\in[0,T]} |X(t)| \Big\|_r^p \lesssim_{p,r} \mathbb{E}\|X(T)\|_r^p.$$

Proof. By Proposition 4.4, $X(t)$ is $\mathscr{F}_t$-measurable, and by Proposition 3.30, each process $\big(X_n(t)\big)_{t\in[0,T]}$ is an L^p martingale with respect to $\mathbb{F}$. Hence, by the continuity of the conditional expectation, we get

$$\begin{aligned}\mathbb{E}[X(t)|\mathscr{F}_s] = \lim_{N\to\infty} \mathbb{E}\Big[\sum_{n=1}^{N} X_n(t)|\mathscr{F}_s\Big] &= \lim_{N\to\infty} \sum_{n=1}^{N} \mathbb{E}[X_n(t)|\mathscr{F}_s] \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} X_n(s) = X(s)\end{aligned}$$

in $L^p(\Omega, \mathscr{F}_s; L^r(U))$ for each fixed $0 \le s \le t \le T$. Thus, $\big(X(t)\big)_{t\in[0,T]}$ is an L^p martingale. By Theorem 3.31, the process $Y_N := \sum_{n=1}^N X_n$ has a continuous version and, by proceeding in the exact same way as in the proof of that theorem, we obtain a continuous version of $\big(X(t)\big)_{t\in[0,T]}$ satisfying the desired maximal inequality. ■

We next consider the localized case.

Lemma 4.6. *Under the assumptions of Proposition 4.4, we have for all $\delta > 0$ and all $\varepsilon > 0$*

$$\mathbb{P}\Big(\big\|\sup_{t\in[0,T]}|X(t)|\big\|_r > \varepsilon\Big) \le \frac{c_{p,r}\delta^p}{\varepsilon^p} + \mathbb{P}\Big(\Big\|\Big(\sum_{n=1}^{\infty}|b_n|^2\Big)^{\frac{1}{2}}\Big\|_{L^r_{\gamma,T}} \ge \delta\Big)$$

and

$$\mathbb{P}\Big(\Big\|\Big(\sum_{n=1}^{\infty}|b_n|^2\Big)^{\frac{1}{2}}\Big\|_{L^r_{\gamma,T}} > \varepsilon\Big) \le \frac{c_{p,r}\delta^p}{\varepsilon^p} + \mathbb{P}\Big(\big\|\sup_{t\in[0,T]}|X(t)|\big\|_r \ge \delta\Big),$$

where $c_{p,r}$ is the constant from Proposition 4.4.

Proof. Since for every stopping time τ we have

$$\mathbb{E}\Big\|\Big(\int_0^T \mathbb{1}_{[0,\tau)}\sum_{n=1}^{\infty}|b_n|^2\,\mathrm{d}t\Big)^{\frac{1}{2}}\Big\|_r^p \le \mathbb{E}\Big\|\Big(\int_0^T \sum_{n=1}^{\infty}|b_n|^2\,\mathrm{d}t\Big)^{\frac{1}{2}}\Big\|_r^p < \infty,$$

we obtain by Proposition 4.4

$$\begin{aligned}\mathbb{E}\|X(\tau)\|_r^p &= \mathbb{E}\Big\|\sum_{n=1}^{\infty}\int_0^{\tau} b_n\,\mathrm{d}\beta_n\Big\|_r^p = \mathbb{E}\Big\|\sum_{n=1}^{\infty}\int_0^{T}\mathbb{1}_{[0,\tau)}b_n\,\mathrm{d}\beta_n\Big\|_r^p \\ &\eqsim_{p,r} \mathbb{E}\Big\|\Big(\int_0^T \mathbb{1}_{[0,\tau)}\sum_{n=1}^{\infty}|b_n|^2\,\mathrm{d}t\Big)^{\frac{1}{2}}\Big\|_r^p \\ &= \mathbb{E}\big\|\mathbb{1}_{[0,\tau)}\|b\|_{l^2}\big\|_{L^r_{\gamma,T}}^p.\end{aligned}$$

Now, the proof can be finished in the same way as in Lemma 3.36. ■

Theorem 4.7. *Let X be an Itô process given by $\mathrm{d}X = \sum_{n=1}^{\infty} b_n\,\mathrm{d}\beta_n$. Then, the process X is well-defined as an element of $L^0(\Omega; L^r(U; C([0,T])))$. Moreover, the assertions from Lemma 4.6 extend to this process.*

Proof. For any fixed $N \in \mathbb{N}$ and each $n = 1,\dots,N$, let $(\tau_k^{(n)})_{k=1}^{\infty}$ be the localizing sequence for $b_n \in L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ from Remark 3.39 (2). Then $b_k^{(n)} := \mathbb{1}_{[0,\tau_k^{(n)})}b_n \in L^p_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ for some $1 < p < \infty$, and $\lim_{k\to\infty} b_k^{(n)} = b_n$ in $L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ for all $n = 1,\dots,N$. By definition we have

$$\lim_{k\to\infty}\int_0^{\cdot} b_k^{(n)}\,\mathrm{d}\beta_n = \int_0^{\cdot} b_n\,\mathrm{d}\beta_n \quad \text{in } L^0(\Omega; L^r(U; C([0,T]))),$$

and therefore, by the definition of this metric space,

$$\lim_{k\to\infty}\sum_{n=1}^{N}\int_0^{\cdot} b_k^{(n)}\,\mathrm{d}\beta_n=\sum_{n=1}^{N}\int_0^{\cdot} b_n\,\mathrm{d}\beta_n \quad \text{in } L^0(\Omega;L^r(U;C([0,T]))).$$

Let $\varepsilon>0$ and $\delta>0$. By passing to a subsequence, we obtain similarly as in the proof of Theorem 3.40

$$\lim_{j\to\infty}\mathbb{P}\Big(\Big\|\sup_{t\in[0,T]}\Big|\sum_{n=1}^{N}\int_0^t b_{k_j}^{(n)}\,\mathrm{d}\beta_n\Big|\Big\|_r>\varepsilon\Big)=\mathbb{P}\Big(\Big\|\sup_{t\in[0,T]}\Big|\sum_{n=1}^{N}\int_0^t b_n\,\mathrm{d}\beta_n\Big|\Big\|_r>\varepsilon\Big)$$

and

$$\lim_{j\to\infty}\mathbb{P}\Big(\Big\|\Big(\sum_{n=1}^{N}|b_{k_j}^{(n)}|^2\Big)^{\frac{1}{2}}\Big\|_{L^r_{\gamma,T}}\geq\delta\Big)=\mathbb{P}\Big(\Big\|\Big(\sum_{n=1}^{N}|b_n|^2\Big)^{\frac{1}{2}}\Big\|_{L^r_{\gamma,T}}\geq\delta\Big).$$

Using this together with Lemma 4.6 for each sum $\sum_{n=1}^{N} b_{k_j}^{(n)}$, we get

$$\mathbb{P}\Big(\Big\|\sup_{t\in[0,T]}\Big|\sum_{n=1}^{N}\int_0^t b_n\,\mathrm{d}\beta_n\Big|\Big\|_r>\varepsilon\Big)\leq\frac{c_{p,r}\delta^p}{\varepsilon^p}+\mathbb{P}\Big(\Big\|\Big(\sum_{n=1}^{N}|b_n|^2\Big)^{\frac{1}{2}}\Big\|_{L^r_{\gamma,T}}\geq\delta\Big)$$

and similarly

$$\mathbb{P}\Big(\Big\|\Big(\sum_{n=1}^{N}|b_n|^2\Big)^{\frac{1}{2}}\Big\|_{L^r_{\gamma,T}}>\varepsilon\Big)\leq\frac{c_{p,r}\delta^p}{\varepsilon^p}+\mathbb{P}\Big(\Big\|\sup_{t\in[0,T]}\Big|\sum_{n=1}^{N}\int_0^t b_n\,\mathrm{d}\beta_n\Big|\Big\|_r\geq\delta\Big).$$

Since

$$\Big\|\Big(\int_0^T\sum_{n=1}^{\infty}|b_n|^2\,\mathrm{d}t\Big)^{\frac{1}{2}}\Big\|_r<\infty \quad \text{almost surely,}$$

we have

$$\lim_{N,M\to\infty}\mathbb{P}\Big(\Big\|\Big(\sum_{n=M}^{N}|b_n|^2\Big)^{\frac{1}{2}}\Big\|_{L^r_{\gamma,T}}\geq\delta\Big)=0.$$

Thus, by the previous estimate, $\big(\sum_{n=1}^{N}\int_0^{\cdot} b_n\,\mathrm{d}\beta_n\big)_{N=1}^{\infty}$ is 'Cauchy in probability', and so also in $L^0(\Omega;L^r(U;C([0,T])))$. Hence, by Proposition 3.37, the series $\sum_{n=1}^{\infty}\int_0^{\cdot} b_n\,\mathrm{d}\beta_n$ is convergent in this space, i.e., the Itô process $\big(X(t)\big)_{t\in[0,T]}$ is well-defined. Finally, a limiting argument also shows that the assertions of Lemma 4.6 are still true for this kind of process. ■

Finally, we extend Theorem 3.45 to Itô processes.

Theorem 4.8. *For each $n \in \mathbb{N}$, let $b_n \colon [0,T] \times \Omega \to L^r(U)$ be measurable and adapted to a filtration $\mathbb{F}$, and let $\big(\langle b_n, g\rangle\big)_{n=1}^{\infty} \in L^0(\Omega; L^2([0,T]; l^2))$ for all $g \in L^{r'}(U)$. Then the following assertions are equivalent:*

(1) $\big\|\big(\int_0^T \sum_{n=1}^{\infty} |b_n|^2 \,\mathrm{d}t\big)^{\frac{1}{2}}\big\|_r < \infty$ *almost surely;*

(2) there exists a process $X \in L^0(\Omega; L^r(U; C([0,T])))$ *such that for all* $g \in L^{r'}(U)$ *we have*

$$\langle X, g\rangle = \sum_{n=1}^{\infty} \int_0^{\cdot} \langle b_n, g\rangle \,\mathrm{d}\beta_n \quad \textit{in } L^0(\Omega; C([0,T])).$$

In this situation, X equals almost surely the Itô process formally given by $\mathrm{d}X = \sum_{n=1}^{\infty} b_n \,\mathrm{d}\beta_n$.

Proof. (1) $\Rightarrow$ (2) : The assumption implies that $b_n \in L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$ for each $n \in \mathbb{N}$. Thus, from Theorem 3.45 we deduce that for any $n \in \mathbb{N}$

$$\Big\langle \int_0^{\cdot} b_n \,\mathrm{d}\beta_n, g \Big\rangle = \int_0^{\cdot} \langle b_n, g\rangle \,\mathrm{d}\beta_n \quad \text{in } L^0(\Omega; C([0,T])).$$

Finally, by taking $X := \sum_{n=1}^{\infty} \int_0^{\cdot} b_n \,\mathrm{d}\beta_n \in L^0(\Omega; L^r(U; C([0,T])))$, we have

$$\langle X, g\rangle = \sum_{n=1}^{\infty} \Big\langle \int_0^{\cdot} b_n \,\mathrm{d}\beta_n, g \Big\rangle = \sum_{n=1}^{\infty} \int_0^{\cdot} \langle b_n, g\rangle \,\mathrm{d}\beta_n \quad \text{in } L^0(\Omega; C([0,T])),$$

where we used that $\sum_{n=1}^{\infty} \int_0^{\cdot} b_n \,\mathrm{d}\beta_n$ also converges in $L^0(\Omega; C([0,T]; L^r(U)))$.

(2) $\Rightarrow$ (1) : For $m \in \mathbb{N}$ let $P_m \colon L^r(U) \to L^r(U)$ be the operator from Lemma A.26 (see also the proof of Theorem 3.29). Then, for any $k = 1, \ldots, N_m$, we have $\mathbb{1}_{U_k^{(m)}} \in L^{r'}(U)$ and

$$\big\langle P_m b_n(t,\omega), \mathbb{1}_{U_k^{(m)}} \big\rangle = \int_{U_k^{(m)}} b_n(t,\omega) \,\mathrm{d}\mu = \big\langle b_n(t,\omega), \mathbb{1}_{U_k^{(m)}} \big\rangle$$

for all $(t,\omega) \in [0,T] \times \Omega$ and each $n \in \mathbb{N}$. By the assumptions, almost surely this yields

$$\begin{aligned} &\Big\| \Big(\int_0^T \sum_{n=1}^{\infty} \Big| \int_{U_k^{(m)}} b_n \,\mathrm{d}\mu \, \mathbb{1}_{U_k^{(m)}} \Big|^2 \mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r \\ &\qquad = \mu(U_k^{(m)})^{\frac{1}{r}} \Big(\int_0^T \sum_{n=1}^{\infty} \big| \big\langle b_n, \mathbb{1}_{U_k^{(m)}} \big\rangle \big|^2 \mathrm{d}t \Big)^{\frac{1}{2}} < \infty. \end{aligned}$$

Therefore, by the definition of P_m, we obtain $\big\|\big(\int_0^T \sum_{n=1}^\infty |P_m b_n|^2\,\mathrm{d}t\big)^{\frac{1}{2}}\big\|_r < \infty$, and by Theorem 4.7 we deduce that the series $\sum_{n=1}^\infty \int_0^\cdot P_m b_n\,\mathrm{d}\beta_n$ converges in $L^0(\Omega; L^r(U; C([0,T])))$. Thus, we obtain

$$\Big\langle \sum_{n=1}^\infty \int_0^\cdot P_m b_n\,\mathrm{d}\beta_n, g\Big\rangle = \sum_{n=1}^\infty \Big\langle \int_0^\cdot P_m b_n\,\mathrm{d}\beta_n, g\Big\rangle \quad \text{almost surely.}$$

Now let $g \in L^{r'}(U)$ be arbitrary, and observe that also $P_m g \in L^{r'}(U)$. Then, by (2) and a small computation similar as in the proof of Theorem 3.29, we get

$$\begin{aligned}\langle P_m X, g\rangle = \langle X, P_m g\rangle &= \sum_{n=1}^\infty \int_0^\cdot \langle b_n, P_m g\rangle\,\mathrm{d}\beta_n = \sum_{n=1}^\infty \int_0^\cdot \langle P_m b_n, g\rangle\,\mathrm{d}\beta_n \\ &= \sum_{n=1}^\infty \Big\langle \int_0^\cdot P_m b_n\,\mathrm{d}\beta_n, g\Big\rangle = \Big\langle \sum_{n=1}^\infty \int_0^\cdot P_m b_n\,\mathrm{d}\beta_n, g\Big\rangle\end{aligned}$$

almost surely. By Corollary A.8, this gives $P_m X = \sum_{n=1}^\infty \int_0^\cdot P_m b_n\,\mathrm{d}\beta_n$ almost surely. Moreover, by the properties of the operators P_m and the dominated convergence theorem, we have $\lim_{m\to\infty} P_m X = X$ in $L^0(\Omega; L^r(U; C([0,T])))$. By taking $c_m := (P_m b_n)_{n=1}^\infty \in L^0(\Omega; L^r(U; L^2([0,T]; l^2)))$, Theorem 4.7 yield that $(c_m)_{m=1}^\infty$ is Cauchy in $L^0(\Omega; L^r(U; L^2([0,T]; l^2)))$ and therefore has a limit $\widetilde{c}$. Similar as in Theorem 3.29 and Theorem 3.45, we can show that

$$\begin{aligned}&\lim_{m\to\infty} c_{m,n}(t,\omega,u) = \lim_{m\to\infty} (P_m b_n(t,\omega))(u) = b_n(t,\omega,u) \quad \text{and} \\ &\lim_{m\to\infty} c_{m,n}(t,\omega,u) = \widetilde{c}_n(t,\omega,u)\end{aligned}$$

almost everywhere and for any $n \in \mathbb{N}$. Finally, this leads to $(b_n)_{n=1}^\infty = \widetilde{c} \in L^0(\Omega; L^r(U; L^2([0,T]; l^2)))$ and this yield $X = \sum_{n=1}^\infty \int_0^\cdot b_n\,\mathrm{d}\beta_n$ almost surely. ■

Remark 4.9. Let D' be a dense linear subspace of $L^{r'}(U)$. Then the previous theorem remains valid if we replace $L^{r'}(U)$ with D'. To prove this, fix a $g \in L^{r'}(U)$ and choose $(g_m)_{m=1}^\infty \subseteq D'$ such that $g = \lim_{m\to\infty} g_m$. Then of course we have $\langle X, g\rangle = \lim_{m\to\infty}\langle X, g_m\rangle$ in $L^0(\Omega; C([0,T]))$. Since (2) holds for every g_m, this convergence and Theorem 4.7 imply that $(a_m)_{m=1}^\infty \subseteq L^0(\Omega; L^2([0,T]; l^2))$, given by $a_m := \big(\langle b_n, g_m\rangle\big)_{n=1}^\infty$, is Cauchy in $L^0(\Omega; L^2([0,T]; l^2))$. Similar as in the previous proof, we can show that the obtaining limit equals $\big(\langle b_n, g\rangle\big)_{n=1}^\infty$ almost surely, and thus $\big(\langle b_n, g\rangle\big)_{n=1}^\infty \in L^0(\Omega; L^2([0,T]; l^2))$. Finally, by this convergence and Theorem 4.7, we have

$$\sum_{n=1}^\infty \int_0^\cdot \langle b_m, g\rangle\,\mathrm{d}\beta_n = \lim_{m\to\infty} \sum_{n=1}^\infty \int_0^\cdot \langle b_n, g_m\rangle\,\mathrm{d}\beta_n = \lim_{m\to\infty}\langle X, g_m\rangle = \langle X, g\rangle$$

in $L^0(\Omega; C([0,T])$, and Theorem 4.8 concludes the proof. ■

We next make a short excursion in the theory of stochastic integration of operator-valued processes with respect to an *H-cylindrical Brownian motion*. Here, we assume that $\big(H, (\cdot|\cdot)_H\big)$ is a real separable Hilbert space.

Definition 4.10. *We call a family $W_H = (W_H(t))_{t\in[0,T]}$ of bounded linear operators from H to $L^2(\Omega)$ an H-*cylindrical Brownian motion *if:*

(1) $W_H h = (W_H(t)h)_{t\in[0,T]}$ is a real-valued Brownian motion for each $h \in H$ with $\|h\|_H = 1$;

(2) $\mathbb{E}\big(W_H(s)g \cdot W_H(t)h\big) = (s \wedge t)(g|h)_H$ for all $s, t \in [0, T]$ and $g, h \in H$.

The procedure is similar to that in this thesis. We first consider *finite rank adapted step processes* $\Phi\colon [0,T] \times \Omega \to \mathscr{B}(H, L^r(U))$ with respect to a given filtration $\mathbb{F} = (\mathscr{F}_t)_{t\in[0,T]}$, i.e.,

$$\Phi(t,\omega) = \sum_{m=1}^{M}\sum_{n=1}^{N} \mathbb{1}_{[t_{n-1},t_n)}(t)\mathbb{1}_{A_{mn}}(\omega)\sum_{j=1}^{k} h_j \otimes x_{jmn},$$

where $0 = t_0 < \ldots < t_N = T$, 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 $L^r(U)$. For $h \in H$ and $x \in L^r(U)$ we denote by $h \otimes x \in \mathscr{B}(H, L^r(U))$ the operator defined by

$$(h \otimes x)h' := (h|h')_H\, x, \quad h' \in H.$$

Moreover, we assume that W_H is an H-cylindrical Brownian motion on $(\Omega, \mathscr{A}, \mathbb{P})$, with the properties that for all $h \in H$

(1) $W_H(t)h$ is $\mathscr{F}_t$-measurable and

(2) $W_H(t)h - W_H(s)h$ is independent of $\mathscr{F}_s$ for $t > s$.

A filtration satisfying these conditions is given by $\mathbb{F}^{W_H} = (\mathscr{F}_t^{W_H})_{t\in[0,T]}$, where $\mathscr{F}_t^{W_H} := \sigma\big(W_H(s)e_n\colon s \le t, n \in \mathbb{N}\big)$ and $(e_n)_{n=1}^{\infty}$ is an orthonormal basis of H.

The *stochastic integral* with respect to W_H of a finite rank adapted step process Φ of the above form is then defined as

$$\int_0^T \Phi \,\mathrm{d}W_H := \sum_{m=1}^{M}\sum_{n=1}^{N} \mathbb{1}_{A_{mn}} \sum_{j=1}^{k}\big(W_H(t_n)h_j - W_H(t_{n-1})h_j\big)x_{jmn}.$$

By approximation, the stochastic integral can now be extended to operator-valued adapted processes. We will not go into any further detail here and refer to [10] and [15] for more information on this theory.

Interestingly, there is a connection between operator-valued stochastic integrals and Itô processes.

Theorem 4.11 (Operator-valued integrals). *Let W_H be an H-cylindrical Brownian motion, and let $B\colon [0,T]\times\Omega \to \mathscr{B}(H, L^r(U))$ be an operator-valued process, which is adapted to the filtration $\mathbb{F}$ and satisfies $B^*g \in L^0(\Omega; L^2([0,T]; H))$ for each $g \in L^{r'}(U)$. Moreover, we assume that*

$$\|B\|_{\gamma([0,T];H,L^r(U))} \eqsim_{p,r} \Big\| \Big(\int_0^T \sum_{n=1}^{\infty} |Be_n|^2 \,\mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r < \infty \quad \textit{almost surely,}$$

where $(e_n)_{n=1}^{\infty}$ is an orthonormal basis of H (cf. [15] for the norm-equivalence and the definition of the norm on the left-hand side). Then the process

$$X(t) := \int_0^t B \,\mathrm{d}W_H$$

equals the $L^r(U)$-valued Itô process given by

$$\mathrm{d}X = \sum_{n=1}^{\infty} b_n \,\mathrm{d}\beta_n,$$

where the sequences $b = (b_n)_{n=1}^{\infty}$ and $\beta = (\beta_n)_{n=1}^{\infty}$ are defined by $b_n := Be_n$ and $\beta_n := W_H e_n$.

Proof. By [15, Corollary 4.5.10], the operator-valued stochastic integral has the following series representation

$$X(t) = \int_0^t B \,\mathrm{d}W_H = \sum_{n=1}^{\infty} \int_0^t Be_n \,\mathrm{d}(W_H e_n)$$

with convergence in $L^0(\Omega; L^r(U))$ for each $t \in [0,T]$. Thus, we obtain

$$X(t) = \sum_{n=1}^{\infty} \int_0^t Be_n \,\mathrm{d}(W_H e_n) = \sum_{n=1}^{\infty} \int_0^t b_n \,\mathrm{d}\beta_n.$$

By the definition of W_H, $(\beta_n)_{n=1}^{\infty} = (W_H e_n)_{n=1}^{\infty}$ is a sequence of independent Brownian motions, and since B is adapted to $\mathbb{F}$, Lemma 3.5.1 in [15] implies that each b_n is adapted with respect to $\mathbb{F}$. Finally, since $\big\| \big(\int_0^T \sum_{n=1}^{\infty} |b_n|^2 \,\mathrm{d}t \big)^{\frac{1}{2}} \big\|_r < \infty$ almost surely, X is an $L^r(U)$-valued Itô process, and the series even converges in $L^0(\Omega; L^r(U; C([0,T])))$, by Theorem 4.7. ■

And a converse direction is also true.

Theorem 4.12. *Let X be an Itô Process given by* $\mathrm{d}X = \sum_{n=1}^{\infty} b_n \,\mathrm{d}\beta_n$, *and assume that* $\big(\langle b_n, g\rangle\big)_{n=1}^{\infty} \in L^0(\Omega; L^2([0,T]; l^2))$ *for all* $g \in L^{r'}(U)$ *and* $\big\|\big(\sum_{n=1}^{\infty} |b_n(\omega,t)|^2\big)^{\frac{1}{2}}\big\|_r < \infty$ *for almost all* $(\omega, t) \in \Omega \times [0,T]$. *Then X can be represented as an operator-valued integral*

$$X = \int_0^{\cdot} B \,\mathrm{d}W_H,$$

where $H = l^2$, and the H-cylindrical Brownian motion and the operator-valued function $B\colon [0,T] \times \Omega \to \mathscr{B}(H, L^r(U))$ *are given by*

$$W_H\big((h_n)_{n=1}^{\infty}\big) = \sum_{n=1}^{\infty} h_n \beta_n,$$
$$B(t,\omega)\big((h_n)_{n=1}^{\infty}\big) = \sum_{n=1}^{\infty} h_n b_n(t,\omega).$$

Proof. We only have to show that W_H is indeed an H-cylindrical Brownian motion, and that B is well-defined with $\|B\|_{\gamma([0,T];H,L^r(U))} < \infty$. For this purpose, let g, $h \in l^2$ and $0 \le s \le t \le T$ be arbitrary. Observe that $\beta_n(t)$ is independent of $\beta_m(s)$ for all $m \neq n$, and that $\beta_n(t) - \beta_n(s)$ is a (centered) Gaussian variable with variance $t - s$ for any $n \in \mathbb{N}$. Thus, we obtain

$$\begin{aligned}
\mathbb{E}\big(W_H(s)g \cdot W_H(t)h\big) &= \sum_{n=1}^{\infty} \mathbb{E} g_n h_n \beta_n(s)\beta_n(t) \;+ \sum_{n\neq m} \mathbb{E} g_m h_n \beta_m(s)\beta_n(t) \\
&= \sum_{n=1}^{\infty} g_n h_h \mathbb{E}\beta_n(s)\beta_n(t) \;+ \sum_{n\neq m} g_m h_n \mathbb{E}\beta_m(s)\mathbb{E}\beta_n(t) \\
&= s(g|h)_{l^2},
\end{aligned}$$

using that

$$\begin{aligned}
\mathbb{E}\beta_n(s)\beta_n(t) &= \tfrac{1}{2}\Big(\mathbb{E}\beta_n(t)^2 + \mathbb{E}\beta_n(s)^2 - \mathbb{E}\big(\beta_n(t) - \beta_n(s)\big)^2\Big) \\
&= \tfrac{1}{2}(t + s - t + s) = s.
\end{aligned}$$

This implies $\mathbb{E}|W_H(t)h|^2 = t\|h\|_{l^2} < \infty$. And since linearity is trivial, this means that $W_H(t)$ is well-defined as a bounded operator from H to $L^2(\Omega)$ with $\|W_H(t)\|_{\mathscr{B}(H,L^2(\Omega))} = \sqrt{t}$. Next, we show that $(W_H(t)h)_{t\in[0,T]}$ is a Brownian motion if we further assume that $\|h\|_{l^2} = 1$. Clearly, we have

$$W_H(0)h = \sum_{n=1}^{\infty} h_n \beta_n(0) = 0 \quad \text{almost surely.}$$

For $N \in \mathbb{N}$ and $0 \leq s \leq t \leq T$ we define

$$X_N^{t,s} := \sum_{n=1}^{N} h_n\big(\beta_n(t) - \beta_n(s)\big).$$

Since $(\beta_n)_{n=1}^\infty$ is a sequence of independent Brownian motions, $X_N^{t,s}$ is a Gaussian random variable with variance $q_N := (t-s)\sum_{n=1}^N h_n^2$. Moreover, we have $\lim_{N\to\infty} X_N^{t,s} = W_H(t)h - W_H(s)h$ in $L^2(\Omega)$, and by Proposition A.17 we infer that $W_H(t)h - W_H(s)h$ is Gaussian with variance

$$q = \lim_{N\to\infty} q_N = (t-s)\|h\|_{l^2}^2 = t - s.$$

Additionally, by the independence of the Brownian motions, $W_H(t)h - W_H(s)h$ is independent of $W_H(r)h$ whenever $0 \leq r \leq s$. So, W_H is an H-cylindrical Brownian motion.

Next, we inspect the operator B. Let $h = (h_n)_{n=1}^\infty \in l^2$. By assumption, we have for all $(t,\omega) \in [0,T] \times \Omega$

$$\|B(t,\omega)h\|_r = \Big\| \sum_{n=1}^\infty h_n b_n(t,\omega) \Big\|_r \leq \|h\|_{l^2} \Big\| \Big(\sum_{n=1}^\infty |b_n(\omega,t)|^2 \Big)^{\frac{1}{2}} \Big\|_r < \infty,$$

which implies that $B(t,\omega)$ is a bounded operator for all $(t,\omega) \in [0,T] \times \Omega$. Moreover, by Fubini's theorem, we obtain

$$\langle Bh, g\rangle = \int_U \sum_{n=1}^\infty h_n b_n g \,\mathrm{d}\mu = \sum_{n=1}^\infty h_n \int_U b_n g \,\mathrm{d}\mu = (h|B^* g)_{l^2}$$

for all $h \in l^2$ and $g \in L^{r'}(U)$. Thus, the adjoint operator $B^*\colon L^{r'}(U) \to l^2$ is given by $B^* g = \big(\langle b_n, g\rangle\big)_{n=1}^\infty$, and almost surely we have

$$\|B^* g\|_{L^2([0,T],l^2)}^2 = \int_0^T \sum_{n=1}^\infty \langle b_n, g\rangle^2 \,\mathrm{d}t < \infty.$$

Finally, the operator B satisfies

$$\|B\|_{\gamma([0,T];H,L^r(U))} \eqsim_{p,r} \Big\| \Big(\int_0^T \sum_{n=1}^\infty |b_n|^2 \,\mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_r < \infty$$

such that Theorem 4.11 concludes the proof. ■

Remark 4.13. As a conclusion we observe that an $L^r(U)$-valued Itô process, given by $\mathrm{d}X = \sum_{n=1}^\infty b_n \,\mathrm{d}\beta_n$, is nothing but an operator-valued integral expanded with respect to an orthonormal basis of a Hilbert space H. ■

4.2 The Itô Formula

By computing a deterministic integral we nearly always use the fundamental theorem of calculus since only a few integrals can be done comfortably by applying the definition of the integral. In the case of a stochastic integral it is very similar, and thus it would be quite neat if we would have an analogue to the fundamental theorem of calculus. In the case where $L^r(U) = \mathbb{R}^N$ for an $N \in \mathbb{N}$, there exists the well-known Itô formula, which can be found in many books about stochastic integration (cf. [1, (5.3.8)] for the proof of the following proposition).

From this point on, we return to general Itô processes as defined in (4.1).

Proposition 4.14 (Itô's formula - the finite-dimensional case). *Fix $M, N \in \mathbb{N}$. Let $\Phi\colon [0,T] \times \mathbb{R}^N \to \mathbb{R}$ be an element of $C^{1,2}([0,T] \times \mathbb{R}^N)$, $(\beta_m)_{m=1}^M$ be an M-dimensional Brownian motion, and let X be an $\mathbb{R}^N$-valued Itô process given by $\mathrm{d}X = f\,\mathrm{d}t + \sum_{m=1}^M b_m\,\mathrm{d}\beta_m$. Then, almost surely we have for all $t \in [0,T]$*

$$\begin{aligned}\Phi\big(t, X(t)\big) = \Phi\big(0, X(0)\big) &+ \int_0^t \partial_t \Phi\big(s, X(s)\big)\,\mathrm{d}s + \int_0^t D_2\Phi\big(s, X(s)\big) f(s)\,\mathrm{d}s \\ &+ \sum_{m=1}^M \int_0^t D_2\Phi\big(s, X(s)\big) b_m(s)\,\mathrm{d}\beta_m(s) \\ &+ \frac{1}{2}\int_0^t \sum_{m=1}^M \Big(D_2^2\Phi\big(s, X(s)\big) b_m(s)\Big) b_m(s)\,\mathrm{d}s.\end{aligned}$$

To extend this formula to functions $\Phi\colon [0,T] \times L^r(U_1) \to L^s(U_2)$ and $L^r(U_1)$-valued Itô processes, we need the following lemma.

Lemma 4.15. *Let X be an Itô process given by $\mathrm{d}X = f\,\mathrm{d}t + \sum_{m=1}^M b_m\,\mathrm{d}\beta_m$, and additionally let $b_m \in L^0(\Omega; L^2([0,T]; L^r(U)))$ for each $m = 1, \ldots, M$. Then there exist sequences $(f_n)_{n=1}^\infty$ and $(b_n^{(m)})_{n=1}^\infty$ of adapted step processes such that*

$$\begin{aligned} f &= \lim_{n\to\infty} f_n \quad \text{in } L^0(\Omega; L^1([0,T]; L^r(U))) \quad \text{and} \\ b_m &= \lim_{n\to\infty} b_n^{(m)} \quad \text{in } L^0(\Omega; L^2([0,T], L^r(U))) \cap L^0(\Omega; L^r_{\gamma,T}) \end{aligned}$$

for all $m = 1, \ldots, M$.

Proof. **(1)** Similar as in the proof of Lemma 3.21, we define for $n \in \mathbb{N}$

$$\widetilde{f}_n(t,\omega) := \big(A_n f_{\delta_n}(\omega)\big)(t) = \sum_{j=1}^{2^n} \mathbb{1}_{[\frac{(j-1)T}{2^n},\frac{jT}{2^n})}(t)\frac{2^n}{T}\int_{\frac{(j-1)T}{2^n}}^{\frac{jT}{2^n}} f(s-\delta_n,\omega)\,\mathrm{d}s,$$

where A_n is the operator from Proposition 3.12, $f_{\delta_n} := f(\cdot-\delta_n)$, and $\delta_n := \frac{T}{2^n}$. As in the proof of that Lemma, we deduce that

$$v_{j,n} := \int_{\frac{(j-1)T}{2^n}}^{\frac{jT}{2^n}} f(s-\delta_n,\cdot)\,\mathrm{d}s$$

is an $\mathscr{F}_{\frac{(j-1)T}{2^n}}$-measurable $L^r(U)$-valued random variable, i.e., $\widetilde{f}_n$ is adapted to the filtration $\mathbb{F}$. Observe that $f, f_{\delta_n} \in L^0(\Omega; L^1([0,T]; L^r(U)))$, and by the dominated convergence theorem, we have

$$\lim_{n\to\infty} \|f_{\delta_n} - f\|_{L^1([0,T];L^r(U))} = 0.$$

Therefore, by Proposition 3.12, almost surely we have

$$\begin{aligned}
&\lim_{n\to\infty} \|\widetilde{f}_n - f\|_{L^1([0,T];L^r(U))} \\
&\quad\le \lim_{n\to\infty}\big(\|\widetilde{f}_n - A_n f\|_{L^1([0,T];L^r(U))} + \|A_n f - f\|_{L^1([0,T];L^r(U))}\big) \\
&\quad\le \lim_{n\to\infty}\big(\|f_{\delta_n} - f\|_{L^1([0,T];L^r(U))} + \|A_n f - f\|_{L^1([0,T];L^r(U))}\big) = 0,
\end{aligned}$$

and hence the convergence also holds in $L^0(\Omega; L^1([0,T]; L^r(U)))$. In general, the processes $\widetilde{f}_n$ are not adapted step processes. Thus, by approximating each $v_{j,n}$ in probability by a sequence of $\mathscr{F}_{\frac{(j-1)T}{2^n}}$-measurable simple functions, as in the proof of Lemma 3.21, we obtain a sequence of adapted step processes $(f_{n,m})_{m=1}^{\infty}$ such that $\lim_{m\to\infty} f_{n,m} = \widetilde{f}_n$ in $L^0(\Omega; L^1([0,T]; L^r(U)))$. By taking an appropriate subsequence $(m_n)_{n=1}^{\infty}$, we finally obtain the required sequence $(f_n)_{n=1}^{\infty} = (f_{n,m_n})_{n=1}^{\infty}$.

(2) Fix an $m \in \{1,\dots,M\}$. We then define the sequence $(b_n^{(m)})_{n=1}^{\infty}$ as in the first part of this proof (just replace f by b_m). Since $b_m \in L^0(\Omega; L^2([0,T]; L^r(U)))$, Proposition 3.12 implies that

$$\lim_{n\to\infty} b_n^{(m)} = b_m \quad \text{in } L^0(\Omega; L^2([0,T]; L^r(U))).$$

Since $b_m \in L^0(\Omega; L^r_{\gamma,T})$, we next can apply an embedding argument similar as in the proof of Lemma 3.21 to obtain

$$\lim_{n\to\infty} b_n^{(m)} = b_m \quad \text{in } L^0(\Omega; L^r_{\gamma,T}).$$

■

In the next theorem we assume that $1 < s < \infty$, and that (U_1, Σ_1, μ_1) and (U_2, Σ_2, μ_2) are σ-finite measure spaces with countably generated σ-algebras Σ_1 and Σ_2.

Theorem 4.16 (Itô's formula). *Assume that* $\Phi\colon [0,T] \times L^r(U_1) \to L^s(U_2)$ *is an element of* $C^{1,2}([0,T] \times L^r(U_1); L^s(U_2))$, $(\beta_m)_{m=1}^\infty$ *is a sequence of independent Brownian motions, and* X *is an* $L^r(U_1)$*-valued Itô process given by* $\mathrm{d}X = f\,\mathrm{d}t + \sum_{m=1}^\infty b_m\,\mathrm{d}\beta_m$. *Further, let* $(b_m)_{m=1}^\infty \in L^0(\Omega; L^2([0,T]; l^2(L^r(U_1))))$. *Then, almost surely we have for all* $t \in [0,T]$

$$\begin{aligned}\Phi\big(t, X(t)\big) = \Phi\big(0, X(0)\big) &+ \int_0^t \partial_t \Phi\big(s, X(s)\big)\,\mathrm{d}s + \int_0^t D_2\Phi\big(s, X(s)\big) f(s)\,\mathrm{d}s \\ &+ \sum_{m=1}^\infty \int_0^t D_2\Phi\big(s, X(s)\big) b_m(s)\,\mathrm{d}\beta_m(s) \qquad (4.2)\\ &+ \frac{1}{2}\int_0^t \sum_{m=1}^\infty \Big(D_2^2\Phi\big(s, X(s)\big) b_m(s)\Big) b_m(s)\,\mathrm{d}s.\end{aligned}$$

Remark 4.17. In the situation of Theorem 4.16, the definition of an Itô process implies that $(b_m)_{m=1}^\infty \in L^0(\Omega; L^r(U_1; L^2([0,T]; l^2)))$. If we furthermore assume that $r \le 2$, Minkowski's integral inequality yield

$$L^r(U_1; L^2([0,T]; l^2)) \subseteq L^2([0,T]; l^2(L^r(U_1))).$$

Therefore, the assumption that $(b_m)_{m=1}^\infty \in L^0(\Omega; L^2([0,T]; l^2(L^r(U_1))))$ is automatically fulfilled. On the other hand, if $r \ge 2$, then we have

$$L^2([0,T]; l^2(L^r(U_1))) \subseteq L^r(U_1; L^2([0,T]; l^2))$$

and the additional assumption will be necessary. ∎

Proof (of Theorem 4.16). The proof is divided into four steps.

(1) We first show that (4.2) is true if Φ is an $\mathbb{R}$-valued function and we have an M-dimensional Brownian motion as well as an Itô process X represented by a simple function x_0 and adapted step processes f and $(b_m)_{m=1}^M$. In this case, the integer

$$N := \max\big\{\dim\big(\mathrm{R}(x_0)\big), \dim\big(\mathrm{R}(f)\big), \dim\big(\mathrm{R}(b_1)\big), \ldots, \dim\big(\mathrm{R}(b_M)\big)\big\}$$

is finite, which means that these processes take their values in an N-dimensional subspace E of $L^r(U_1)$. Therefore, we have an isomorphism $I\colon E \to \mathbb{R}^N$.

Since f and each b_m are adapted step processes, almost surely we trivially have

$$\int_0^T \|I(f)\|_{\mathbb{R}^N}\,\mathrm{d}t < \infty \quad \text{and} \quad \Big\| \Big(\int_0^T \sum_{m=1}^M |I(b_m)|^2\,\mathrm{d}t \Big)^{\frac{1}{2}} \Big\|_{\mathbb{R}^N} < \infty$$

Additionally, by a direct computation using the linearity of I, almost surely we obtain

$$I(X) = I(x_0) + \int_0^{\cdot} I(f(s))\,\mathrm{d}s + \sum_{m=1}^M \int_0^{\cdot} I(b_m(s))\,\mathrm{d}\beta_m(s).$$

So, $I(X)$ is a well-defined Itô process. We next define the function

$$\widetilde{\Phi}\colon [0,T]\times\mathbb{R}^N \to \mathbb{R}, \quad \widetilde{\Phi}(t,x) = \Phi(t, I^{-1}(x)).$$

Then $\widetilde{\Phi} \in C^{1,2}([0,T]\times\mathbb{R}^N)$ with

$$\begin{aligned}
&\partial_t\widetilde{\Phi}(t,x) = \partial_t\Phi\big(t, I^{-1}(x)\big),\\
&D_2\widetilde{\Phi}(t,x)\colon \mathbb{R}^N \to \mathbb{R}, \quad h \mapsto D_2\Phi\big(t, I^{-1}(x)\big)I^{-1}(h),\\
&D_2^2\widetilde{\Phi}(t,x)\colon \mathbb{R}^N \to \mathscr{B}(\mathbb{R}^N,\mathbb{R}) \quad h \mapsto \Big(D_2^2\Phi\big(t, I^{-1}(x)\big)I^{-1}(h)\Big)I^{-1}.
\end{aligned}$$

Finally, Proposition 4.14 leads to

$$\begin{aligned}
\Phi\big(t, X(t)\big) = \widetilde{\Phi}\big(t, I(X(t))\big) &= \widetilde{\Phi}\big(0, I(X(0))\big) + \int_0^t \partial_t\widetilde{\Phi}\big(s, I(X(s))\big)\,\mathrm{d}s\\
&+ \int_0^t D_2\widetilde{\Phi}\big(s, I(X(s))\big)I(f(s))\,\mathrm{d}s\\
&+ \sum_{m=1}^\infty \int_0^t D_2\widetilde{\Phi}\big(s, I(X(s))\big)I(b_m(s))\,\mathrm{d}\beta_m(s)\\
&+ \frac{1}{2}\int_0^t \sum_{m=1}^\infty \Big(D_2^2\widetilde{\Phi}\big(s, I(X(s))\big)I(b_m(s))\Big)I(b_m(s))\,\mathrm{d}s\\
&= \Phi\big(0, X(0)\big) + \int_0^t \partial_t\Phi\big(s, X(s)\big)\,\mathrm{d}s + \int_0^t D_2\Phi\big(s, X(s)\big)f(s)\,\mathrm{d}s\\
&+ \sum_{m=1}^\infty \int_0^t D_2\Phi\big(s, X(s)\big)b_m(s)\,\mathrm{d}\beta_m(s)\\
&+ \frac{1}{2}\int_0^t \sum_{m=1}^\infty \Big(D_2^2\Phi\big(s, X(s)\big)b_m(s)\Big)b_m(s)\,\mathrm{d}s.
\end{aligned}$$

(2) We next extend (4.2) to the case where X is represented by arbitrary adapted processes f and $b_1, \dots, b_M$. By the assumption of the theorem we have $b_m \in L^0(\Omega; L^2([0,T]; L^r(U_1))) \cap L^0(\Omega; L^r_{\gamma,T})$ for any $m = 1, \dots, M$.

Define the sequence $(X_n)_{n=1}^\infty$ in $L^0(\Omega; C([0,T]; L^r(U_1)))$ by

$$X_n(t) := x_{0,n} + \int_0^t f_n \, \mathrm{d}s + \sum_{m=1}^M \int_0^t b_n^{(m)} \, \mathrm{d}\beta_m,$$

where $(x_{0,n})_{n=1}^\infty$ is a sequence of $\mathscr{F}_0$-measurable simple functions with $x_0 = \lim_{n\to\infty} x_{0,n}$ almost surely, and $(f_n)_{n=1}^\infty$ and $(b_n^{(m)})_{n=1}^\infty$ are taken from Lemma 4.15. By Theorem 3.40 we have $X = \lim_{n\to\infty} X_n$ in $L^0(\Omega; C([0,T]; L^r(U_1)))$. By passing to a subsequence we may assume that there exists a set $\Omega_0 \subseteq \Omega$ of full measure such that

$$X = \lim_{n\to\infty} X_n \quad \text{in } C([0,T]; L^r(U_1)) \text{ on } \Omega_0. \tag{$\star_1$}$$

From this we deduce that

$$\lim_{n\to\infty} \Phi\big(t, X_n(t)\big) - \Phi\big(0, X_n(0)\big) = \Phi\big(t, X(t)\big) - \Phi\big(0, X(0)\big) \quad \text{on } \Omega_0.$$

For a continuous function $\Psi\colon [0,T] \times L^r(U_1) \to E$, where E is some Banach space, and $\omega \in \Omega_0$ fixed, the set

$$\bigcup_{n=1}^\infty \big\{\Psi\big(s, X_n(s,\omega)\big) \colon s \in [0,T]\big\} \cup \big\{\Psi\big(s, X(s,\omega)\big) \colon s \in [0,T]\big\}$$

is bounded. By applying this to the functions $\partial_t \Phi$, $D_2\Phi$ and $D_2^2\Phi$, the random variable

$$\begin{aligned} K = K(\omega) := \max_{t\in[0,T]} \big\{ &|\partial_t \Phi\big(t, X(t)\big)|, \|D_2\Phi\big(t, X(t)\big)\|, \|D_2^2\Phi\big(t, X(t)\big)\|, \\ &|\partial_t \Phi\big(t, X_n(t)\big)|, \|D_2\Phi\big(t, X_n(t)\big)\|, \|D_2^2\Phi\big(t, X_n(t)\big)\| \big\} \end{aligned}$$

is finite for each fixed $\omega \in \Omega_0$. Thus, by $(\star_1)$ and the dominated convergence theorem, we obtain on Ω_0

$$\lim_{n\to\infty} \int_0^t \partial_t \Phi\big(s, X_n(s)\big) \, \mathrm{d}s = \int_0^t \partial_t \Phi\big(s, X(s)\big) \, \mathrm{d}s.$$

Since

$$\|D_2\Phi(\cdot, X) f\|_{L^1([0,T])} \le K \|f\|_{L^1([0,T]; L^r(U_1))} < \infty,$$

$(\star_1)$ and the dominated convergence theorem yield

$$\lim_{n\to\infty} \|D_2\Phi(\cdot, X_n) f - D_2\Phi(\cdot, X) f\|_{L^1([0,T])} = 0.$$

And by Lemma 4.15 we have

$$\lim_{n\to\infty} \|D_2\Phi(\cdot, X_n) f_n - D_2\Phi(\cdot, X_n) f\|_{L^1([0,T])} \le \lim_{n\to\infty} K \|f_n - f\|_{L^1([0,T]; L^r(U_1))} = 0.$$

Together these estimates give $\lim_{n\to\infty}\|D_2\Phi(\cdot,X_n)f_n - D_2\Phi(\cdot,X)f\|_{L^1([0,T])} = 0$, and this implies that on Ω_0

$$\lim_{n\to\infty}\int_0^t D_2\Phi\big(s,X_n(s)\big)f_n(s)\,\mathrm{d}s = \int_0^t D_2\Phi\big(s,X(s)\big)f(s)\,\mathrm{d}s.$$

Now fix an $m\in\{1,\dots,M\}$, and observe that $b_m(\omega)\in L^2([0,T];L^r(U_1))$ for each $\omega\in\Omega_0$. Thus, similar as we just did, we obtain by Lemma 4.15, $(\star_1)$, and the dominated convergence theorem,

$$\begin{aligned}
&\lim_{n\to\infty}\|D_2\Phi(\cdot,X_n)b_n^{(m)} - D_2\Phi(\cdot,X)b_m\|_{L^2([0,T])}\\
&\le \lim_{n\to\infty}\big(\|D_2\Phi(\cdot,X_n)(b_n^{(m)}-b_m)\|_{L^2([0,T])} + \big\|\big(D_2\Phi(\cdot,X_n)-D_2\Phi(\cdot,X)\big)b_m\big\|_{L^2([0,T])}\big)\\
&\le \lim_{n\to\infty}\big(K\|b_n^{(m)}-b_m\|_{L^2([0,T];L^r(U_1))} + \big\|\big(D_2\Phi(\cdot,X_n)-D_2\Phi(\cdot,X)\big)b_m\big\|_{L^2([0,T])}\big) = 0.
\end{aligned}$$

By Theorem 3.40 (for the $\mathbb{R}$-valued case), this implies

$$\lim_{n\to\infty}\int_0^t D_2\Phi\big(s,X_n(s)\big)b_n^{(m)}(s)\,\mathrm{d}\beta_m(s) = \int_0^t D_2\Phi\big(s,X(s)\big)b_m(s)\,\mathrm{d}\beta_m(s)$$

on Ω_0. For each summand in the last term in (4.2) we have

$$\begin{aligned}
&\big\|\big(D_2^2\Phi(\cdot,X)b_m\big)b_m - \big(D_2^2\Phi(\cdot,X_n)b_n^{(m)}\big)b_n^{(m)}\big\|_{L^1([0,T])}\\
&\quad\le \big\|\big(D_2^2\Phi(\cdot,X)b_m\big)b_m - \big(D_2^2\Phi(\cdot,X_n)b_m\big)b_m\big\|_{L^1([0,T])}\\
&\qquad + \big\|\big(D_2^2\Phi(\cdot,X_n)b_m\big)b_m - \big(D_2^2\Phi(\cdot,X_n)b_n^{(m)}\big)b_n^{(m)}\big\|_{L^1([0,T])}.
\end{aligned}$$

Using the Cauchy-Schwarz inequality, we get

$$\|D_2^2\big(\Phi(\cdot,X)b_m\big)b_m\|_{L^1([0,T])} \le K\|b_m\|^2_{L^2([0,T];L^r(U_1))} < \infty,$$

and thus, by $(\star_1)$ and the dominated convergence theorem, the first term tends to 0 on Ω_0 . For the second term, by the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\big\|\big(D_2^2\Phi(\cdot,X_n)b_m\big)b_m - \big(D_2^2\Phi(\cdot,X_n)b_n^{(m)}\big)b_n^{(m)}\big\|_{L^1([0,T])}\\
&\quad\le \big\|\big(D_2^2\Phi(\cdot,X_n)b_m\big)b_m - \big(D_2^2\Phi(\cdot,X_n)b_n^{(m)}\big)b_m\big\|_{L^1([0,T])}\\
&\qquad + \big\|\big(D_2^2\Phi(\cdot,X_n)b_n^{(m)}\big)b_m - \big(D_2^2\Phi(\cdot,X_n)b_n^{(m)}\big)b_n^{(m)}\big\|_{L^1([0,T])}\\
&\quad\le K\|b_n^{(m)}-b_m\|_{L^2([0,T];L^r(U_1))}\|b_m\|_{L^2([0,T];L^r(U_1))}\\
&\qquad + K\|b_n^{(m)}-b_m\|_{L^2([0,T];L^r(U_1))}\|b_n^{(m)}\|_{L^2([0,T];L^r(U_1))},
\end{aligned}$$

which tends to 0 on Ω_0 by Lemma 4.15. Finally, putting all these estimates together and applying **(1)** to each X_n, we have shown that (4.2) is true for arbitrary adapted processes f and $b_1,\dots,b_M$.

(3) We next assume that X is an arbitrary $L^r(U_1)$-valued Itô process given by $\mathrm{d}X = f\,\mathrm{d}t + \sum_{m=1}^\infty b_m\,\mathrm{d}\beta_m$, where $(\beta_m)_{m=1}^\infty$ is a sequence of independent Brownian motions. For $M \in \mathbb{N}$ we define

$$X_M := x_0 + \int_0^\cdot f\,\mathrm{d}t + \sum_{m=1}^M \int_0^\cdot b_m\,\mathrm{d}\beta_m.$$

Since $\big\|\big(\int_0^T \sum_{m=1}^\infty |b_m|^2\,\mathrm{d}t\big)^{\frac{1}{2}}\big\|_r < \infty$, Theorem 4.7 implies that $X = \lim_{M\to\infty} X_M$ in $L^0(\Omega; C([0,T]; L^r(U_1)))$. Hence, by passing to a subsequence, we may choose $\Omega_0 \subseteq \Omega$ of full measure such that

$$X = \lim_{M\to\infty} X_M \quad \text{in } C([0,T]; L^r(U_1)) \text{ on } \Omega_0. \tag{$\star_2$}$$

Let $K\colon \Omega \to \mathbb{R}$ be as in the second part of this proof. Then, by $(\star_2)$, we have

$$\lim_{M\to\infty} \Phi\big(t, X_M(t)\big) - \Phi\big(0, X_M(0)\big) = \Phi\big(t, X(t)\big) - \Phi\big(0, X(0)\big) \quad \text{on } \Omega_0.$$

Additionally, by $(\star_2)$ and the dominated convergence theorem, we obtain on Ω_0

$$\begin{aligned} \lim_{M\to\infty} \int_0^t \partial_t \Phi\big(s, X_M(s)\big)\,\mathrm{d}s &= \int_0^t \partial_t \Phi\big(s, X(s)\big)\,\mathrm{d}s, \\ \lim_{M\to\infty} \int_0^t D_2\Phi\big(s, X_M(s)\big) f(s)\,\mathrm{d}s &= \int_0^t D_2\Phi\big(s, X(s)\big) f(s)\,\mathrm{d}s. \end{aligned}$$

Next, we take a look at the stochastic integral terms. Here, we have

$$\lim_{M\to\infty} \int_0^T \sum_{m=M+1}^\infty \big|D_2\Phi\big(s, X_M(s)\big) b_m\big|^2\,\mathrm{d}s \le \lim_{M\to\infty} K^2 \int_0^T \sum_{m=M+1}^\infty \|b_m(s)\|_r^2\,\mathrm{d}s = 0,$$

by the assumption we made to the sequence $(b_m)_{m=1}^\infty$. In addition, by $(\star_2)$ and the dominated convergence theorem, we obtain

$$\lim_{M\to\infty} \int_0^T \sum_{m=1}^\infty \Big|\Big(D_2\Phi\big(s, X_M(s)\big) - D_2\Phi\big(s, X(s)\big)\Big) b_m\Big|^2\,\mathrm{d}s = 0.$$

Let $(\widetilde{b}_m)_{m=1}^\infty$ be given by $\widetilde{b}_m = b_m$ for $1 \le m \le M$ and $\widetilde{b_m} = 0$ for $m > M$. Then, by the previous estimates, we obtain

$$\begin{aligned} &\lim_{M\to\infty} \int_0^T \sum_{m=1}^\infty \big|D_2\Phi\big(s, X_M(s)\big)\widetilde{b}_m - D_2\Phi\big(s, X(s)\big) b_m\big|^2\,\mathrm{d}s \\ &\le \lim_{M\to\infty} \Big(\int_0^T \sum_{m=M+1}^\infty \big|D_2\Phi\big(s, X_M(s)\big) b_m\big|^2\,\mathrm{d}s \\ &\qquad + \int_0^T \sum_{m=1}^\infty \Big|\Big(D_2\Phi\big(s, X_M(s)\big) - D_2\Phi\big(s, X(s)\big)\Big) b_m\Big|^2\,\mathrm{d}s\Big) = 0. \end{aligned}$$

By Theorem 4.7, this leads to

$$\lim_{M\to\infty}\sum_{m=1}^{M}\int_0^t D_2\Phi\big(s,X_M(s)\big)b_m\,\mathrm{d}\beta_m=\sum_{m=1}^{\infty}\int_0^t D_2\Phi\big(s,X(s)\big)b_m\,\mathrm{d}\beta_m\quad\text{on }\Omega_0.$$

For the last term in (4.2), we have

$$\begin{aligned}&\Big\|\sum_{m=1}^{\infty}\big(D_2^2\Phi(\cdot,X)b_m\big)b_m-\sum_{m=1}^{M}\big(D_2^2\Phi(\cdot,X_M)b_m\big)b_m\Big\|_{L^1([0,T])}\\&\quad\le\Big\|\sum_{m=1}^{\infty}\big(D_2^2\Phi(\cdot,X)b_m\big)b_m-\sum_{m=1}^{\infty}\big(D_2^2\Phi(\cdot,X_M)b_m\big)b_m\Big\|_{L^1([0,T])}\\&\qquad+\Big\|\sum_{m=1}^{\infty}\big(D_2^2\Phi(\cdot,X_M)b_m\big)b_m-\sum_{m=1}^{M}\big(D_2^2\Phi(\cdot,X_M)b_m\big)b_m\Big\|_{L^1([0,T])}.\end{aligned}$$

Again, by ($\star_2$) and the dominated convergence theorem, the first term tends to 0 on Ω_0. For the second term, we have

$$\begin{aligned}&\Big\|\sum_{m=1}^{\infty}\big(D_2^2\Phi(\cdot,X_M)b_m\big)b_m-\sum_{m=1}^{M}\big(D_2^2\Phi(\cdot,X_M)b_m\big)b_m\Big\|_{L^1([0,T])}\\&\quad=\Big\|\sum_{m=M+1}^{\infty}\big(D_2^2\Phi(\cdot,X_M)b_m\big)b_m\Big\|_{L^1([0,T])}\\&\quad\le K\Big\|\Big(\sum_{m=M+1}^{\infty}\|b_m\|_r^2\Big)^{\frac12}\Big\|_{L^2([0,T])}^2,\end{aligned}$$

which also tends to 0 on Ω_0, by the assumption we made to the sequence $(b_m)_{m=1}^{\infty}$. Again, by collecting all estimates we have established in this part, and by applying **(2)** to each X_M, we get the desired result.

(4) Finally, we show (4.2) as posted in the theorem, i.e., Φ is now an $L^s(U_2)$-valued function. For all $g\in L^{s'}(U_2)$ we almost surely have

$$\begin{aligned}&\int_0^T\sum_{m=1}^{\infty}\Big|\big\langle D_2\Phi\big(s,X(s)\big)b_m(s),g\big\rangle\Big|^2\,\mathrm{d}s\\&\qquad\le\|g\|_{L^{s'}(U_2)}^2\max_{t\in[0,T]}\Big\|D_2\Phi\big(t,X(t)\big)\Big\|^2\int_0^T\sum_{m=1}^{\infty}\|b_m(s)\|_r^2\,\mathrm{d}s<\infty,\end{aligned}$$

since $t\mapsto D_2\Phi\big(t,X(t)\big)$ is continuous. Now fix $g\in L^{s'}(U_2)$. Observe that differentiating Φ and applying $\langle\cdot,g\rangle$ are commuting operations.

With this in mind, part **(3)** applied to the $\mathbb{R}$-valued function $\langle \Phi, g\rangle$ leads to

$$\begin{aligned}\langle \Phi\big(t, X(t)\big), g\rangle &= \langle \Phi\big(0, X(0)\big), g\rangle + \Big\langle \int_0^t \partial_t \Phi\big(s, X(s)\big)\,\mathrm{d}s, g\Big\rangle \\ &+ \Big\langle \int_0^t D_2\Phi\big(s, X(s)\big) f(s)\,\mathrm{d}s, g\Big\rangle + \sum_{m=1}^\infty \int_0^t \langle D_2\Phi\big(s, X(s)\big) b_m(s), g\rangle\,\mathrm{d}\beta_m(s) \\ &+ \Big\langle \frac{1}{2}\int_0^t \sum_{m=1}^\infty \Big(D_2^2\Phi\big(s, X(s)\big) b_m(s)\Big) b_m(s)\,\mathrm{d}s, g\Big\rangle.\end{aligned}$$

An application of Theorem 4.8 to the pathwise continuous process

$$\begin{aligned}Y := \Phi\big(\cdot, X(\cdot)\big) - \Phi\big(0, X(0)\big) - \int_0^\cdot \partial_t\Phi\big(s, X(s)\big)\,\mathrm{d}s - \int_0^\cdot D_2\Phi\big(s, X(s)\big) f(s)\,\mathrm{d}s \\ - \frac{1}{2}\int_0^\cdot \sum_{m=1}^\infty \Big(D_2^2\Phi\big(s, X(s)\big) b_m(s)\Big) b_m(s)\,\mathrm{d}s\end{aligned}$$

shows that $\sum_{m=1}^\infty \int_0^t D_2\Phi\big(s, X(s)\big) b_m(s)\,\mathrm{d}\beta_m(s)$ is well-defined and equals almost surely Y, which means that (4.2) holds. ■

Corollary 4.18. *Let $1 < r_1 < \infty$, and define $r_2 := r_1'$. For $i = 1, 2$, let X_i be an $L^{r_i}(U)$-valued Itô process given by $\mathrm{d}X_i = f_i\,\mathrm{d}t + \sum_{m=1}^\infty b_{m,i}\,\mathrm{d}\beta_m$, which satisfies the assumptions of Theorem 4.16. Then, almost surely we have for all $t \in [0, T]$*

$$\begin{aligned}\langle X_1(t), X_2(t)\rangle = \langle X_1(0), X_2(0)\rangle &+ \int_0^t \langle X_1(s), f_2(s)\rangle + \langle f_1(s), X_2(s)\rangle\,\mathrm{d}s \\ &+ \sum_{m=1}^\infty \int_0^t \langle X_1(s), b_{m,2}(s)\rangle + \langle b_{m,1}(s), X_2(s)\rangle\,\mathrm{d}\beta_m(s) \qquad (4.3) \\ &+ \int_0^t \sum_{m=1}^\infty \langle b_{m,1}(s), b_{m,2}(s)\rangle\,\mathrm{d}s.\end{aligned}$$

Proof. Define $\Phi\colon [0, T] \times L^{r_1}(U) \times L^{r_1'}(U) \to \mathbb{R}$ by $\Phi(t, x, y) := \langle x, y\rangle$ (and observe that Theorem 4.16 could be proved for these functions in the exact same way with the same formula). Then Φ is of class $C^{1,2}$ with

$$\begin{aligned}&\partial_t\Phi(t, x, y) = 0, \\ &D_2\Phi(t, x, y)\colon L^{r_1}(U) \times L^{r_1'}(U) \to \mathbb{R}, \quad (h, h') \mapsto \langle h, y\rangle + \langle x, h'\rangle\,, \\ &D_2^2\Phi(t, x, y)\colon L^{r_1}(U) \times L^{r_1'}(U) \to \big(L^{r_1}(U) \times L^{r_1'}(U)\big)^*, \ (h, h') \mapsto \langle h, \cdot\rangle + \langle \cdot, h'\rangle\,.\end{aligned}$$

And finally, an application of Theorem 4.16 gives (4.3). ■

4.3 An Infinite Dimensional Version of the Geometric Brownian Motion

The final section is dedicated to an application of the theory developed in this thesis to a stochastic partial differential equation. In this context, we will use some notions and results from the theory of semigroups[1]. More precisely, we want to study the problem

$$U(t) = u_0 + \int_0^t AU(s)\,\mathrm{d}s + \sum_{n=1}^{N} \int_0^t B_n U(s)\,\mathrm{d}\beta_n(s), \quad t \in [0,T], \tag{4.4}$$

which we also write as

$$\begin{aligned} \mathrm{d}U(t) &= AU(t)\,\mathrm{d}t + \sum_{n=1}^{N} B_n U(t)\,\mathrm{d}\beta_n(t), \quad t \in [0,T], \\ U(0) &= u_0. \end{aligned}$$

Here, the processes $\beta_n = \big(\beta_n(t)\big)_{t\in[0,T]}$ are independent Brownian motions defined on some probability space $(\Omega, \mathscr{A}, \mathbb{P})$ and adapted to some filtration $\mathbb{F} = (\mathscr{F}_t)_{t\in[0,T]}$. We assume that the initial random variable $u_0\colon \Omega \to L^r(U)$ is $\mathscr{F}_0$-measurable, and regarding the operators $A\colon D(A) \subseteq L^r(U) \to L^r(U)$ and $B_n\colon D(B_n) \subseteq L^r(U) \to L^r(U)$ we make the following hypotheses:

(H1) The operator A is closed and densely defined.

(H2) The operators B_n generate commuting C_0-groups $G_n = \big(G_n(t)\big)_{t\in\mathbb{R}}$ on $L^r(U)$, which also commute with A.

(H3) We have $D(A) \subseteq \bigcap_{n=1}^{N} D(B_n^2)$.

Defining $D(C) := D(A)$ and $C := A - \frac{1}{2}\sum_{n=1}^{N} B_n^2$, we further assume:

(H4) The operator C generates a C_0-semigroup $S = \big(S(t)\big)_{t\geq 0}$ on $L^r(U)$, and $u_0 \in D(C)$ almost surely.

We call an $L^r(U)$-valued process $U = \big(U(t)\big)_{t\in[0,T]}$ a *strong solution* of (4.4) on the interval $[0,T]$ if $U \in C([0,T]; L^r(U))$ almost surely, $U(0) = u_0$, and the following conditions are satisfied:

(1) For almost all $\omega \in \Omega$, $U(t,\omega) \in D(A)$ for almost all $t \in [0,T]$ and the path $t \mapsto AU(t,\omega)$ belongs to $L^1([0,T]; L^r(U))$.

(2) For each $n = 1, \ldots, N$, the process $B_n U$ is an element of $L^0_{\mathbb{F}}(\Omega; L^r_{\gamma,T})$.

(3) Almost surely, U solves (4.4) on $[0,T]$.

[1]For more information on this topic we refer to [4] and [13].

Lemma 4.19. *If (H2) holds, then* $D' := \bigcap_{n=1}^{N} D(B_n^{*2})$ *is dense in* $L^{r'}(U)$.

Proof. Since $L^r(U)$ is reflexive, an application of the Hahn-Banach theorem says that a linear subspace $E \subseteq L^{r'}(U)$ is dense in $L^{r'}(U)$ if and only if for all $f \in L^r(U) \setminus \{0\}$ there exists a $g \in E$ with $\langle f, g \rangle \neq 0$. Thus fix an $f \in L^r(U) \setminus \{0\}$ and some $\lambda \in \bigcap_{n=1}^{N} \varrho(B_n)$. Take $\widetilde{f} := \prod_{n=1}^{N} R(\lambda, B_n)^2 f$. Since $\widetilde{f} \neq 0$, we can find a $\widetilde{g} \in L^{r'}(U)$ such that $\langle \widetilde{f}, \widetilde{g} \rangle \neq 0$. By (H2), the resolvents $R(\lambda, B_n^*)$ commute, which implies that $g := \prod_{n=1}^{N} R(\lambda, B_n^*)^2 \widetilde{g} \in \bigcap_{n=1}^{N} D(B_n^{*2})$. Finally, by construction, we have $\langle f, g \rangle = \langle \widetilde{f}, \widetilde{g} \rangle \neq 0$, and this concludes the proof. ■

Theorem 4.20. *Assuming (H1)-(H4), the problem* (4.4) *has a strong solution on* $[0, T]$.

Proof. Define $G \colon \mathbb{R}^N \to \mathscr{B}(L^r(U))$ as

$$G(x) := \prod_{n=1}^{N} G_n(x_n)$$

and the process $G_\beta \colon \Omega \times [0, T] \to \mathscr{B}(L^r(U))$ by

$$G_\beta(\omega, t) := G\big(\beta_1(\omega, t), \ldots, \beta_N(\omega, t)\big),$$

which is adapted and pathwise strongly continuous. Next, we introduce the following pathwise problem:

$$\begin{aligned} V'(t) &= CV(t), \quad t \in [0, T], \\ V(0) &= u_0. \end{aligned} \tag{4.5}$$

By (H4), the unique solution of this problem is $V(t) = S(t)u_0$. Moreover, almost surely we have $V \in C^1([0, T]; L^r(U))$ and $V(t) \in D(C) = D(A)$ for all $t \in [0, T]$. Put $U := G_\beta V$. By (H2) we then have $U(t) \in D(A)$, and since almost surely $CV \in C([0, T]; L^r(U))$ and since G_β is strongly continuous, we have $CU = CG_\beta V = G_\beta CV \in C([0, T]; L^r(U)) \subseteq L^1([0, T]; L^r(U))$. Let $\lambda \in \varrho(C) \cap \mathbb{R}$. Then we have

$$B_n^2 U = B_n^2 R(\lambda, C)(\lambda I - C) G_\beta V = \lambda B_n^2 R(\lambda, C) G_\beta V - B_n^2 R(\lambda, C) C G_\beta V,$$

which is almost surely continuous. Therefore, $B_n^2 U \in L^1([0, T]; L^r(U))$ almost surely, and this implies that $AU = CU + \frac{1}{2} \sum_{n=1}^{N} B_n^2 U \in L^1([0, T]; L^r(U))$ almost surely.

Let D' as in the previous lemma, and let $g \in D'$ be fixed. The function $\Phi \colon \mathbb{R}^N \to L^{r'}(U)$ defined by $\Phi(x) := G(x)^* g$ is twice continuously differentiable with

$$\frac{\partial \Phi}{\partial x_n}(x) = G(x)^* B_n^* g \quad \text{and} \quad \frac{\partial^2 \Phi}{\partial x_n^2}(x) = G(x)^* B_n^{*2} g.$$

We next define the $\mathbb{R}^N$-valued Itô process X^β by

$$X^\beta(t) := \sum_{n=1}^N \int_0^t e_n \, \mathrm{d}\beta_n = \big(\beta_1(t), \dots, \beta_N(t)\big),$$

where $(e_n)_{n=1}^N$ is the standard basis of $\mathbb{R}^N$. Hence, we have $\Phi(X^\beta(t)) = G_\beta(t)^* g$, and Itô's formula yield

$$G_\beta(t)^* g = G_\beta(0)^* g + \sum_{n=1}^N \int_0^t G_\beta(s)^* B_n^* g \, \mathrm{d}\beta_n(s) + \frac{1}{2} \sum_{n=1}^N \int_0^t G_\beta(s)^* B_n^{*2} g \, \mathrm{d}s.$$

Moreover, since V is the solution of (4.5), we have

$$V(t) = V(0) + \int_0^t CV(s) \, \mathrm{d}s.$$

Thus, by applying (4.3) to the Itô processes V and $G_\beta^* g$, we almost surely obtain for all $t \in [0, T]$

$$\begin{aligned}
\langle U(t), g\rangle - \langle U(0), g\rangle &= \langle V(t), G_\beta(t)^* g\rangle - \langle V(0), G_\beta(0)^* g\rangle \\
&= \int_0^t \frac{1}{2} \sum_{n=1}^N \big\langle V(s), G_\beta(s)^* B_n^{*2} g\big\rangle + \langle CV(s), G_\beta(s)^* g\rangle \ \mathrm{d}s \\
&\quad + \sum_{n=1}^N \int_0^t \langle V(s), G_\beta(s)^* B_n^* g\rangle \ \mathrm{d}\beta_n(s) \\
&= \int_0^t \langle A G_\beta(s) V(s), g\rangle \ \mathrm{d}s + \sum_{n=1}^N \int_0^t \langle B_n G_\beta(s) V(s), g\rangle \ \mathrm{d}\beta_n(s) \\
&= \int_0^t \langle AU(s), g\rangle \ \mathrm{d}s + \sum_{n=1}^N \int_0^t \langle B_n U(s), g\rangle \ \mathrm{d}\beta_n(s),
\end{aligned}$$

where we used (H2) and the definition of C. Since $U = G_\beta V$ is almost surely continuous and since $AU \in L^1([0,T]; L^r(U))$, the process

$$Y := U - U(0) - \int_0^\cdot AU(s) \, \mathrm{d}s$$

is almost surely continuous.

Finally, thanks to Lemma 4.19, we are now able to apply Remark 4.9, which yields that $B_n U \in L^0(\Omega; L^r_{\gamma,T})$ for each $n = 1, \ldots, N$, and also implies that we almost surely have

$$U(t) = u_0 + \int_0^t AU(s)\,\mathrm{d}s + \sum_{n=1}^{N} \int_0^t B_n U(s)\,\mathrm{d}\beta_n(s) \quad \text{for all } t \in [0,T]. \qquad \blacksquare$$

Appendix A

Appendix

A.1 Integration in Banach Spaces

In this section we want to give a brief introduction to the Bochner integral, which is a generalization of the Lebesgue integral to the Banach space valued setting. First we are going to show some measurability results and after that we will construct the Bochner integral.

A.1.1 Measurability

In what follows, let $(A, \mathscr{A})$ be a measurable space, E be an arbitrary real Banach space, and $\mathscr{B}(E)$ be the Borel σ-algebra of E. We call a function $f\colon A \to E$ *$\mathscr{A}$-simple* if it is of the form $f = \sum_{n=1}^{N} \mathbb{1}_{A_n} x_n$ with $A_n \in \mathscr{A}$ and $x_n \in E$ for all $1 \leq n \leq N$.

Definition A.1. *A function $f\colon A \to E$ is called* $\mathscr{A}$-measurable *if*

$$f^{-1}(B) \in \mathscr{A} \quad \text{for all } B \in \mathscr{B}(E),$$

and strongly $\mathscr{A}$-measurable *if there exists a sequence of $\mathscr{A}$-simple functions $f_n\colon A \to E$ such that*

$$\lim_{n\to\infty} f_n = f \quad \text{pointwise on } A.$$

Next, we give a characterization of strong $\mathscr{A}$-measurability of E-valued functions (cf. [3, Chapter II.1, Theorem 2]). Here we call a function $f\colon A \to E$ *separably valued* if there exists a separable closed subspace E_0 of E such that $f(a) \in E_0$ for all $a \in A$.

Theorem A.2 (Pettis measurability theorem I). *For any function $f\colon A \to E$ 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 $x^* \in E^*$.

As a consequence we obtain the following result.

Proposition A.3. *For a function $f\colon A \to E$ the following assertions are equivalent:*

(1) f is strongly $\mathscr{A}$-measurable;

(2) f is $\mathscr{A}$-measurable and separably valued.

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.

Proof. (1) $\Rightarrow$ (2): By Theorem A.2, f is separably valued. To show $\mathscr{A}$-measurability of f it suffices to prove that $f^{-1}(U) \in \mathscr{A}$ for any open subset $U \subseteq E$. Let U be open and choose a sequence of $\mathscr{A}$-simple functions f_n converging pointwise to f. Since U is open, we then get

$$f^{-1}(U) \subseteq \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} f_k^{-1}(U).$$

For the converse inclusion, let $U_r := \{x \in U : d(x, U^C) > r\}$. Then $U = \bigcup_{m=1}^{\infty} U_{\frac{1}{m}}$ and for each fixed $m \in \mathbb{N}$ we have

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

Putting both estimates together, we obtain

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

Since $f_k^{-1}(U_{\frac{1}{m}}) \in \mathscr{A}$, it follows that f is $\mathscr{A}$-measurable.

(2) $\Rightarrow$ (1): By the $\mathscr{A}$-measurability of f we infer that $\langle f, x^*\rangle$ is $\mathscr{A}$-measurable for all $x^* \in E^*$ and the result now follows from Theorem A.2. ∎

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 $(A, \mathscr{A}, \nu)$. A function $f\colon A \to E$ is called ν-simple if it is of the form $f = \sum_{n=1}^{N} \mathbb{1}_{A_n} x_n$ with $x_n \in E$, and the sets $A_n \in \mathscr{A}$ satisfy $\nu(A_n) < \infty$.

Definition A.4. *A function $f\colon A \to E$ is called* strongly ν-measurable *if there exists a sequence of ν-simple functions $f_n\colon A \to E$ such that*

$$\lim_{n\to\infty} f_n = f \quad \nu\text{-almost everywhere.}$$

Using the σ-finiteness of ν, we can show that every strongly $\mathscr{A}$-measurable function f is strongly ν-measurable. For this purpose, let $(f_n)_{n=1}^{\infty}$ be a sequence of $\mathscr{A}$-simple functions converging to f pointwise, and $A_1 \subseteq A_2 \subseteq \ldots$ in $\mathscr{A}$ with $A = \bigcup_{n=1}^{\infty} A_n$ and $\nu(A_n) < \infty$ for each $n \in \mathbb{N}$. Then also $\lim_{n\to\infty} \mathbb{1}_{A_n} f_n = f$ pointwise and each $\mathbb{1}_{A_n} f_n$ is ν-simple.

A similar result is also true for the converse direction. Here we call two functions *ν-versions* of each other if they agree ν-almost everywhere.

Proposition A.5. *For a function $f\colon A \to E$ the following assertions are equivalent:*

(1) f is strongly ν-measurable;

(2) f has a ν-version which is strongly $\mathscr{A}$-measurable.

Proof. (1) $\Rightarrow$ (2): Let $(f_n)_{n=1}^{\infty}$ be a sequence of ν-simple functions converging to f outside the null set $N \in \mathscr{A}$. Then the functions $\mathbb{1}_{N^C} f_n$ are $\mathscr{A}$-simple and satisfy $\lim_{n\to\infty} \mathbb{1}_{N^C} f_n = \mathbb{1}_{N^C} f$ pointwise on A. Thus $\mathbb{1}_{N^C} f$ is a strongly $\mathscr{A}$-measurable ν-version of f.

(2) $\Rightarrow$ (1): Let $\widetilde{f}$ be a strongly $\mathscr{A}$-measurable ν-version of f and $(\widetilde{f}_n)_{n=1}^{\infty}$ be a sequence of $\mathscr{A}$-simple functions converging to $\widetilde{f}$. Then $\lim_{n\to\infty} \widetilde{f}_n = f$ ν-almost everywhere. Moreover, write $A = \bigcup_{n=1}^{\infty} A_n$ with $A_1 \subseteq A_2 \subseteq \ldots \in \mathscr{A}$ and $\nu(A_n) < \infty$ for all $n \in \mathbb{N}$. By taking $f_n := \mathbb{1}_{A_n} \widetilde{f}_n$, we obtain a sequence of ν-simple functions $(f_n)_{n=1}^{\infty}$ with $\lim_{n\to\infty} f_n = f$ ν-almost everywhere. ■

By combining this proposition and Theorem A.2, we get the following result.

Theorem A.6 (Pettis measurability theorem II). *For any function $f\colon A \to E$ the following assertions are equivalent:*

(1) f is strongly ν-measurable;

(2) f has a ν-version $\widetilde{f}$ which is separably valued and $\langle \widetilde{f}, x^* \rangle$ is $\mathscr{A}$-measurable for all $x^* \in E^*$.

As a consequence of this theorem, we get the following results (cf. [3, Chapter II.1, Corollary 3]).

Corollary A.7. *(1) The ν-almost everywhere limit of a sequence of strongly ν-measurable E-valued functions is strongly ν-measurable.*

(2) If $f\colon A \to E$ is strongly ν-measurable and $\Phi\colon E \to F$ is continuous, where F is another Banach space, then $\Phi \circ f$ is strongly ν-measurable.

We finish the discussion of ν-measurability with a quite useful corollary (cf. [3, Chapter II.2, Corollary 7]).

Corollary A.8. *If f and g are strongly ν-measurable E-valued functions with $\langle f, x^* \rangle = \langle g, x^* \rangle$ ν-almost everywhere for every $x^* \in E^*$, then $f = g$ ν-almost everywhere.*

A.1.2 The Bochner Integral

We next concentrate on the construction of the *Bochner integral* and give a short introduction of the *Lebesgue-Bochner spaces.*

Definition A.9. *A function $f\colon A \to E$ is called ν-*Bochner integrable *if there exists a sequence of ν-simple functions $f_n\colon A \to E$ such that the following two conditions hold:*

(1) $\lim_{n\to\infty} f_n = f$ ν-almost everywhere (i.e., f is strongly ν-measurable);

(2) $\lim_{n\to\infty} \int_A \|f_n - f\|_E \,\mathrm{d}\nu = 0$.

From the definition it is easy to see that every ν-simple function is ν-Bochner integrable. For $f = \sum_{n=1}^N \mathbb{1}_{A_n} x_n$ we put

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

Similar to the Lebesgue integral, we can check that this definition is independent of the representation of f (and that implies that the integral is linear) . If f is ν-Bochner integrable, the limit

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

exists in E and is called the *Bochner integral* of f with respect to ν. It is routine to check that this definition is independent of the approximating sequence of f. Moreover, if f is ν-Bochner integrable and g is a ν-version of f, then g is ν-Bochner integrable, and the Bochner integrals of f and g agree.

In the next proposition we collect some properties regarding ν-Bochner integrability (cf. [3, Chapter II.2, Theorem 2,4 and 6]).

Proposition A.10. *Let $f\colon A \to E$ be strongly ν-measurable. Then the following assertions hold:*

(1) f is ν-Bochner integrable if and only if $\int_A \|f\|_E \, \mathrm{d}\nu < \infty$. In this case, we have

$$\left\| \int_A f \, \mathrm{d}\nu \right\|_E \leq \int_A \|f\|_E \, \mathrm{d}\nu.$$

(2) If f is ν-Bochner integrable and $T\colon E \to F$ is a bounded linear operator, where F is another Banach space, then $Tf\colon A \to F$ is ν-Bochner integrable and

$$T \int_A f \, \mathrm{d}\nu = \int_A Tf \, \mathrm{d}\nu.$$

Especially, we have

$$\left\langle \int_A f \, \mathrm{d}\nu, x^* \right\rangle = \int_A \langle f, x^* \rangle \, \mathrm{d}\nu$$

for all $x^* \in E^*$.

Next, we give an analogue of the dominated convergence theorem (cf. [3, Chapter II.2, Theorem 3]).

Proposition A.11 (Dominated convergence theorem). *Let $(f_n)_{n=1}^{\infty}$ be a sequence of E-valued ν-Bochner integrable functions. Assume that there exist a function $f\colon A \to E$ and a ν-Bochner integrable function $g\colon A \to \mathbb{R}$ such that:*

(1) $\lim_{n\to\infty} f_n = f$ *ν-almost everywhere;*

(2) $\|f_n\|_E \leq |g|$ *ν-almost everywhere.*

Then f is ν-Bochner integrable, and we have

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

In particular, we have

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

We finish this section with an introduction of the L^p spaces for Banach space valued functions, the so called *Lebesgue-Bochner spaces*. For $1 \leq p < \infty$ we define $L^p(A;E)$ as the linear space of all equivalence classes of strongly ν-measurable functions $f\colon A \to E$ satisfying

$$\int_A \|f\|_E^p \,\mathrm{d}\nu < \infty,$$

identifying functions which are equal ν-almost everywhere. As in the scalar case, we can show that the space $L^p(A;E)$ endowed with the norm

$$\|f\|_p := \Big(\int_A \|f\|_E^p \,\mathrm{d}\nu\Big)^{\frac{1}{p}}$$

is a Banach space. By the definition of the Bochner integral and Proposition A.10 (1), it is easy to see that ν-simple functions are dense in $L^p(A;E)$.

We define $L^\infty(A;E)$ as the linear space of all equivalence classes of strongly ν-measurable functions $f\colon A \to E$ such that there exists an $r \geq 0$ with

$$\nu\big(\|f\|_E > r\big) = 0.$$

Endowed with the norm

$$\|f\|_\infty := \inf\big\{r \geq 0\colon \nu\big(\|f\|_E > r\big) = 0\big\},$$

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

Remark A.12. In the case of equivalence classes, we will just say that f is *strongly measurable* if f has a strongly ν-measurable representative (in this case, every representative is strongly ν-measurable and by Proposition A.5 there even exists a representative which is strongly $\mathscr{A}$-measurable). So, in what follows, we will omit the prefix '$\nu-$' from our terminology if no confusion can arise. Especially, if E is separable, we will just say that a strongly measurable function f is *measurable* (motivated by Proposition A.3). ■

A.2 Gaussian Random Variables

Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space and E be an arbitrary real Banach space. An E-valued *random variable* is an E-valued strongly measurable function $X\colon \Omega \to E$. For E-valued random variables, the definitions from the $\mathbb{R}$-valued case carries over nearly verbatim. For example, the Bochner integral of an integrable random variable X is called its *mean value* or *expectation*, and is denoted by

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

Moreover, the *distribution* of an E-valued random variable X is the Borel probability measure $\mathbb{P}_X$ on E defined by

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

Definition A.13. *The* Fourier transform *of a Borel probability measure μ on E is the function $\widehat{\mu}\colon E^* \to \mathbb{C}$ defined by*

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

The Fourier transform *of a random variable $X\colon \Omega \to E$ is the Fourier transform of its distribution $\mathbb{P}_X$.*

Note that $\widehat{\mu}$ is well-defined since $\big|\exp\big(-i \langle x, x^*\rangle\big)\big| = 1$, which implies that the integral is absolutely convergent. By a change of variable, the Fourier transform of a random variable X on E is given by

$$\widehat{X}(x^*) = \mathbb{E}\exp\big(-i \langle X, x^*\rangle\big) = \int_E \exp\big(-i \langle x, x^*\rangle\big) \,\mathrm{d}\mathbb{P}_X(x).$$

We next show a uniqueness result (cf. [12, Chapter IV, Theorem 3.1 and Lemma 5.2]).

Theorem A.14. *Let X_1 and X_2 be E-valued random variables who satisfy*

$$\widehat{X_1}(x^*) = \widehat{X_2}(x^*) \quad \text{for all } x^* \in E^*.$$

Then X_1 and X_2 are identically distributed.

Next, we proceed with *Gaussian random variables.* An $\mathbb{R}$-valued random variable γ is called *Gaussian* if there exists a number $q \geq 0$ such that its Fourier transform is given by

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

Note that this definition is consistent with the one given in Chapter 1.

An E-valued random variable X is said to be *Gaussian* if the $\mathbb{R}$-valued random variable $\langle X, x^*\rangle$ is Gaussian for all $x^* \in E^*$. Based on the uniqueness theorem of the Fourier transform, we can now show the following properties of E-valued Gaussian random variables.

Proposition A.15. *Let X and Y be independent and identically distributed E-valued Gaussian random variables. Then $U := \frac{1}{\sqrt{2}}(X+Y)$ and $V := \frac{1}{\sqrt{2}}(X-Y)$ are independent and have the same distribution as X and Y.*

Proof. We have $\widehat{X}(x^*) = \widehat{Y}(x^*) = \exp\big(-\frac{1}{2}q(x^*)\big)$, where $q(x^*) = \mathbb{E}\langle X, x^*\rangle^2 = \mathbb{E}\langle Y, x^*\rangle^2$. By the independence of X and Y, we obtain

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

So, by Theorem A.14, U has the same distribution as X and Y. By a similar computation we get the same result for V and the independence of U and V. ■

As a further application we prove next that if the E-valued random variables $X_1, \ldots, X_N$ are jointly Gaussian, that is, if the E^N-valued random variable $X := (X_1, \ldots, X_N)$ is Gaussian, then $X_1, \ldots, X_N$ are independent if and only if they are uncorrelated in the sense that

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

Proposition A.16. *Let $X_1, \dots, X_N$ be E-valued random variables such that the E^N-valued random variable $(X_1, \dots, X_N)$ is Gaussian. Then the following assertions are equivalent:*

(1) $X_1, \dots, X_N$ are independent;

(2) $X_1, \dots, X_N$ are uncorrelated.

Proof. For the $\mathbb{R}$-valued case see [7, Lemma 13.1] or [9, Satz 7.33].

(1) $\Rightarrow$ (2): Since $X_1, \dots, X_N$ are independent, it follows that the random variables $\langle X_1, x_1^* \rangle, \dots, \langle X_N, x_N^* \rangle$ are independent for any $x_1^*, \dots, x_N^* \in E^*$. The implication therefore follows from the corresponding implication in the $\mathbb{R}$-valued case.

(2) $\Rightarrow$ (1): Note that $\big(\langle X_1, x_1^* \rangle, \dots, \langle X_N, x_N^* \rangle\big)$ is again an $\mathbb{R}^N$-valued Gaussian random variable for any $x_1^*, \dots, x_N^* \in E^*$. So, by (2) and the foregoing remark, the $\mathbb{R}$-valued random variables $\langle X_1, x_1^* \rangle, \dots, \langle X_N, x_N^* \rangle$ are independent. Now we obtain

$$
\begin{aligned}
\widehat{\mathbb{P}}_{(X_1,\dots,X_N)}(x_1^*, \dots, x_N^*) &= \mathbb{E} \exp\Big(-i \sum_{n=1}^{N} \langle X_n, x_n^* \rangle\Big) \\
&= \prod_{n=1}^{N} \mathbb{E} \exp\big(-i \langle X_n, x_n^* \rangle\big) \\
&= \prod_{n=1}^{N} \widehat{\mathbb{P}}_{X_n}(x_n^*) = \widehat{\mathbb{P}_{X_1} \otimes \ldots \otimes \mathbb{P}_{X_N}}(x_1^*, \dots, x_N^*).
\end{aligned}
$$

Thus, Theorem A.14 implies $\mathbb{P}_{(X_1,\dots,X_N)} = \mathbb{P}_{X_1} \otimes \ldots \otimes \mathbb{P}_{X_N}$, which shows that $X_1, \dots, X_N$ are independent. ■

Finally, the next result shows that limits of Gaussian random variables are again Gaussian.

Proposition A.17. *If $(X_n)_{n=1}^{\infty}$ is a sequence of E-valued Gaussian random variables and X is a random variable with*

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

then X is Gaussian.

Proof. Let $x^* \in E^*$ be fixed. By passing to an appropriate subsequence, we have $\lim_{k\to\infty} \langle X_{n_k}, x^*\rangle = \langle X, x^*\rangle$ almost surely. Therefore, the dominated convergence theorem implies

$$\mathbb{E}\exp\big(-i\xi\,\langle X, x^*\rangle\big) = \lim_{k\to\infty} \mathbb{E}\exp\big(-i\xi\langle X_{n_k}, x^*\rangle\big) = \lim_{k\to\infty} \exp\big(-\tfrac{1}{2}\xi^2 q_{n_k}(x^*)\big),$$

using that X_{n_k} is Gaussian for each $k \in \mathbb{N}$. From this we infer that the limit $q(x^*) := \lim_{k\to\infty} q_{n_k}(x^*)$ exists (observe that $\big((q_{n_k}(x^*)\big)_{k=1}^{\infty}$ is a non-negative, bounded sequence). This leads to

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

and Theorem A.14 finally yield that $\langle X, x^*\rangle$ is Gaussian. ∎

A.3 Conditional Expectations and Martingales

Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space and $\mathscr{G}$ be a sub-σ-algebra of $\mathscr{A}$. For $1 \le p \le \infty$ we denote by $L^p(\Omega, \mathscr{G})$ the subspace of all $\xi \in L^p(\Omega)$ having a $\mathscr{G}$-measurable representative. Note that $L^p(\Omega, \mathscr{G})$ is a closed subspace of $L^p(\Omega)$. We next want to show that $L^p(\Omega, \mathscr{G})$ is the range of a contractive projection in $L^p(\Omega)$.

For $p = 2$ we have the orthogonal decomposition

$$L^2(\Omega) = L^2(\Omega, \mathscr{G}) \oplus L^2(\Omega, \mathscr{G})^{\perp},$$

and we can choose the orthogonal projection $P_{\mathscr{G}}$ onto $L^2(\Omega, \mathscr{G})$. As everyone does, we also 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 ξ with respect to $\mathscr{G}$. Note that $\mathbb{E}[\xi|\mathscr{G}]$ is an element of $L^2(\Omega, \mathscr{G})$ and therefore an equivalence class of random variables.

Lemma A.18. *For all $\xi \in L^2(\Omega)$ and $G \in \mathscr{G}$ we have*

$$\int_G \mathbb{E}[\xi|\mathscr{G}]\,\mathrm{d}\mathbb{P} = \int_G \xi\,\mathrm{d}\mathbb{P}.$$

Proof. Let $G \in \mathscr{G}$. Since $\mathbb{1}_G \in L^2(\Omega, \mathscr{G})$ and $\xi - \mathbb{E}[\xi|\mathscr{G}] \in L^2(\Omega, \mathscr{G})^{\perp}$, we have

$$\int_\Omega \mathbb{1}_G\big(\xi - \mathbb{E}[\xi|\mathscr{G}]\big)\,\mathrm{d}\mathbb{P} = 0,$$

which gives the desired identity. ∎

As a consequence, we get the following properties for $\xi \in L^2(\Omega)$:

(1) If $\xi \geq 0$ almost surely, then $\mathbb{E}[\xi|\mathscr{G}] \geq 0$ almost surely.

(2) By taking $G = \Omega$ in the previous lemma, we get

$$\mathbb{E}\big(\mathbb{E}[\xi|\mathscr{G}]\big) = \mathbb{E}\xi.$$

(3) Let ξ^+ and ξ^- be the positive and negative part of ξ, respectively. By (2) we then have

$$\mathbb{E}\big|\mathbb{E}[\xi|\mathscr{G}]\big| \leq \mathbb{E}\big(\mathbb{E}[\xi^+|\mathscr{G}] + \mathbb{E}[\xi^-|\mathscr{G}]\big) = \mathbb{E}\big(\mathbb{E}\big[|\xi|\big|\mathscr{G}\big]\big) = \mathbb{E}|\xi|,$$

which shows that the map $\xi \mapsto \mathbb{E}[\xi|\mathscr{G}]$ is L^1-bounded.

Since $L^2(\Omega)$ is dense in $L^1(\Omega)$, we can extend the conditional expectation operator to a contractive projection on $L^1(\Omega)$, which we also denote by $\mathbb{E}[\,\cdot\,|\mathscr{G}]$. The properties (1)–(3) then still hold for this projection.

Using this estimate and a conditional version of Jensen's inequality, we get the next theorem (cf. [3, Chapter V.1, Lemma 3] for more details).

Theorem A.19 (L^p-contractivity). *For all $1 \leq p \leq \infty$, 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 $G \in \mathscr{G}$ we have*

$$\int_G \mathbb{E}[\xi|\mathscr{G}]\,\mathrm{d}\mathbb{P} = \int_G \xi\,\mathrm{d}\mathbb{P}.$$

Our next aim is to extend these contractive operators from $L^p(\Omega)$ to $L^p(\Omega; E)$, where E is some Banach space. For this purpose, fix $1 \leq p < \infty$. Then, by the definition of the Lebesgue-Bochner spaces, the set

$$D := \Big\{X = \sum_{n=1}^{N} \xi_n x_n \colon\ \xi_n \in L^p(\Omega),\, x_n \in E,\, N \in \mathbb{N}\Big\}$$

is dense in $L^p(\Omega; E)$. Suppose next that $T \in \mathscr{B}\big(L^p(\Omega)\big)$. On the above set, we then define a linear operator $T \otimes I$ by

$$(T \otimes I)X = \sum_{n=1}^{N} T\xi_n \cdot x_n, \quad X \in D.$$

For positive operators T we then have the following result (cf. [5]).

Proposition A.20. *If T is a positive operator on $L^p(\Omega)$, then $T \otimes I$ extends uniquely to a bounded operator on $L^p(\Omega; E)$, and we have $\|T \otimes I\| = \|T\|$.*

Returning to the positive conditional expectation operator, we obtain the following extension of Theorem A.19.

Theorem A.21 (L^p-contractivity). *For $1 \le p \le \infty$, the linear 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*

$$\mathbb{E}[X|\mathscr{G}] := \big(\mathbb{E}[\,\cdot\,|\mathscr{G}] \otimes I\big)X$$

is the unique element of $L^p(\Omega, \mathscr{G}; E)$ with the property that for all $G \in \mathscr{G}$ we have

$$\int_G \mathbb{E}[X|\mathscr{G}]\,\mathrm{d}\mathbb{P} = \int_G X\,\mathrm{d}\mathbb{P}.$$

Having defined the conditional expectation for random variables with values in a Banach space, we next collect some of its properties.

(1) Let $X \in L^1(\Omega; E)$. Then

$$\mathbb{E}\big(\mathbb{E}[X|\mathscr{G}]\big) = \mathbb{E}X. \tag{A.1}$$

(2) If $X \in L^1(\Omega, \mathscr{G}; E)$, then we almost surely have

$$\mathbb{E}[X|\mathscr{G}] = X. \tag{A.2}$$

Especially, if $\xi \in L^p(\Omega)$ and $X \in L^{p'}(\Omega, \mathscr{G}; E)$, then almost surely

$$\mathbb{E}[\xi X|\mathscr{G}] = \mathbb{E}[\xi|\mathscr{G}]\,X, \tag{A.3}$$

and if $\xi \in L^p(\Omega, \mathscr{G})$ and $X \in L^{p'}(\Omega; E)$, then almost surely

$$\mathbb{E}[\xi X|\mathscr{G}] = \xi\,\mathbb{E}[X|\mathscr{G}]. \tag{A.4}$$

(3) If $X \in L^1(\Omega; E)$ is independent of $\mathscr{G}$ (that is, X is independent of $\mathbb{1}_G$ for all $G \in \mathscr{G}$), then almost surely

$$\mathbb{E}[X|\mathscr{G}] = \mathbb{E}X. \tag{A.5}$$

(4) If $X \in L^1(\Omega; E)$ and $\mathscr{H}$ is a sub-σ-algebra of $\mathscr{G}$, then almost surely

$$\mathbb{E}\big[\mathbb{E}[X|\mathscr{G}]\big|\mathscr{H}\big] = \mathbb{E}\big[\mathbb{E}[X|\mathscr{H}]\big|\mathscr{G}\big] = \mathbb{E}[X|\mathscr{H}]. \tag{A.6}$$

(5) Let $\mathscr{H}$ be a sub-σ-algebra of $\mathscr{G}$ and $X \in L^1(\Omega; E)$. Assume that $\mathscr{H}$ is independent of $\sigma(X, \mathscr{G}) := \sigma\big(\sigma(X) \cup \mathscr{G}\big)$ (i.e., the indicator functions $\mathbb{1}_H$ and $\mathbb{1}_G$ are independent for all $H \in \mathscr{H}$ and $G \in \sigma(X, \mathscr{G})$). Then we almost surely have

$$\mathbb{E}[X|\mathscr{G}, \mathscr{H}] := \mathbb{E}[X|\sigma(\mathscr{G} \cup \mathscr{H})] = \mathbb{E}[X|\mathscr{G}]. \tag{A.7}$$

(6) Let $(\widetilde{\Omega}, \widetilde{\mathscr{A}}, \widetilde{\mathbb{P}})$ be another probability space and $\widetilde{\mathscr{G}}$ be a sub-σ-algebra of $\widetilde{\mathscr{A}}$. Suppose that $\xi \in L^1(\Omega)$ and $X \in L^1(\widetilde{\Omega}; E)$, then almost surely

$$\mathbb{E}[\xi X|\mathscr{G} \otimes \widetilde{\mathscr{G}}] = \mathbb{E}[\xi|\mathscr{G}]\, \mathbb{E}[X|\widetilde{\mathscr{G}}]. \tag{A.8}$$

(7) Let X, $Y \in L^1(\Omega; E)$ be independent and identically distributed. Then we have

$$\mathbb{E}[X - Y|X + Y] := \mathbb{E}[X - Y|\sigma(X + Y)] = 0. \tag{A.9}$$

Proof. The assertions **(1)** - **(4)** follow directly from the uniqueness part of Theorem A.21.

(5) We first consider an $\mathbb{R}$-valued random variable $\xi \geq 0$. We further assume that $\mathbb{E}\xi > 0$ (since otherwise $\xi = 0$ almost surely and there is nothing to prove) and note that $\mathcal{C} := \{G \cap H : G \in \mathscr{G}, H \in \mathscr{H}\}$ is a generating system of $\sigma(\mathscr{G} \cup \mathscr{H})$, which is closed under taking finite intersections. By the independence of $\mathscr{H}$ and $\sigma(\xi, \mathscr{G})$, we have for $G \cap H \in \mathcal{C}$

$$\begin{aligned}\int_{G\cap H} \mathbb{E}[\xi|\mathscr{G}, \mathscr{H}]\, \mathrm{d}\mathbb{P} = \int_{G\cap H} \xi\, \mathrm{d}\mathbb{P} &= \mathbb{E}[\mathbb{1}_G \mathbb{1}_H \xi] \\ &= \mathbb{E}\mathbb{1}_H\, \mathbb{E}[\mathbb{1}_G \xi] = \mathbb{E}\mathbb{1}_H \mathbb{E}\big(\mathbb{E}[\mathbb{1}_G \xi|\mathscr{G}]\big) \\ &= \mathbb{E}\big(\mathbb{1}_{G\cap H}\mathbb{E}[\xi|\mathscr{G}]\big) = \int_{G\cap H} \mathbb{E}[\xi|\mathscr{G}]\, \mathrm{d}\mathbb{P}.\end{aligned}$$

By Dynkin's lemma, applied to the probability measures

$$\mu_1(C) := \frac{1}{\mathbb{E}\xi} \int_C \mathbb{E}[\xi|\mathscr{G}, \mathscr{H}]\, \mathrm{d}\mathbb{P} \quad \text{and} \quad \mu_2(C) := \frac{1}{\mathbb{E}\xi} \int_C \mathbb{E}[\xi|\mathscr{G}]\, \mathrm{d}\mathbb{P},$$

it follows that $\mu_1 = \mu_2$ on $\sigma(\mathcal{C}) = \sigma(\mathscr{G}, \mathscr{H})$, and this shows the desired estimate for positive random variables ξ. For arbitrary $\mathbb{R}$-valued random variables we consider positive and negative parts separately. Finally, the vector-valued case follows from the definition of the conditional expectation operator for 'simple' functions and approximation.

(6) By a similar argument as above, it suffices to prove the estimate for $\mathbb{R}$-valued random variables ρ. Therefore, we first consider ξ, $\rho \geq 0$, and we may assume that $\mathbb{E}\xi > 0$ and $\mathbb{E}\rho > 0$. Note that $\mathcal{D} := \{G_1 \times G_2 : G_1 \in \mathscr{G}, G_2 \in \widetilde{\mathscr{G}}\}$ is a generating system of $\mathscr{G} \otimes \widetilde{\mathscr{G}}$, which is closed under finite intersections. Then, by Fubini's theorem, we obtain for $G_1 \times G_2 \in \mathcal{D}$

$$\begin{aligned}\int_{G_1\times G_2} \mathbb{E}[\xi\rho|\mathscr{G} \otimes \widetilde{G}]\,\mathrm{d}(\mathbb{P}\otimes\widetilde{\mathbb{P}}) &= \int_{G_1\times G_2} \xi\rho\,\mathrm{d}(\mathbb{P}\otimes\widetilde{\mathbb{P}}) = \int_{G_1} \xi\,\mathrm{d}\mathbb{P} \int_{G_2} \rho\,\mathrm{d}\widetilde{\mathbb{P}} \\ &= \int_{G_1} \mathbb{E}[\xi|\mathscr{G}]\,\mathrm{d}\mathbb{P} \int_{G_2} \mathbb{E}[\rho|\widetilde{\mathscr{G}}]\,\mathrm{d}\widetilde{\mathbb{P}} \\ &= \int_{G_1\times G_2} \mathbb{E}[\xi|\mathscr{G}]\,\mathbb{E}[\rho|\widetilde{\mathscr{G}}]\,\mathrm{d}(\mathbb{P}\otimes\widetilde{\mathbb{P}}).\end{aligned}$$

As in **(5)**, we apply Dynkin's lemma with the probability measures

$$\begin{aligned}\nu_1(D) &:= \frac{1}{\mathbb{E}\xi\mathbb{E}\rho}\int_D \mathbb{E}[\xi\rho|\mathscr{G}\otimes\widetilde{\mathscr{G}}]\,\mathrm{d}(\mathbb{P}\otimes\widetilde{\mathbb{P}}) \quad \text{and} \\ \nu_2(D) &:= \frac{1}{\mathbb{E}\xi\mathbb{E}\rho}\int_D \mathbb{E}[\xi|\mathscr{G}]\mathbb{E}[\rho|\widetilde{\mathscr{G}}]\,\mathrm{d}(\mathbb{P}\otimes\widetilde{\mathbb{P}})\end{aligned}$$

to obtain $\nu_1 = \nu_2$ on $\mathscr{G} \otimes \widetilde{\mathscr{G}}$. This then shows the estimate for positive random variables ξ and ρ. By splitting the random variables in positive and negative parts and making a simple computation, we finally get

$$\mathbb{E}[\xi\rho|\mathscr{G}\otimes\widetilde{\mathscr{G}}] = \mathbb{E}[\xi|\mathscr{G}]\mathbb{E}[\rho|\widetilde{\mathscr{G}}]$$

for arbitrary $\mathbb{R}$-valued random variables ξ and ρ.

(7) Since X and Y are independent and identically distributed, we have

$$\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y = \mathbb{P}_Y \otimes \mathbb{P}_X = P_{(Y,X)}.$$

Let $A \in \sigma(X+Y)$ be arbitrary. Then there exists a set $B \in \mathscr{B}(E)$ satisfying $A = (X+Y)^{-1}(B)$, and this leads to

$$\begin{aligned}\int_A X\,\mathrm{d}\mathbb{P} &= \int_\Omega \mathbb{1}_A X\,\mathrm{d}\mathbb{P} = \int_E \mathbb{1}_B(x,y)x\,\mathrm{d}\mathbb{P}_{(X,Y)} \\ &= \int_E \mathbb{1}_B(y,x)y\,\mathrm{d}\mathbb{P}_{(Y,X)} = \int_\Omega \mathbb{1}_A Y\,\mathrm{d}\mathbb{P} = \int_A Y\,\mathrm{d}\mathbb{P}.\end{aligned}$$

From this we infer that

$$\mathbb{E}[X|X+Y] = \mathbb{E}[Y|X+Y] \quad \text{almost surely,}$$

and the linearity of the conditional expectation operator finally shows the desired estimate. ∎

Now that the existence and properties of conditional expectations were discussed, we next introduce E-valued *martingales*.

Let I be a partially ordered set. A *filtration* with index set I is a family $(\mathscr{F}_i)_{i\in I}$ of sub-σ-algebras of $\mathscr{A}$ such that $\mathscr{F}_i \subseteq \mathscr{F}_j$ whenever $i \leq j$. A family of E-valued random variables is said to be *adapted* to the filtration $(\mathscr{F}_i)_{i\in I}$ if each X_i is strongly $\mathscr{F}_i$-measurable (or more precisely, if each X_i has a strongly $\mathscr{F}_i$-measurable representative).

Definition A.22. *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 to $(\mathscr{F}_i)_{i\in I}$ and if for all $i \leq j$ we have*

$$\mathbb{E}[M_j|\mathscr{F}_i] = M_i \quad \text{almost surely.}$$

If in addition $\mathbb{E}\|M_i\|_E^p < \infty$ for all $i \in I$ and some $1 \leq p < \infty$, then we call $(M_i)_{i\in I}$ an E-valued L^p martingale.

Example A.23 (Martingale transform). A sequence of $\mathbb{R}$-valued random variables $v = (v_n)_{n=1}^N$ 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 $n = 1, \ldots, N$ (with the understanding that $\mathscr{F}_0 = \{\emptyset, \Omega\}$). If $M = (M_n)_{n=1}^N$ is an E-valued martingale with respect to $(\mathscr{F}_n)_{n=1}^N$, then the sequence $v * M = \big((v * M)_n\big)_{n=1}^N$ defined by

$$(v * M)_n := \sum_{j=1}^{n} v_j(M_j - M_{j-1}), \quad n = 1, \ldots, N,$$

is called the *martingale transform* of M by v (with the understanding that $M_0 = 0$).

If we further assume that each v_n is bounded, then $v * M$ is indeed a martingale with respect to $(\mathscr{F}_n)_{n=1}^N$. In fact, by the boundedness the random variables $v_j(M_j - M_{j-1})$ are integrable and so also $(v * M)_n$. Clearly, $v * M$ is adapted to $(\mathscr{F}_n)_{n=1}^N$. Moreover, by (A.4) and the $\mathscr{F}_{n-1}$-measurability of $(v * M)_{n-1}$ and v_n, we obtain

$$\mathbb{E}\big[(v * M)_n \big| \mathscr{F}_{n-1}\big] = (v * M)_{n-1} + v_n \mathbb{E}[M_n - M_{n-1}|\mathscr{F}_{n-1}] = (v * M)_{n-1}. \quad \blacksquare$$

Finally, we collect some important results of E-valued L^p martingales, starting with the famous martingale inequality by Doob (cf. [8, Proposition 4.1.1]).

Theorem A.24 (Doob). *Let $1 < p < \infty$, and let $(M_n)_{n=1}^N$ be an E-valued martingale with respect to $(\mathscr{F}_n)_{n=1}^N$. Define*

$$M^*\colon \Omega \to [0,\infty), \quad M^*(\omega) := \max_{n=1}^{N} \|M_n(\omega)\|_E.$$

If $M_N \in L^p(\Omega; E)$, then

$$\mathbb{E}|M^*|^p \le \Big(\frac{p}{p-1}\Big)^p \mathbb{E}\|M_N\|_E^p.$$

Given a filtration $(\mathscr{F}_n)_{n=1}^\infty$ on $(\Omega, \mathscr{A}, \mathbb{P})$, we denote by $\mathscr{F}_\infty$ the σ-algebra generated by $(\mathscr{F}_n)_{n=1}^\infty$, that is, $\mathscr{F}_\infty$ is the smallest σ-algebra containing each sub-σ-algebra $\mathscr{F}_n$. The next result can be found in [3, Chapter V.2, Theorem 1 and 8].

Theorem A.25 (Martingale convergence theorem). *Let $1 \le p < \infty$, 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.

Fix $1 \le r < \infty$, and assume that (U, Σ, μ) is a σ-finite measure space and Σ is countably generated. In this case we have the following result.

Lemma A.26. *Let $(M_n)_{n=1}^N$ be an $L^r(U)$-valued L^r martingale with respect to a filtration $(\mathscr{F}_n)_{n=1}^N$. Then $\big(M_n(u)\big)_{n=1}^N$ is an $\mathbb{R}$-valued L^r martingale with respect to $(\mathscr{F}_n)_{n=1}^N$ for μ-almost all $u \in U$.*

Proof. Since (U, Σ, μ) is a σ-finite measure space, U is a countable union of disjoint sets $U_k \in \Sigma$, $k \in \mathbb{N}$, of finite measure. Take $U^{(K)} := \bigcup_{k=1}^K U_k$ and let $(U_j^K)_{j=1}^{N_K}$ be a partition of $U^{(K)}$ in such a way that $\Sigma_K := \sigma\big(U_j^{(K)}, j = 1, \dots, N_K\big) \subseteq \Sigma_M$ whenever $K \le M$, and $\Sigma = \sigma(\Sigma_K, K \in \mathbb{N})$. Next, we define for $K \in \mathbb{N}$

$$P_K := \mathbb{E}[\,\cdot\,|\Sigma_K]\colon L^r(U) \to L^r(U).$$

Let $f \in L^r(U)$. Since Σ_K is generated by N_K disjoint sets, P_K has the following representation

$$(P_K f)(u) = \sum_{j=1}^{N_K} \frac{\int_{U_j^{(K)}} f \,\mathrm{d}\mu}{\mu(U_j^{(K)})} \mathbb{1}_{U_j^{(K)}}(u).$$

Let $\varepsilon > 0$. Since $f \in L^r(U)$ and $\lim_{n\to\infty} \mathbb{1}_{U^{(n)}}(u)f(u) = f(u)$ for all $u \in U$, the dominated convergence theorem yield an $N \in \mathbb{N}$ such that

$$\|\mathbb{1}_{U^{(K)}} f - f\|_r < \frac{\varepsilon}{6} \quad \text{for all } K \geq N.$$

Moreover, for that N the martingale convergence theorem gives an $M(N) \in \mathbb{N}$ such that

$$\left\|\mathbb{E}[\mathbb{1}_{U^{(N)}} f|\Sigma_K] - \mathbb{1}_{U^{(N)}} f\right\|_r < \frac{\varepsilon}{6} \quad \text{for all } K \geq M(N),$$

since $\mu(U^{(N)}) < \infty$. Thus for $K \geq \max\{N, M(N)\}$, we obtain

$$\begin{aligned}
\left\|\mathbb{E}[f|\Sigma_K] - f\right\|_r &\leq \left\|\mathbb{E}[\mathbb{1}_{U^{(K)}} f|\Sigma_K] - \mathbb{E}[\mathbb{1}_{U^{(N)}} f|\Sigma_K]\right\|_r + \left\|\mathbb{E}[\mathbb{1}_{U^{(N)}} f|\Sigma_K] - \mathbb{1}_{U^{(N)}} f\right\|_r \\
&\quad + \left\|\mathbb{1}_{U^{(N)}} f - \mathbb{1}_{U^{(K)}} f\right\|_r + \left\|\mathbb{1}_{U^{(K)}} f - f\right\|_r \\
&\leq 2\left\|\mathbb{1}_{U^{(N)}} f - \mathbb{1}_{U^{(K)}} f\right\|_r + \frac{\varepsilon}{3} \\
&\leq 2\left\|\mathbb{1}_{U^{(N)}} f - f\right\|_r + 2\left\|\mathbb{1}_{U^{(K)}} f - f\right\|_r + \frac{\varepsilon}{3} \\
&\leq \varepsilon,
\end{aligned}$$

where we used that $\mathbb{E}[f|\Sigma_K] = \mathbb{E}[\mathbb{1}_{U^{(K)}} f|\Sigma_K]$ by construction, and that the conditional expectation operator is contractive (observe that $\mu(U^{(K)}) < \infty$).

For each $n = 1, \ldots, N$ we next define the random variable $M_n^{(k)} \colon \Omega \to L^r(U)$ by

$$M_n^{(k)}(\omega) := P_k M_n(\omega) \quad \omega \in \Omega,\ k \in \mathbb{N}.$$

Let $A \in \mathscr{F}_{n-1}$ be arbitrary. Since $(M_n)_{n=1}^N$ is a martingale with respect to $(\mathscr{F}_n)_{n=1}^N$, Theorem A.21 leads to

$$\int_A M_n \,\mathrm{d}\mathbb{P} = \int_A \mathbb{E}[M_n|\mathscr{F}_{n-1}] \,\mathrm{d}\mathbb{P} = \int_A M_{n-1} \,\mathrm{d}\mathbb{P}.$$

Therefore, by Fubini's theorem, we obtain for each fixed $k \in \mathbb{N}$

$$\begin{aligned}
\int_A M_n^{(k)} \,\mathrm{d}\mathbb{P} &= \sum_{j=1}^{N_k} \frac{1}{\mu(U_j^{(k)})} \mathbb{1}_{U_j^{(k)}} \int_A \int_{U_j^{(k)}} M_n \,\mathrm{d}\mu \,\mathrm{d}\mathbb{P} \\
&= \sum_{j=1}^{N_k} \frac{1}{\mu(U_j^{(k)})} \mathbb{1}_{U_j^{(k)}} \int_{U_j^{(k)}} \int_A M_{n-1} \,\mathrm{d}\mathbb{P} \,\mathrm{d}\mu \\
&= \int_A \sum_{j=1}^{N_k} \frac{1}{\mu(U_j^{(k)})} \mathbb{1}_{U_j^{(k)}} \int_{U_j^{(k)}} M_{n-1} \,\mathrm{d}\mu \,\mathrm{d}\mathbb{P} \\
&= \int_A M_{n-1}^{(k)} \,\mathrm{d}\mathbb{P}.
\end{aligned}$$

And again by Theorem A.21, this implies

$$\mathbb{E}[M_n^{(k)}|\mathscr{F}_{n-1}] = M_{n-1}^{(k)} \quad \text{almost surely.}$$

Next, for any $k \in \mathbb{N}$ and each fixed $u \in U$, almost surely we have

$$\begin{aligned}\mathbb{E}[M_n^{(k)}(u)|\mathscr{F}_{n-1}] &= \sum_{j=1}^{N_k} \frac{1}{\mu(U_j^{(k)})}\mathbb{E}\left[\int_{U_j^{(k)}} M_n \,\mathrm{d}\mu \,\middle|\, \mathscr{F}_{n-1}\right] \mathbb{1}_{U_j^{(k)}}(u) \\ &= \mathbb{E}[M_n^{(k)}|\mathscr{F}_{n-1}](u) \qquad (\triangle) \\ &= M_{n-1}^{(k)}(u).\end{aligned}$$

Having shown all these auxiliary results, we are now in the position to prove the assertion claimed in the lemma. Fix an $n \in \{1, \ldots, N\}$. Since $\lim_{k\to\infty} M_n^{(k)} = M_n$ and $\lim_{k\to\infty} M_{n-1}^{(k)} = M_{n-1}$ almost surely in $L^r(U)$, and since M_n, $M_{n-1} \in L^r(\Omega; L^r(U))$, the dominated convergence theorem and Fubini's theorem yield

$$\lim_{k\to\infty} M_n^{(k)} = M_n \;\text{ and }\; \lim_{k\to\infty} M_{n-1}^{(k)} = M_{n-1} \;\text{ in } L^r(\Omega; L^r(U)) \simeq L^r(U; L^r(\Omega)).$$

Thus, we may find an appropriate subsequence $(k_j)_{j\in\mathbb{N}}$ and a μ-nullset $U_0 \subseteq U$ such that

$$\lim_{j\to\infty} M_n^{(k_j)}(u) = M_n(u) \;\text{ and }\; \lim_{j\to\infty} M_{n-1}^{(k_j)}(u) = M_{n-1}(u) \;\text{ in } L^r(\Omega) \text{ for all } u \notin U_0.$$

Especially, $M_n(u) \in L^r(\Omega)$ for all $u \notin U_0$ and is therefore integrable. We next observe that each $M_n^{(k)}(u)$ is $\mathscr{F}_n$-measurable by construction, which implies that $M_n(u)$ is $\mathscr{F}_n$-measurable as a limit of $\mathscr{F}_n$-measurable random variables for all $u \notin U_0$. Finally, by ($\triangle$), Theorem A.19, and the above convergence, we almost surely get for all $u \notin U_0$

$$\mathbb{E}[M_n(u)|\mathscr{F}_{n-1}] = \lim_{j\to\infty} \mathbb{E}[M_n^{(k_j)}(u)|\mathscr{F}_{n-1}] = \lim_{j\to\infty} M_{n-1}^{(k_j)}(u) = M_{n-1}(u),$$

which concludes the proof. ■

Bibliography

[1] L. Arnold. *Stochastische Differentialgleichungen – Theorie und Anwendung [Stochastic Differential Equations – Theory and Application].* R. Oldenbourg-Verlag, München/Wien, 1995.

[2] Z. Brzeźniak, J.M.A.M. van Neerven, M.C. Veraar and L. Weis. Itô's formula in UMD Banach spaces and regularity of solutions of the Zakai equation. *J. Differ. Equations*, 245:30-58, 2008.

[3] J. Diestel and J.J. Uhl, Jr. *Vector Measures.* American Mathematical Society, Providence, Rhode Island, 1977.

[4] K.-J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations.* Graduate Texts in Mathematics, vol. 194, Springer-Verlag, New York, 2000.

[5] M. Haase. Convexity inequalities for positive operators. *Positivity*, 11(1):57-63, 2006.

[6] K. Itô. Stochastic integral. *Proc. Imp. Acad. Tokyo*, 20:519-524, 1944.

[7] O. Kallenberg. *Foundations of Modern Probability,* second edition. Probability and its Applications, Springer-Verlag, New York, 2002.

[8] S. Kwapień and W.A. Woyczyński. *Random Series and Stochastic Integrals: Single and Multiple.* Probability and its Applications, Birkhäuser Boston Inc., Boston, MA, 1992.

[9] D. Meintrup and S. Schäffler. *Stochastik – Theorie und Anwendungen [Stochastic Calculus – Theory and Applications]* Springer-Verlag, Berlin/Heidelberg, 2005.

[10] J.M.A.M. van Neerven, M.C. Veraar and L. Weis. Stochastic integration in UMD Banach spaces. *Ann. Probab.*, 35:1438-1478, 2007.

[11] J.M.A.M. van Neerven and L. Weis. Stochastic integration of functions with values in a Banach space. *Studia Math.*, 166(2):131-170, 2005.

[12] K.R. Parthasarathy. *Probability Measures on Metric Spaces.* Academic Press, New York/London, 1967.

[13] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983.

[14] J.M. Steele. *Stochastic Calculus and Financial Applications.* Applications of Mathematics, vol. 45, Springer-Verlag, New York, 2001.

[15] M.C. Veraar. *Stochastic Integration in Banach Spaces and Applications to Parabolic Evolution Equations.* Ph.D. thesis, Delft University of Technology, 2006.

[16] N. Wiener. Differential space. *J. Math. Phys.* 2:131-174, 1923.

Erklärung

Hiermit versichere ich, dass ich die Arbeit selbständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung des Karlsruher Instituts für Technologie zur Sicherung guter wissenschaftlicher Praxis in der gültigen Fassung beachtet habe.

Karlsruhe, den 3. August 2012

Markus Antoni