TECHNISCHE UNIVERSITÄT WIEN

DIPLOMA THESIS

Aspects of Stochastic Integration beyond Standard Assumptions

supervised by

Univ-Prof. Dr. Josef Teichmann,

written by

Lorenz Riess, BSc
Student number: 01612419

at

TU Wien.

Vienna, February 22, 2021

(Signature Author)

(Signature Supervisor)

Kurzfassung

Diese Arbeit beschäftigt sich mit der Theorie stochastischer Integration und versucht einige Resultate über übliche Annahmen hinaus zu verallgemeinern.

Ein entscheidender Schritt der Definition eines stochastischen Integrals ist es, dieses für Martingale zu definieren. Folgt man dem funktional-analytischen Zugang von Philip E. Protter, dann ist eine Ungleichung von Burkholder die entscheidende Zutat. Diese Arbeit verallgemeinert diese Ungleichung für stochastische Prozesse mit Werten in bestimmten Banachräumen, speziell wenn der Integrand Werte in solchen Banachräumen annimmt und der Integrator reellwertig ist. Für einen solchen Banachraum können wir ein stochastisches Integral für alle càglàd Prozesse mit Werten in diesem Banachraum definieren, wenn als Integrator ein reellwertiges Semimartingal verwendet wird.

Ein weiterer großer Teil dieser Arbeit fokussiert sich auf das Bichteler-Dellacherie Theorem, das eine Charakterisierung jener stochastischen Prozesse im reellen Fall liefert, die als Integrator verwendet werden können. Es besagt, dass ein stochastisches Integral genau für càdlàg Semimartingale definiert werden kann und, dass sich diese als Summe eines lokalen Martingals und eines Prozesses mit endlicher Variation schreiben lassen. Dieses Resultat verallgemeinern wir, indem wir die càdlàg Annahme der Pfade fallen lassen, die analogen Voraussetzungen verwenden und den Prozess auf dem Level von Versionen noch immer als Summe eines lokalen Martingals und eines Prozesses mit endlicher Variation schreiben können. Genauer finden wir ein lokales Martingal mit càdlàg Pfaden und einen Prozess von endlicher Variation, sodass für jeden Zeitpunkt der ursprüngliche Prozess fast sicher als Summe der beiden Prozesse geschrieben werden kann.

Die gleichen Ideen können auf die Doob-Meyer Zerlegung angewandt werden, wodurch wir diese ebenso auf Supermartingale ohne stetige Pfade erweitern können.

Abstract

This thesis deals with the theory of stochastic integration and tries to generalize some results beyond standard assumptions.

One crucial part of defining a stochastic integral is the step to define it for martingales. If one follows a functional analytic approach introduced by Philip E. Protter, an inequality due to Burkholder is the necessary ingredient. This thesis generalizes this inequality to a setting in which stochastic processes with values in certain Banach spaces are considered, in particular when the integrand is Banach space-valued and the integrator real-valued. For such a Banach space we can define a stochastic integral of all càglàd processes with values in that Banach space against a real-valued semimartingale.

Another major part of this thesis focuses on the Bichteler-Dellacherie theorem which is the characterization of stochastic processes which can be taken as integrator in the real-valued case. It tells that a stochastic integral can be defined precisely for càdlàg semimartingales and that those decompose into a local martingale and a finite variation process. We can extend this result by dropping the càdlàg assumption on the paths, using the analogue assumptions of the theorem and still decomposing the process on a level of versions. In particular we can find a càdlàg local martingale and a finite variation process such that for each time point the original process is the sum of those two, almost surely.

The same ideas can be applied to the Doob-Meyer decomposition and we can again generalize this to supermartingales without continuity assumptions on its paths.

Acknowledgement

To start with I sincerely thank my supervisor Josef Teichmann from ETH Zurich for his mentoring, all the insightful discussions and for trusting in me to write this thesis about a very challenging but highly interesting topic. Although this thesis arose during times of physical distance, I want to thank him for taking lots of time for online meetings, his flexibility and theoretical input he gave me. This joint work also taught me lots about the methods and approaches of research, and how to tackle new problems.

Furthermore, I want to thank my parents and siblings for the strong support during my studies and making them possible.

Finally, I want to thank my friends who have always been there for me, especially in tougher times.

Lorenz

Statement of Originality

I hereby declare that I have authored the present master thesis independently and did not use any sources other than specified. I have not yet submitted the work to any other examining authority in the same comparable form. It has not been published yet.

Vienna, July 26, 2021

Lorenz Riess

Contents

1 Introduction

Stochastic integration theory deals with stochastic processes and defining a meaningful integral for such. The goal is to define an integral for a large class of stochastic processes and deduce properties of this integral and the emerging integral process. This theory started in the early 1940's in which Kiyosi Itô developed an integral in which the integrator was a Brownian motion, see also [It44] for one of his first papers. In the following this was extended to square integrable martingales, then martingales and afterwards semimartingales in general.

The stochastic integral has a variety of applications. An example is financial mathematics in which one considers a financial market as integrator and a strategy as integrand. The stochastic integral at some time point then describes the wealth of an agent investing in this strategy.

Before starting with explaining the focus of this thesis we want to mention that we assume the reader to be familiar with basic probability theory and basic theory of stochastic processes. Concerning these topics we can recommend the books by Olav Kallenberg, [Kal02], and also for some more measure theoretic facts the book [Els18] by Jürgen Elstrodt.

This thesis is devoted to inspecting some aspects of stochastic integration theory and pushing those a little further. We focus on the functional analytic way of defining the stochastic integral by Philip Protter (see [Pro04]) which will be introduced in Chapter 2. It will be done by saying a stochastic process X can be taken as integrator if it meets a certain condition and is then referred to as "good integrator". The condition states that a process is a good integrator if the map of simple predictable processes to their discrete stochastic integral against X is continuous when considering the so-called ucp metric which will be introduced later on. After defining this condition in more detail one question arises: Which stochastic processes are these "good integrators"? This question is answered by looking at it from two directions:

i) Show that all semimartingales whose paths are right-continuous and have left limits satisfy this definition.

ii) Prove that if a stochastic process is a good integrator, i.e. it satisfies the definition, it has to be a semimartingale.

Recall that a semimartingale is a stochastic process X which can be written as

$X = M + A$ where M is a local martingale and A a finite variation process. This thesis focuses on generalizing these two points in slightly different directions.

The crucial part in proving i) is to show that all martingales satisfy the definition of being a good integrator. For this sake there is a very elegant inequality due to Burkholder which states that the probability of the discrete stochastic integral exceeding some constant can be estimated by some constant and the martingale at end time. One part of this work is devoted to generalizing this inequality to a Banach space setting. In particular when the integrand is Banach space-valued. For this sake we introduce a Burkholder property for Banach spaces in Chapter 3. There, the inequality is stated in the following way which is one result of this thesis.

Theorem 1.1 (see also Theorem 3.9). *Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space with right-continuous filtration and $(E, \|\cdot\|)$ a Banach space with the Burkholder property. Additionally, let X be a real-valued martingale. We then have for all Banach space-valued simple predictable processes H which are uniformly bounded, i.e. there are stopping times $0 \leq \tau_0 \leq \ldots \leq \tau_n$ and Banach space-valued random variables $H_0, \ldots, H_{n-1}$ which are uniformly bounded and such that H_i is measurable w.r.t. F_{τ_i} and $H = \sum_{i=0}^{n-1} H_i \mathbb{1}_{(\tau_i, \tau_{i+1}]}$, time points $T > 0$ and $c > 0$ that it holds*

$$c\mathbb{P}[\sup_{0\leq t\leq T} \|\sum_{i=0}^{n-1} H_i(X_{\tau_i\wedge t} - X_{\tau_{i+1}\wedge t})\| > c] \leq (10 + 8C_E^2) \max_{0\leq i\leq n-1} \|H_i\|_\infty \mathbb{E}[|X_t|].$$

Here, C_E is a constant which only depends on the Banach space E.

This theorem is a generalization of the corresponding inequality in the real-valued case due to Burkholder, see also [Mey72, Theorem 47] for example.

Chapter 3 is mostly devoted to Banach spaces in which that inequality holds. Having this inequality a stochastic integral for Banach space-valued integrand against real-valued semimartingales can be defined similar as Protter's functional analytic approach in [Pro04]. We formulate this in the following theorem.

Theorem 1.2 (see also Theorem 3.19). *Let $(E, \|\cdot\|)$ be a Banach space satisfying the Burkholder property. Then a stochastic integral for all $E-$valued stochastic processes which are left-continuous and have right limits against a real-valued semimartingale can be defined.*

This gives new insights to stochastic integration theory in a Banach space setting and can be applied to martingale type 2 spaces for example (see also Chapter 8.6 for a definition).

The very short Chapter 4 considers the Burkholder inequality in a setting in which the integrand is real-valued and the integrator Banach space-valued.

The second major part of this thesis is devoted to point ii), which is generalizing the Bichteler-Dellacherie theorem which tells that if a right-continuous stochastic process, which has left limits as well, is a good integrator, then it already has to be a semimartingale. The Bichteler-Dellacherie theorem is celebrated in stochastic integration theory since it characterizes the processes which can be taken as integrators.

In Chapter 7 we consider a stochastic process without continuity assumption on its paths and prove that it can still be decomposed into a local martingale and a finite variation process if an analogue to the good integrator property holds. We are able to prove the following theorem on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with right-continuous filtration.

Theorem 1.3 (see also Theorem 7.6)**.** *Let X be a stochastic process such that for each $t \geq 0$ the sets*

$$\{\sum_{i=0}^{n-1} H_i(X_{t_{i+1}\wedge t} - X_{t_i\wedge t}) \mid H_i \; uniformly \; bounded \; by \; 1 \; and \; \mathcal{F}_{t_i} - measurable\}$$

are bounded in probability. Then there exists a local martingale M which is right-continuous and has left limits and a process of finite variation A such that for all $t \geq 0$ we have

$$X_t = M_t + A_t, \; almost \; surely.$$

The assumption of the theorem is the analogue of the condition for being a good integrator. In particular, if a process satisfies the condition of the theorem and, in addition, has paths which are right-continuous and have left limits, then it is a good integrator. Therefore, the theorem is an extension and generalization of the celebrated Bichteler-Dellacherie theorem (see for example [Bic79, Theorem 1].

Similar techniques as those which will be used in proving the generalized Bichteler-Dellacherie theorem lead to a generalization of the Doob-Meyer decomposition. The Doob-Meyer decomposition tells that a supermartingale can be written as a local martingale minus a non-decreasing process. Again, we generalize this to non-negative supermartingales without continuity assumptions on its paths. This is Theorem 7.8 of this thesis.

Chapters 5 and 6 are used to present two major results of stochastic analysis which are crucial in our proof of the generalized Bichteler-Dellacherie theorem. These are the Nikisin-Yan and the Girsanov-Meyer theorem. The Nikisin-Yan theorem can be applied to a convex set of random variables which is bounded from above in probability and allows one to change the measure to an equivalent one under which the expectation of all random variables in this set is bounded from above uniformly. The Girsanov-Meyer theorem is another theorem about equivalent measure changes, while keeping semimartingale properties.

The last Chapter is a rather extensive appendix in which facts concerning probability theory and theory of stochastic processes in Banach spaces are given, which are needed for Chapters 3 and 4. Next to this, a guided tour through already existing theory of stochastic integration in Banach spaces is given which makes use of so called UMD (unconditional martingale difference) Banach spaces. There, we review the literature and gained some insights and ideas for this thesis as well.

2 Stochastic Integration via the Good Integrator Property

In this chapter an introduction into stochastic integration of real-valued stochastic processes is given. The approach of Philip E. Protter in [Pro04] is used where he introduces the so called good integrator property. We will not prove every result but rather introduce this approach since the method will be important for the next chapters. The interested reader may find rigorous proofs in [Pro04]. Some results stated in this chapter will be generalized to a more general setting in Chapter 3. This means that actually some of the theorems presented here will follow from those slightly more general theorems.

Protter's way of introducing stochastic integrals is a functional analytic approach, which enables us to define a stochastic integral quite quickly and general when talking about càdlàg processes and their martingale transforms. càdlàg here comes from the French expression "continue à droite, limite à gauche", which means that we consider stochastic processes whose paths are right continuous and have left limits. If we talk about càglàd processes we talk about processes whose paths are left continuous and have right limits. The martingale transform refers to the discrete stochastic integral. Actually Protter's functional analytic approach is so general and in some sense complete that any càdlàg semimartingale can be taken as an integrator for the stochastic integral. In fact these are the only càdlàg processes which can be taken as such. This is the statement of the famous Bichteler-Dellacherie theorem.

As a personal comment I was taught this introduction into stochastic integration during my exchange year at ETH Zurich by Josef Teichmann in the course Mathematical Finance.

Let us fix some basic definitions and notations in the following which will be used throughout this chapter.

We fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and assume the filtration to be right-continuous. For two càdlàg or two càglàd adapted processes X and Y, we define the so called ucp metric (uniform convergence on compacts in probability) via

$$d(X,Y) := \sum_{n\geq 1} \frac{1}{2^n} \mathbb{E}\left[|X-Y|_n^* \wedge 1\right],$$

where we use the * notation for the running maximum process (also more precisely running supremum process), i.e.

$$|X|_t^* := \sup_{0 \le s \le t} |X_s|.$$

The supremum over all time points will be denoted by

$$|X|^* := \sup_{t \ge 0} |X_t|.$$

As the name suggests, the convergence of this metric is equivalent to convergence in probability uniformly on compacts, that is for a sequence of càdlàg (or càglàd) adapted stochastic processes $(X_n)_{n \ge 1}$ and a càdlàg (or càglàd) adapted stochastic process X we have

$$d(X_n, X) \to 0 \iff \forall t \ge 0,\ \forall \epsilon > 0 :\ \mathbb{P}[|X_n - X|_t^* > \epsilon] \to 0.$$

Having the ucp metric at hand we define the two spaces

$$\mathcal{D} := \{X \mid X \text{ càdlàg and adapted}\},$$
$$\mathcal{L} := \{Y \mid Y \text{ càglàd and adapted}\}.$$

Equipped with the ucp metric these are complete topological vector spaces. Note that in the following if we talk about stochastic processes we will also assume them to be adapted, also if we do not mention it explicitly. Sometimes we even only say processes in this case.

The idea of our approach of defining a stochastic integral will be to introduce a stochastical integral as a linear operator on a dense subset of $\mathcal{L}$ to $\mathcal{D}$. This map can then be extended onto the closure of this dense subset, which will be $\mathcal{L}$ if the operator meets certain continuity conditions. If one is already familiar with some integration theory or stochastic integration, one would not be surprised by the choice of this dense subset. We will introduce it in the following lines.

For this sake, given stopping times $0 = \tau_0 \le \tau_1 \le \ldots \le \tau_n < \infty$ and random variables $H_0, \ldots, H_{n-1}$ such that H_i is measurable with respect to $\mathcal{F}_{\tau_i}$, we define a simple predictable process as

$$H := H_0 \mathbb{1}_{\{0\}} + \sum_{i=0}^{n-1} H_i \mathbb{1}_{(\tau_i, \tau_{i+1}]}.$$

We will call a simple predictable process also a simple predictable strategy in the following.

The set of all such simple predictable strategies, which will be our choice of the upwards mentioned dense subset, will be denoted by $\mathcal{S}$. We furthermore define

$$\mathcal{S}_u := \{H = H_0 \mathbb{1}_{\{0\}} + \sum_{i=0}^{n-1} H_i \mathbb{1}_{(\tau_i, \tau_{i+1}]} \mid H_i \in \mathbb{L}^\infty,\ i = 0, \dots n-1\},$$

which is the set of all simple predictable strategies which are uniformly bounded. For a process in this set we set

$$\|H\|_\infty := \sup_{t \geq 0} \|H_t\|_{\mathbb{L}^\infty}$$

to have a slightly shorter notation.

Now, we are already at the point to define the good integrator property which tells us if a process can be taken as stochastic integrator.

Definition 2.1 (Good Integrator)**.** *A càdlàg and adapted stochastic process X is called good integrator, also said to have the good integrator property, if the map*

$$J_X : (\mathcal{S}, d) \to (\mathcal{D}, d) : H = H_0 \mathbb{1}_{\{0\}} + \sum_{i=0}^{n-1} H_i \mathbb{1}_{(\tau_i, \tau_{i+1}]} \mapsto (H \cdot X) \tag{2.1}$$

is continuous. Here we set

$$(H \cdot X)_t := H_0 X_0 + \sum_{i=0}^{n-1} H_i (X_{\tau_{i+1} \wedge t} - X_{\tau_i \wedge t})$$

for $t \geq 0$.

Remark 2.2. *As a short remark on the definition, one can see right away that the map defined in* (2.1) *is linear. Therefore, it is sufficient to only check continuity at* 0 *since we deal with topological vector spaces, see also [BKW17, Proposition 2.1.11].*

Also, if a process X justifies the good integrator property, by the linearity of the map J_X it is also uniformly continuous. Therefore, the map J_X can be extended to the closure of $\mathcal{S}$ with respect to the ucp metric which is precisely the space $\mathcal{L}$, the space of all càglàd processes. That the closure of $\mathcal{S}$ in the ucp topology is $\mathcal{L}$ can be seen in [Pro04, Chapter 2, Theorem 10].

In the second part of the remark we use the fact that a uniformly continuous map $f : (Y, d_1) \to (Z, d_2)$, where (Z, d_2) is a complete metric space and (Y, d_1) a metric space, can be uniquely extended to the closure of Y with respect to the topology induced by the metric d_1. This extension agrees with the original map f on Y and is uniformly continuous as well. See also [BKW17, Satz 1.1.1] for a proof of this statement.

Having defined the good integrator property we can quickly define quadratic covariation as well. Taking a good integrator X, the process X_- denotes the càglàd version of X and is an element of $\mathcal{L}$, which means we can look at the stochastic integral $(X_- \cdot X)$ or also $(Y_- \cdot X)$ for another good integrator Y.

The quadratic covariation between two good integrators X and Y is then defined as

$$[X, Y] := XY - (X_- \cdot Y) - (Y_- \cdot X).$$

Analogously, the quadratic variation of X is defined as

$$[X, X] := X^2 - 2(X_- \cdot X).$$

Different to some usual approaches in which one first treats stochastic integration with respect to martingales only, mostly even with continuous paths, and then generalizes this to local martingales and then semimartingales, we use a completely functional analytic definition. Intuitively spoken the definition says, that X is a good integrator if the discrete stochastic integral does not change a lot if also the integrand does not change by a lot. It will turn out that with this definition which sounds very intuitive, we can define a stochastic integral right away as a generalization of the discrete stochastic integral just by extending a continuous linear map. Also the definition of quadratic variation and quadratic covariation can be done very elegantly.

As drawback by this general definition we do not know the processes which fit this definition immediately. We will however show that all semimartingales fulfill this definition and the Bichteller-Delacherie theorem will do the rest in characterizing the good integrator property as the property fulfilled only by semimartingales.

For our definition of the good integrator property, there also is an equivalent statement which characterizes this property. We state this in the following theorem, see also Theorem 3.16 for an extended and more general version.

Theorem 2.3. *Let X be a càdlàg and adapted stochastic process. X is a good integrator if and only if for all $t \geq 0$, the map*

$$I_{X^t} : (\mathcal{S}_u, \|\cdot\|_\infty) \to (\mathbb{L}^0, d_{\mathbb{P}}) : H \mapsto (H \cdot X)_t$$

is continuous, where $d_{\mathbb{P}}$ denotes the metric induced by convergence in probability.

This characterization of the good integrator property reduces the domain to simple predictable processes which are uniformly bounded. We also do not have to use the ucp metric since here we have to check continuity for maps with co-domain the set of random variables. However, for every time point we have a different map. Usually it is easier to check for a process to fulfill the good integrator property by checking that the assumptions of Theorem 2.3 are met. That checking this is sufficient will be proved in the next chapter in a more general setting. However, the proof is just an extension of the real-valued case.

At this point the reader might wonder if it is easy to check that a certain process fits our definition of a good integrator. Therefore, we give some basic examples which however show already how quickly one gets quite big classes of functions fulfilling the good integrator property. The first two examples use the characterization of the good integrator property, the third example will use the definition almost right away.

Example 2.4. *Let A be a càdlàg stochastic process of finite variation. Without loss of generality we assume $A_0 = 0$. Given a simple predictable and uniformly bounded strategy $H \in \mathcal{S}_u$, one gets*

$$|(H \cdot A)_t| \leq \sum_{i=0}^{n-1} |H_i||A_{\tau_{i+1}\wedge t} - A_{\tau_i \wedge t}| \leq \|H\|_\infty Var(A)_t, \text{ almost surely,}$$

where $Var(A)$ denotes the variation process of A.

Given a sequence of simple predictable uniformly bounded strategies $(H_n)_{n\geq 1}$ tending to 0 we have by using this inequality for every $t \geq 0$ the convergence of

$$|I_{A^t}(H_n)| = |(H_n \cdot A)_t| \leq \|H_n\|_\infty Var(A)_t \to 0, \text{ almost surely.}$$

Since almost sure convergence implies convergence in probability, A and therefore all processes of finite variation are good integrators.

Example 2.5. *Let now X be an $\mathbb{L}^2-$ martingale. Again, without loss of generality assume $X_0 = 0$. We can use Itô 's insight, which is orthogonality of a martingales increments, to derive for a simple predictable and uniformly bounded strategy $H \in \mathcal{S}_u$ the following inequality for every $t \geq 0$:*

$$\mathbb{E}\left[(H \cdot X)_t^2\right] = \sum_{i,j=0}^{n-1} \mathbb{E}[H_i(X_{\tau_{i+1}\wedge t} - X_{\tau_i\wedge t})H_j(X_{\tau_{j+1}\wedge t} - X_{\tau_j\wedge t})] =$$

$$\sum_{i=0}^{n-1} \mathbb{E}[H_i^2(X_{\tau_{i+1}\wedge t} - X_{\tau_i\wedge t})^2] \leq \|H\|_\infty^2 \sum_{i=0}^{n-1} \mathbb{E}[(X_{\tau_{i+1}\wedge t} - X_{\tau_i\wedge t})^2] \leq \|H\|_\infty^2 \mathbb{E}[X_t^2].$$

Let now $(H_n)_{n\geq 1}$ be a sequence of uniformly bounded predictable strategies such that $\|H_n\|_\infty \to 0$. We can use Chebyshev's inequality to prove for $\epsilon > 0$ arbitrary,

$$\mathbb{P}[|I_{X^t}(H_n)| > \epsilon] \leq \frac{1}{\epsilon^2}\mathbb{E}[(H_n \cdot X)_t^2] \leq \frac{1}{\epsilon^2}\|H_n\|_\infty^2 \mathbb{E}[X_t^2] \to 0.$$

This proves that X and therefore all $\mathbb{L}^2-$ martingales are good integrators.

With our approach of the definition of a stochastic integral, the crucial step in showing that all semimartingales are good integrators lies in proving that all martingales are such. By a stopping argument one then gets that also local martingales are good integrators. Afterwards, by writing a semimartingale as a local martingale plus a finite variation process, the linearity of our definition of a good integrator brings us to the goal of showing that all semimartingales are good integrators.

For this step we will use an inequality, from which one directly sees why martingales are good integrators. The study of this inequality will be a major part of this thesis. We will look into it more detailed and try to tackle it on a more general level in Chapters 3 and 4. Therefore, we will just state it. It will be a consequence of Theorem 3.9 and Remark 3.10. The inequality can also be seen in [Mey72, Theorem 47, Chapter 2] from which we choose the formulation.

Theorem 2.6 (Burkholder inequality)**.** *Let X be a martingale. Then we have for all simple predictable and uniformly bounded strategies $H \in \mathcal{S}_u$ and $c > 0$,*

$$c\mathbb{P}[|(H \cdot X)|_t^* > c] \leq 18\|H\|_\infty \mathbb{E}[|X_t|],$$

for all $t \geq 0$.

The Burkholder inequality can be applied right away to prove the good integrator property for martingales. Note that in this case we prove uniform convergence on compacts in probability right away and will not use that it follows from some stronger convergence. For example when we proved the good integrator for finite variation processes we proved convergence almost surely and deduced convergence in probability from it. For $\mathbb{L}^2-$martingales we showed $\mathbb{L}^2-$convergence and applied Chebvyhsev's inequality.

Example 2.7. *Let X be a martingale. For a sequence $(H^n)_{n\geq 1}$ with $H_n \in \mathcal{S}_u$ such that $\|H^n\|_\infty \to 0$, we have for arbitrary $c > 0$ and $t \geq 0$*

$$\mathbb{P}[|(H^n \cdot X)|_t^* > c] \leq \frac{18}{c}\|H^n\|_\infty \mathbb{E}[|X_t|] \to 0.$$

This proves that X and therefore all martingales are good integrators.

It is remarkable that the Burkholder inequality can be applied almost right away to our definition of a good integrator. The only thing different is that we use uniformly bounded predictable simple strategies, rather than just predictable simple strategies. This not only shows in a very elegant and quick way how we can define the stochastic integral but also the importance of this inequality. Actually this inequality also plays a major role in Josef Teichmann's proof of the Fundamental Theorem of Asset Pricing (FTAP), see [CT15, Lemma 4.7], which is arguably the most important theorem in financial mathematics. These are some of the reasons why this thesis focuses a lot on this inequality. We tried and worked out different approaches in generalizing this inequality to more general spaces in Chapters 3 and 4.

Now that we know all semimartingales are good integrators we can talk about the Bichteler-Dellacherie theorem which tells us that semimartingales are already all processes which are good integrators. The Bichteler-Dellacherie theorem in our setting can simply be formulated as follows.

Theorem 2.8. *Let X be a càdlàg and adapted process. Then X is a good integrator if and only if it is the sum of a càdlàg local martingale and a càdlàg finite variation process.*

We chose this formulation since this precisely tells us that the good integrator property is the right definition for the stochastic integral. Originally, the theorem popped up in its first version by Klaus Bichteler in [Bic79, Theorem 1]. Today, there are many different proofs for this theorem, including for example [BSV11, Theorem 1.2] and [BS12, Theorem BD]. The celebrated Bichteler-Dellacherie theorem precisely tells which processes are useful for stochastic integration. The second major part of this thesis is therefore devoted to this theorem. In Chapter 7 we try to tackle the theorem in a more general version and drop the càdlàg assumption. We orient ourselves at a proof by Christophe Stricker, see his paper [Str84] and generalize it.

But before this theorem we turn back out attention to the Burkholder inequality and the good integrator property in the next two chapters.

3 Banach Space-valued Integrands, Real-valued Integrators

This chapter focuses on stochastic integration of Banach space-valued integrands with respect to real-valued integrators.

The first section is devoted to the Burkholder inequality which we already saw in the previous chapter. A new definition of so called Burkholder spaces will be introduced which essentially defines Banach spaces in which the Burkholder inequality works.

The second section of this chapter focuses on generalizing the good integrator property to this setting and we give examples of Banach spaces in which this theory is applicable.

In the third section we tried to work out the good integrator property for UMD (Unconditional Martingale Difference) spaces, see also Chapter 8.4 for more on UMD spaces. We compare our approach to other approaches of introducing a stochastic integral in a Banach spaces setting like in [vVW15] for example. Unfortunately, we did not fully succeed in this chapter yet, so we cannot give a positive to the question if the good integrator theory works in UMD spaces. However, the interested reader might get some ideas how one could tackle this open question, so it is still presented.

3.1 Burkholder Spaces

In this section we find a sufficient condition to get a Burkholder inequality for Banach space-valued integrands and real-valued integrators. For this sake we will define a certain property which is fulfilled for certain Banach spaces. However, a necessary criterion for a Burkholder inequality is not yet found and it is left open, whether the definition written in this section characterizes those Banach spaces which have a Burkholder inequality. Since the constants for the inequality are the same if we take as Banach space just the real numbers, the generalized Burkholder inequality developed in this section is an exact extension of the real-valued one, which we already saw in the previous chapter. We will present examples of Banach spaces for which a Burkholder inequality can be applied, i.e. in our setting Banach spaces which fulfil the property defined in this section.

For the rest of this section consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is a right-continuous filtration if we do not mention another one.

We start with the definition of the sufficient property.

Definition 3.1 (Burkholder space)**.** *For $p > 1$ a Banach space E is said to be a Burkholder p space, or short a $(BH)_p$ space, if there exists a constant $C_{E,p} > 0$ such that for all real-valued $\mathbb{L}^p-$ martingale difference sequences $(d_i)_{i=1}^n$ w.r.t. a filtration $(\mathcal{F}_i)_{i=0}^n$ and random variables $H_1, \dots, H_n \in \mathbb{L}^\infty(\Omega; E)$ such that H_i is measurable w.r.t. $\mathcal{F}_{i-1}$, one has*

$$\mathbb{E}\left[\Big\| \sum_{i=1}^n H_i d_i \Big\|^p\right] \leq C_{E,p}^p \max_{1\leq i\leq n} \|H_i\|^p_{\mathbb{L}^\infty(\Omega;E)} \mathbb{E}\left[\Big|\sum_{i=1}^n d_i\Big|^p\right]. \tag{3.1}$$

If the property holds for $p = 2$, E will just be called a Burkholder space or (BH) space and we set $C_E := C_{E,2}$. We will also sometimes simply say that E has the Burkholder property in this case.

Remark 3.2. *For completeness, as a short remark recall the definition of $\mathbb{L}^p-$ martingale difference sequences. An $\mathbb{L}^p-$ martingale difference sequence $(d_i)_{i=1}^n$ with respect to a filtration $(\mathcal{F}_i)_{i=0}^n$ is a family of random variables $(d_i)_{i=1}^n$ such that d_i is measurable w.r.t. $\mathcal{F}_i$ and*

$$\mathbb{E}[d_i \mid \mathcal{F}_{i-1}] = 0, \quad i = 1, \dots, n.$$

Given an $\mathbb{L}^p-$ martingale $(M_i)_{i=0}^n$, one can define its $\mathbb{L}^p-$ martingale difference sequence via

$$d_i := M_i - M_{i-1}, \quad i = 1, \dots, n.$$

Therefore, every $\mathbb{L}^p-$ martingale defines an $\mathbb{L}^p-$ martingale difference sequence.

It also works vice versa, i.e. given an $\mathbb{L}^p-$ martingale difference sequence $(d_i)_{i=1}^n$, one can define

$$M_i := \sum_{j=1}^i d_j, \quad i = 1, \dots, n$$

to get an $\mathbb{L}^p-$ martingale $(M_i)_{i=0}^n$ starting at 0, since for $i \in \{1, \dots, n\}$ it holds

$$\mathbb{E}[M_i \mid \mathcal{F}_{i-1}] = \mathbb{E}\Big[\sum_{j=1}^i d_j \mid \mathcal{F}_{i-1}\Big] = \sum_{j=1}^{i-1} d_j + \mathbb{E}[d_i \mid \mathcal{F}_{i-1}] = M_{i-1}.$$

In the following if we just write martingale difference sequence, we will always refer to a $\mathbb{L}^2-$ martingale difference sequence.

Looking at the definition of Burkholder spaces, in the case of $p = 2$ one can think of Ito's insight (see also Example 2.5) and that inequality (3.1) refers to some kind of orthogonality of the increments of the martingale corresponding to the martingale difference sequence. This is also perfectly visualized by the following example. To fix some notation beforehand right away, given a martingale $(M_i)_{i=0}^n$ and a stochastic process H with Banach space-valued random variables $H_0, \dots, H_{n-1}$ such that H_i is measurable with respect to $\mathcal{F}_i$ we denote the discrete stochastic integral by $(H \cdot M)$, i.e.

$$(H \cdot M)_i := \sum_{j=0}^{i-1} H_j (M_{j+1} - M_j), \quad i = 0, \dots, n.$$

Fixed this, we present the example.

Example 3.3. *Every Hilbert space H is a Burkholder space with constant $C_H = 1$. Denote the inner product of H by $\langle \cdot , \cdot \rangle$, then we get with the bilinearity of the inner product for an $\mathbb{L}^2-$ martingale $(M_i)_{i=0}^n$ and random variables $H_0, \dots, H_{n-1} \in \mathbb{L}^\infty(\Omega; H)$ such that H_i is measurable w.r.t. $\mathcal{F}_i$ the equality*

$$\begin{aligned} \mathbb{E}[\|(H \cdot M)_n\|^2] = \mathbb{E}[\|\sum_{i=0}^{n-1} H_i (M_{i+1} - M_i)\|^2] = \\ = \mathbb{E}\left[\left\langle \sum_{i=0}^{n-1} H_i (M_{i+1} - M_i) , \sum_{i=0}^{n-1} H_i (M_{i+1} - M_i) \right\rangle\right] = \\ = \sum_{i \neq j} \mathbb{E}[[\langle H_i , H_j \rangle (M_{i+1} - M_i)(M_{j+1} - M_j)] + \sum_{i=0}^{n-1} \mathbb{E}\left[\|H_i\|^2 (M_{i+1} - M_j)^2\right]. \end{aligned}$$

For the left term we can use orthogonality in the following way: assume $i < j$, then

$$\begin{aligned} \mathbb{E}\left[\langle H_i , H_j \rangle (M_{i+1} - M_i)(M_{j+1} - M_j)\right] = \\ = \mathbb{E}[\langle H_i , H_j \rangle (M_{i+1} - M_i) \underbrace{\mathbb{E}[(M_{j+1} - M_j) \mid \mathcal{F}_j]}_{=0}] = 0. \end{aligned}$$

In total this leads to

$$\mathbb{E}[\|(H\cdot M)_n\|^2] = \sum_{i=0}^{n-1} \mathbb{E}\left[\|H_i\|^2 (M_{i+1}-M_j)^2\right] \leq$$
$$\leq \max_{0\leq i\leq n-1} \|H_i\|^2_{\mathbb{L}^\infty(\Omega;E)} \sum_{i=0}^{n-1} \mathbb{E}[(M_{i+1}-M_i)^2] =$$
$$= \max_{0\leq i\leq n-1} \|H_i\|^2_{\mathbb{L}^\infty(\Omega;E)} \mathbb{E}[M_n^2 - M_0^2] =$$
$$= \max_{0\leq i\leq n-1} \|H_i\|^2_{\mathbb{L}^\infty(\Omega;E)} \mathbb{E}\left[\big| \sum_{i=0}^{n-1} M_{i+1} - M_i \big|^2 \right].$$

Since every $\mathbb{L}^2-$martingale difference sequence comes from an $\mathbb{L}^2-$martingale this shows that H is a Burkholder space with constant $C_H = 1$ if we look at the martingale difference associated to $(M_i)_{i=0}^n$, i.e. $d_i = M_i - M_{i-1}$, for $i \in \{1,\dots,n\}$.

Next, we will give an example of a Banach space fulfilling our property which is not as usual as a Hilbert space and might not be familiar to the reader. Therefore, we will also recall the definition of such a space in the example.

Example 3.4. *Let E be a martingale type 2 space, for details see Definition 8.13. Essentially we get a constant $\mu \geq 0$ such that for all finite $E-$ valued martingale difference sequences $(d_i)_{i=1}^n$ it holds*

$$\mathbb{E}\left[\big\| \sum_{i=1}^{n} d_i \big\|^2 \right] \leq \mu^2 \sum_{i=1}^{n} \mathbb{E}[\|d_i\|^2]. \tag{3.2}$$

Let now X be an $\mathbb{L}^2-$ martingale w.r.t. a filtration $(\mathcal{F}_i)_{i=0}^n$ and $H_1,\dots,H_n \in \mathbb{L}^\infty(\Omega;E)$ such that H_i is measurable w.r.t. $\mathcal{F}_i$. Writing the discrete stochastic integral $(H\cdot X)$ as a telescopic sum, i.e.

$$(H\cdot X)_n = \sum_{i=1}^{n} (H\cdot X)_i - (H\cdot X)_{i-1} = \sum_{i=1}^{n} H_{i-1}(X_i - X_{i-1}),$$

we get by looking at the martingale difference sequence $d_i := (H \cdot X)_i - (H \cdot X)_{i-1}$*,*

$$\begin{aligned}\mathbb{E}[\|(H \cdot X)_n\|^2] = \mathbb{E}\left[\Big\|\sum_{i=1}^n d_i\Big\|^2\right] &\overset{(3.2)}{\leq} \mu^2 \sum_{i=1}^n \mathbb{E}[\|d_i\|^2] = \\ &= \mu^2 \sum_{i=1}^n \mathbb{E}[\|H_{i-1}(X_i - X_{i-1})\|^2] \leq \\ &\leq \mu^2 \max_{0\leq i\leq n-1} \|H_i\|^2_{\mathbb{L}^\infty(\Omega;E)} \sum_{i=1}^n \mathbb{E}[(X_i - X_{i-1})^2] \leq \\ &\leq \mu^2 \max_{0\leq i\leq n-1} \|H_i\|^2_{\mathbb{L}^\infty(\Omega;E)} \mathbb{E}[X_n^2 - X_0^2].\end{aligned}$$

This proves that every martingale type 2 *space is a Burkholder space with constant* $C_E = \mu$*.*

After these examples we are almost at the point to tackle the Burkholder inequality. In the proof however, another well known inequality for non-negative and bounded supermartingales is needed. For this sake this inequality will be introduced. First a definition is needed and for completeness a proof of the result is given as well. We will work very closely to [Mey72, Chapter 2] and take some results from there. Also notice that in the following we work with processes in discrete time and take as time points the non-negative integers.

Let us first introduce the notion of a potential taken from [Mey72, Definition 34].

Definition 3.5 (Potential)**.** *Let* X *be a non-negative supermartingale. Then it is a potential if it does not dominate any non-negative martingale except* 0*.*

A potential can be characterized by the following theorem. Its precise proof can be found in [Mey72, Chapter 2, Theorem 35].

Theorem 3.6. *A non-negative supermartingale* X *is a potential if and only if*

$$\lim_{n\to\infty} \mathbb{E}[X_n] = 0.$$

By this theorem one can deduce for a potential X that $\lim_{n\to\infty} X_n =: X_\infty$ exists and is equal to 0 almost surely. In general we will write for stochastic processes Y where $\lim_{n\to\infty} Y_n$ exists almost surely, $Y_\infty := \lim_{n\to\infty} Y_n$.

Recall the Doob-Meyer decomposition on our countable index set of the non-negative integers for a non-negative supermartingale X, i.e. $X = M - A$ where M is a martingale and A a predictable non-decreasing process. This decomposition is unique and can actually be defined via

$$A_t := \sum_{s<t} \mathbb{E}[X_s - X_{s+1} \mid \mathcal{F}_s]$$

for $t \in \mathbb{N}$ and $M := X + A$. Since X is a supermartingale, A is non-decreasing. To see that M is a martingale one just has to calculate

$$\begin{aligned}\mathbb{E}[M_{t+1} \mid \mathcal{F}_t] &= \mathbb{E}\big[X_{t+1} + \sum_{s<t+1} \mathbb{E}[X_s - X_{s+1} \mid \mathcal{F}_s] \mid \mathcal{F}_t\big] = \\ &= \mathbb{E}[X_{t+1} \mid \mathcal{F}_t] + \sum_{s<t} \mathbb{E}[X_s - X_{s+1} \mid \mathcal{F}_s] + \mathbb{E}[X_t - X_{t+1} \mid \mathcal{F}_t] = \\ &= X_t + A_t = M_t.\end{aligned}$$

In fact, such a decomposition holds for any integrable adapted stochastic process with index set the non-negative integers into a unique martingale and a unique predictable process starting at 0 (see also [Tei19, Section 2]).

If in the following X is a potential, then also the random variables A_∞, M_∞ and X_∞ of its unique decomposition exist almost surely.

We are now at the point to state and give a proof of a bound for the moments of A_∞. This is taken and rephrased from [Mey72, Chapter 2, Theorem 45.2].

Theorem 3.7. *Let X be a potential bounded by a constant $c \geq 0$. Then, for any integer $p \geq 1$ it holds*

$$\mathbb{E}[A_\infty^p] \leq p! c^{p-1} \mathbb{E}[X_0] \leq p! c^p. \tag{3.3}$$

Proof. We write A_∞ as a telescopic sum (note $A_0 = 0$), i.e.

$$\sum_{i \geq 0} A_{i+1} - A_i.$$

This can be raised to the power p to obtain

$$A_\infty^p = \sum_{i_1,\ldots,i_p \geq 0} (A_{i_1+1} - A_{i_1}) \cdots (A_{i_p+1} - A_{i_p}).$$

From the integers $i_1, \ldots, i_p$ we then pick the largest one, call it j and single it out. It can occur at p positions, so we get

$$A_\infty^p = p \sum_{j \geq 0} \sum_{j_1,\ldots,j_{p-1} \leq j} (A_{j_1+1} - A_{j_1}) \cdots (A_{j_{p-1}+1} - A_{j_{p-1}})(A_{j+1} - A_j).$$

Because A is non-negative and non-decreasing all terms are non-negative and one can apply Fubini's theorem to exchange summation. By summing over j first, we arrive at

$$A_\infty^p = p \sum_{j_1,\ldots,j_{p-1}} (A_{j_1+1} - A_{j_1}) \cdots (A_{j_{p-1}+1} - A_{j_{p-1}})(A_\infty - A_{j_1 \vee \ldots \vee j_{p-1}}). \tag{3.4}$$

If we take the conditional expectation w.r.t. $\mathcal{F}_{j_1 \vee \ldots \vee j_{p-1}}$ for each summand, actually every but the last term of the product is measurable, therefore by recognizing the decomposition of X, we get for each summand

$$\begin{aligned} &\mathbb{E}[(A_{j_1+1} - A_{j_1}) \cdots (A_{j_{p-1}+1} - A_{j_{p-1}})(A_\infty - A_{j_1 \vee \ldots \vee j_{p-1}}) \mid \mathcal{F}_{j_1 \vee \ldots \vee j_{p-1}}] \\ &= (A_{j_1+1} - A_{j_1}) \cdots (A_{j_{p-1}+1} - A_{j_{p-1}})\mathbb{E}[(A_\infty - A_{j_1 \vee \ldots \vee j_{p-1}}) \mid \mathcal{F}_{j_1 \vee \ldots \vee j_{p-1}}] \\ &= (A_{j_1+1} - A_{j_1}) \cdots (A_{j_{p-1}+1} - A_{j_{p-1}})X_{j_1 \vee \ldots \vee j_{p-1}}. \end{aligned}$$

Arrived at this point use the bound c of X and the form of (3.4) to arrive at

$$\mathbb{E}[A_\infty^p] \leq cp\mathbb{E}[A_\infty^{p-1}].$$

This argument can be iterated to get

$$\mathbb{E}[A_\infty^p] \leq c^{p-1}p!\mathbb{E}[A_\infty].$$

As a last step look at

$$\mathbb{E}[A_\infty] = \mathbb{E}[X_0] - \mathbb{E}[X_\infty] = \mathbb{E}[X_0],$$

which finally gives

$$\mathbb{E}[A_\infty^p] \leq p!c^{p-1}\mathbb{E}[X_0] \leq p!c^p.$$

This is the desired inequality and ends the proof. □

This bound can be generalized to non-negative and bounded supermartingales via the following theorem. This is again taken and reformulated from Meyers work, in this case [Mey72, Chapter 2, Theorem 46].

Theorem 3.8. *Assume X is a non-negative supermartingale bounded by some constant $c > 0$. Having its Doob decomposition $X = M - A$, in particular $M_\infty = A_\infty + X_\infty$, for an integer $p \geq 1$ we get the bound*

$$\mathbb{E}[M_\infty^p] \leq p!c^{p-1}\mathbb{E}[X_0].$$

Proof. For $n \in \mathbb{N}$ let us define a new supermartingale via $\tilde{X}_i := X_i$ if $i \leq n$ and $\tilde{X}_i = 0$ if $i > n$ w.r.t. the filtration $\mathcal{F}_0, \ldots, \mathcal{F}_n, \mathcal{F}, \mathcal{F}, \ldots$. Clearly, $\tilde{X}$ is a potential and for its decomposition $\tilde{X} = \tilde{M} - \tilde{A}$ there is the inequality

$$\tilde{A}_\infty = A_n + \mathbb{E}[X_n - X_\infty \mid \mathcal{F}_n] + X_\infty \geq A_n + X_\infty.$$

With this we are in the position to apply the previous theorem, in particular inequality (3.3) to arrive at

$$\mathbb{E}[(A_n + X_\infty)^p] \leq \mathbb{E}[\tilde{A}_\infty^p] \leq p!c^{p-1}\mathbb{E}[\tilde{X}_0] = p!c^{p-1}\mathbb{E}[X_0].$$

Finally let n tend to infinity and apply dominated convergence to get

$$\mathbb{E}[M_\infty^p] = \mathbb{E}[(A_\infty + X_\infty)^p] \leq p!c^{p-1}\mathbb{E}[X_0].$$

□

The previous theorem will be crucial in proving the first main result of this thesis which is the generalization of the Burkholder inequality. Having the estimate of Theorem 3.8 and the definition of the Burkholder property in mind we are in the position to formulate and prove it. The idea was essentially to generalize the proof of Paul-André Meyer in [Mey72, Chapter 2, Theorem 47]. For this we switch back to our usual filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ fixed in the beginning of this section.

Theorem 3.9 (Burkholder inequality)**.** *Let X be a real-valued martingale and $(E, \|\cdot\|)$ a Burkholder space. Define*

$$\mathcal{S} := \Big\{H = \sum_{i=0}^{n-1} H_i \mathbb{1}_{(\tau_i, \tau_{i+1}]} \mid H_i \in \mathbb{L}^0(\Omega; E) : \ H_i \text{ is } \mathcal{F}_{\tau_i} - \text{strongly measurable} \\ \text{and } 0 \leq \tau_0 \leq \ldots \leq \tau_n \text{ are stopping times}\Big\},$$

and set

$$\mathcal{S}_u := \{H \in \mathcal{S} \mid H_i \in \mathbb{L}^\infty(\Omega; E), \ i = 0, \ldots n-1\}.$$

Then we have for all $T > 0$, $c > 0$ and $H \in \mathcal{S}_u$:

$$c\mathbb{P}[\|(H \cdot X)\|_T^* > c] \leq (10 + 8C_E^2)\|H\|_\infty \mathbb{E}[|X_T|], \tag{3.5}$$

where for $H \in \mathcal{S}_u$ written as $H = \sum_{i=0}^{n-1} H_i \mathbb{1}_{(\tau_i, \tau_{i+1}]}$, we set

$$\|H\|_\infty := \max_{0 \leq i \leq n-1} \|H\|_{\mathbb{L}^\infty(\Omega; E)}.$$

Proof. Before proving the general result, let us do some reductions first.

Assume $X \geq 0$, let $c > 0$ and $H \in \mathcal{S}_u$ for which we additionally assume $\|H\|_\infty \leq 1$. Since H has finitely many values it is enough to consider a discrete martingale with finitely many time points.

Having this reduction, define $Z := X \wedge c$ which is by Jensen's inequality a supermartingale. Therefore it admits a discrete Doob Meyer decomposition $Z = M - A$ where M is a martingale and A a non-negative non-decreasing process with $A_0 = 0$.

Using that on the set $\{|X|^*_{t_n} < c\}$ it holds $X = Z$, we get

$$\begin{aligned} c\mathbb{P}[\|(H\cdot X)\|^*_{t_n} > c] &\leq c\mathbb{P}[|X|^*_{t_n} \geq c] + \mathbb{P}[|X|^*_{t_n} < c,\ \|(H\cdot X)\|^*_{t_n} > c] = \\ &= c\mathbb{P}[|X|^*_{t_n} \geq c] + c\mathbb{P}[\|(H\cdot Z)\|^*_{t_n} > c]. \end{aligned}$$

As a next step we use that A is increasing, $A_0 = 0$ and $\|H\|_\infty \leq 1$, so we get for any $j \in \{1,\dots,n\}$ the inequality

$$\begin{aligned} \|(H\cdot Z)_{t_j}\| &\leq \|(H\cdot M)_{t_j}\| + \|(H\cdot A)_{t_j}\| = \\ &= \|(H\cdot M)_{t_j}\| + \|\sum_{i=0}^{j-1} H_i(A_{t_{i+1}} - A_{t_i})\| \leq \\ &\leq \|(H\cdot M)_{t_j}\| + \sum_{i=0}^{j-1}(A_{t_{i+1}} - A_{t_i}) = \\ &= \|(H\cdot M)_{t_j}\| + A_{t_j}. \end{aligned}$$

Since the norm is convex and A non-decreasing actually $\|(H\cdot M)\| + A$ is a submartingale and therefore by Jensen's inequality also its square. We apply Doob's maximal inequality to estimate

$$\begin{aligned} c\mathbb{P}[\|(H\cdot Z)\|^*_{t_n} > c] &\leq \frac{c^2}{c}\mathbb{P}[(|\|(H\cdot M)\| + A|^*_{t_n})^2 > c^2] \leq \\ &\overset{\text{Doob}}{\leq} \frac{1}{c}\mathbb{E}[(\|(H\cdot M)_{t_n}\| + A_{t_n})^2] \leq \\ &\overset{(a+b)^2\leq 2(a^2+b^2)}{\leq} \frac{2}{c}\left(\mathbb{E}[\|(H\cdot M)_{t_n}\|^2] + \mathbb{E}[A^2_{t_n}]\right). \end{aligned}$$

Z is a non-negative supermartingale which is by its definition bounded by c and $Z = M - A$. We will now apply Theorem 3.8 with $p = 2$. It also holds that $Z = M - A \geq 0$, implying $A \leq M$. With this we arrive at

$$\mathbb{E}[A^2_{t_n}] \leq \mathbb{E}[M^2_{t_n}] \leq 2c\mathbb{E}[Z_0] \leq 2c\mathbb{E}[X_0] = 2c\mathbb{E}[X_{t_n}],$$

which also implies $M \in \mathbb{L}^2$.

As next step we can use the Burkholder property of E for the martingale difference sequence $(d_i)_{i=0}^{n-1} := (M_{t_{i+1}} - M_{t_i})_{i=0}^{n-1}$ to obtain (using $\|H_i\|_{\mathbb{L}^\infty(\Omega;E)} \leq 1$)

$$\begin{aligned} \mathbb{E}[\|(H\cdot M)_{t_n}\|^2] = \mathbb{E}\big[\|\sum_{i=0}^{n-1} H_i d_i\|^2\big] \leq C_E^2\mathbb{E}\left[|\sum_{i=0}^{n-1} M_{t_{i+1}} - M_{t_i}|^2\right] = \\ = C_E^2\mathbb{E}[(M_{t_n} - M_0)^2] \leq C_E^2\mathbb{E}[M^2_{t_n}] \leq 2cC_E^2\mathbb{E}[X_{t_n}]. \end{aligned}$$

Plugging this in we can estimate

$$\begin{aligned} c\mathbb{P}[\|(H\cdot X)\|^*_{t_n}\geq c] &\leq c\mathbb{P}[|X|^*_{t_n}\geq c]+c\mathbb{P}[\|(H\cdot Z)\|^*_{t_n}\geq c]\leq \\ &\leq \mathbb{E}[X_{t_n}]+\frac{2}{c}\left(\mathbb{E}[\|(H\cdot M)\|^2_{t_n}]+\mathbb{E}[A^2_{t_n}]\right)\leq \\ &\leq \mathbb{E}[X_{t_n}]+\frac{2}{c}(2cC_E^2\mathbb{E}[Xt_n]+2c\mathbb{E}[X_{t_n}])= \\ &= (5+4C_E^2)\mathbb{E}[X_{t_n}]. \end{aligned}$$

After this, for a general X, one can split it into a difference of two non-negative martingales and apply what we just proved to get the same constant multiplied by two, i.e.

$$c\mathbb{P}[\|(H\cdot X)\|^*_{t_n}\geq c]\leq (10+8C_E^2)\mathbb{E}[|X_{t_n}|].$$

□

After having proved this result this section is almost done. We give a final remark justifying that we called our Burkholder inequality a generalization of the original one.

Remark 3.10. *The original Burkholder inequality, which covers the real-valued case, is the following inequality. For a martingale X with $X_0=0$ it holds*

$$c\mathbb{P}[|(H\cdot X)|^*_T\geq c]\leq 18\|H\|_\infty\mathbb{E}[|X_T|],\quad \forall T>0,\ \forall H\in\mathcal{S}_u.$$

This is in accordance with Theorem 3.9 because if $E=\mathbb{R}$ (or any Hilbert space), then $C_E=1$ and $10+8C_E^2=18$. Therefore, Theorem 3.9 is a natural extension of the Burkholder inequality. Also, clearly $\mathbb{R}$ satisfies the Burkholder property since, in particular, it is a Hilbert space.

3.2 Good Integration

After having generalized the Burkholder inequality in the previous section, we remember from Chapter 2 that this was the crucial inequality to get that martingales are good integrators in the real-valued case.

In the following let us try to generalize this idea to the setting of this Chapter where Banach space-valued strategies are considered. Recall that in the real-valued case processes which can be used for stochastic integration can be characterized by the good integrator property, see [Pro04] for all details. We will present an analogous definition for processes where we look at the simple stochastic integral of

Banach space-valued processes against real-valued processes. The results proved in this section also show the corresponding theorems presented in Chapter 2 since they only correspond to a special choice of the Banach space, i.e. $\mathbb{R}$.

For the remainder of this section consider a separable Banach space $(E, \|\cdot\|)$. As earlier define the spaces of càglàd and càdlàg adapted stochastic processes as

$$\begin{aligned} \mathcal{L} &:= \{H : \Omega \times \mathbb{R}_+ \to E \mid H(\omega, \cdot) \text{ càglàd}, \ H \text{ adapted}\}, \\ \mathcal{D} &:= \{X : \Omega \times \mathbb{R}_+ \to E \mid X(\omega, \cdot) \text{ càdlàg }, \ X \text{ adapted}\}, \end{aligned}$$

where we identify processes up to indistinguishability. With adapted here we always mean strongly adapted, i.e. an $E-$valued stochastic process X is adapted with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if and only if for every $t \geq 0$ it holds that X_t is strongly $\mathcal{F}_t-$measurable, see also Chapter 8.1 in the appendix for more details.

Remember the spaces of simple predictable strategies $\mathcal{S}$ and simple predictable and uniformly bounded strategies $\mathcal{S}_u$ which were defined in Theorem 3.9.

We will equip these spaces with a metric, namely the straight-forward generalization of the ucp-metric from the real-valued case. This means that we just replace absolute values by norms.

Definition 3.11. *For $X,\ Y \in \mathcal{L}$ or $X,\ Y \in \mathcal{D}$ define the ucp-metric d to be*

$$d(X, Y) := \sum_{n\geq 1} \frac{1}{2^n} \mathbb{E}[\|X - Y\|_n^* \wedge 1] \quad (\leq 1). \tag{3.6}$$

It is clear that this defines a metric, since d is symmetric and the triangle inequality follows from the inequality of the norm. Since $d(X, X) = 0 \iff X_t = 0 \ a.s.$ for all $t \geq 0$, d satisfies definiteness as well. Equipped with the ucp metric, $\mathcal{L}$ and $\mathcal{D}$ are topological vector spaces.

We will give a prove that $\mathcal{L}$ and $\mathcal{D}$ are complete topological vector spaces with respect to d. It will be an application of the next lemma. It is formulated as extra lemma since it is basically the desired result on a pathwise level.

Lemma 3.12. *Let $(F, \|\cdot\|)$ be a Banach space. Define the linear spaces*

$$\begin{aligned} \mathcal{C} &:= \{f : \mathbb{R}_{\geq 0} \to F \mid f \text{ càdlàg}\} \\ \mathcal{B} &:= \{f : \mathbb{R}_{\geq 0} \to F \mid f \text{ càglàd}\}. \end{aligned}$$

Also define the metric

$$\rho(f, g) := \sum_{k\geq 1} \frac{1}{2^k} \left(\|f - g\|_k^* \wedge 1\right), \quad f, g : \mathbb{R}_{\geq 0} \to F.$$

In the definition of ρ, we set

$$\|f\|_k^* := \sup_{t\leq k}\|f(t)\|, \quad f \in \mathcal{B} \text{ or } f \in \mathcal{C}.$$

Then $(\mathcal{B}, \rho)$ and $(\mathcal{C}, \rho)$ are complete vector spaces.

Proof. We will prove the statement for $(\mathcal{C}, \rho)$, for $(\mathcal{B}, \rho)$ it works analogously.

That ρ defines a metric is not hard since it is just the ucp metric for paths. In fact, symmetry is clear and the triangle inequality follow from the triangle inequality of the norm since for $f, g, h \in \mathcal{C}$ we have

$$\rho(f,h) \leq \sum_{k\geq 1}\frac{1}{2^k}\sup_{t\leq k}(\|f(t)-g(t)\| + \|g(t)-h(t)\|) \wedge 1 \leq \rho(f,g) + \rho(g,h).$$

If $f = g$, of course $\rho(f,g) = 0$ and if $\rho(f,g) = 0$ we have

$$\|f-g\|_k^* = 0, \quad \forall k \in \mathbb{N}.$$

This implies $f = g$. Therefore, definiteness of ρ is proved and it actually defines a metric.

To prove completeness, let $(f_n)_{n\geq 1}$ be a Cauchy sequence in $\mathcal{D}$. Convergence uniformly on compacts which is induced by ρ implies pointwise convergence, which we can formalize for our problem. This means we need to show that $(f_n(t))_{n\geq 1}$ is a Cauchy sequence in F. For this sake, let a point $t \in \mathbb{R}_{\geq 0}$ be given and fix $k \in \mathbb{N}$ with $k > t$. For a given $\epsilon > 0$ let $n_0 \in \mathbb{N}$ such that for $m, n \geq n_0$ we have $\rho(f_n, f_m) < \epsilon/2^k$. Use this n_0 to get for $m, n \geq n_0$

$$\|f_n(t) - f_m(t)\| \leq 2^k\rho(f_n, f_m) < \epsilon.$$

Therefore, $(f_n(t))_{n\geq 1}$ is a Cauchy sequence in F and there exists a point $f(t) \in F$ such that $f_n(t) \to f(t)$.

Having this pointwise limit we are left with showing that $\rho(f_n, f) \to 0$ and $f \in \mathcal{D}$. For proving convergence with respect to ρ let $\epsilon > 0$ be given and choose $n_0 \in \mathbb{N}$ such that for all $m, n \geq n_0$, it holds $\rho(f_n, f_m) < \epsilon$. Now look at

$$\begin{aligned}
\rho(f_n, f) &= \sum_{k\geq 1}\frac{1}{2^k}\sup_{t\leq k}\lim_{m\to\infty}\|f_n(t) - f_m(t)\| \wedge 1 \leq \\
&\overset{(\star)}{\leq} \sum_{k\geq 1}\frac{1}{2^k}\liminf_{m\to\infty}\sup_{t\leq k}\|f_n(t) - f_m(t)\| \wedge 1 \leq \\
&\leq \limsup_{m\to\infty}\underbrace{\rho(f_n, f_m)}_{<\epsilon\ \forall m\geq n_0} \leq \epsilon.
\end{aligned}$$

Let us now show the inequality ($\star$). Note that

$$\forall s \leq k, \forall m \in \mathbb{N}: \; \|f_n(s) - f_m(s)\| \leq \sup_{t \leq k} \|f_n(t) - f_m(t)\|.$$

This implies, noting that $\liminf_{m\to\infty} f_m(s) = \lim_{m\to\infty} f_m(s)$, the following

$$\forall s \leq k: \; \lim_{m\to\infty} \|f_n(s) - f_m(s)\| \leq \liminf_{m\to\infty} \sup_{t \leq k} \|f_n(t) - f_m(t)\|,$$

which finally implies

$$\sup_{t \leq k} \lim_{m\to\infty} \|f_n(t) - f_m(t)\| \leq \liminf_{m\to\infty} \sup_{t \leq k} \|f_n(t) - f_m(t)\|.$$

This proves ($\star$). All in all, we now know $\rho(f_n, f) \to 0$. To prove that $f \in \mathcal{C}$, i.e. that it has càdlàg paths, note that since we have uniform convergence on compacts limits can be interchanged and therefore for $t \in \mathbb{R}_{\geq 0}$ it holds

$$\lim_{s\downarrow t} f(s) = \lim_{s\downarrow t} \lim_{n\to\infty} f_n(s) = \lim_{n\to\infty} \lim_{s\downarrow t} f_n(s) = \lim_{n\to\infty} f_n(t) = f(t)$$
$$\lim_{s\uparrow t} f(s) = \lim_{s\uparrow t} \lim_{n\to\infty} f_n(s) = \lim_{n\to\infty} \lim_{s\uparrow t} f_n(s) = \lim_{n\to\infty} f_n(s-).$$

The last limit exists since $f_n(\cdot-) \in \mathcal{B}$ and there it is a Cauchy sequence and $(f_n(s-))_{n\geq 1}$ has a limit as well.

This finally shows $f \in \mathcal{C}$ and finishes the prove that $(\mathcal{C}, \rho)$ is complete. □

As mentioned before the lemma, showing that the spaces of càdlàg adapted stochastic and càglàd processes are complete with respect to the ucp metric will be an application of the lemma on a pathwise level.

Proposition 3.13. *$\mathcal{D}$ and $\mathcal{L}$ are complete topological vector spaces with respect to the ucp metric d.*

Proof. We will prove the result for $\mathcal{D}$ since the proof works analogously for $\mathcal{L}$.

To prove completeness let $(X^n)_{n\geq 1}$ be a Cauchy sequence of càdlàg processes in $\mathcal{D}$. One can pick a subsequence $Z^n := X^{m_n}$ such that it holds

$$d(Z^n, Z^m) \leq \frac{1}{2^n}, \quad \forall m \geq n.$$

For this subsequence by monotone convergence we get

$$\mathbb{E}\Big[\sum_{n\geq 1} \sum_{k\geq 1} \frac{1}{2^k} \|Z^n - Z^{n+1}\|_k^* \wedge 1\Big] = \sum_{n\geq 1} d(Z^n, Z^{n+1}) \leq 1.$$

This tells that there exists a measurable set $\tilde{\Omega} \subset \Omega$ with $\mathbb{P}[\tilde{\Omega}] = 1$ and using the metric ρ of the previous lemma for $\omega \in \tilde{\Omega}$ it holds

$$\sum_{n\geq 1} \rho(Z^n(\omega), Z^{n+1}(\omega)) < \infty.$$

Since the remainder of a converging series tends towards zero, for $\omega \in \tilde{\Omega}$ the sequence $(Z^n(\omega))_{n\geq 1}$ is a Cauchy sequence in

$$\mathcal{C} := \{f : \mathbb{R}_{\geq 0} \to \mathbb{E} \mid f \text{ càdlàg}\}.$$

At this point the previous Lemma can be applied to get càdlàg functions $X(\omega) \in \mathcal{C}$ such that $\rho(Z^n(\omega), X(\omega)) \to 0$ for $\omega \in \tilde{\Omega}$. Since this convergence implies almost sure convergence of $Z_t^n \to X_t$, the random variable X_t is measurable. In particular if the processes Z^n are adapted, X is adapted as well.

At this point X is a càdlàg adapted stochastic process, therefore in $\mathcal{D}$ and we are left with showing that the original sequence $(X^n)_{n\geq 1}$ converges to X with respect to the ucp metric. First we show $Z^n \xrightarrow{d} X$. This can be proved by applying dominated convergence since almost surely it holds $\rho(Z^n, X) \leq 1$ and also $\rho(Z^n, X) \to 0$ almost surely. Therefore, also using monotone convergence it holds

$$d(Z^n, X) = \sum_{k\geq 1} \frac{1}{2^k} \mathbb{E}[\|Z^n - X\|_k^* \wedge 1] = \mathbb{E}[\rho(Z^n, X)] \to 0.$$

To prove that $(X^n)_{n\geq 1}$ also converges to X we basically just have to apply the triangle inequality to get

$$\limsup_{n\to\infty} d(X^n, X) \leq \limsup_{n,m\to\infty} \big(d(X^n, Z^m) + d(Z^m, X)\big) = 0.$$

This tells $X^n \xrightarrow{d} X$ and finishes the proof of the completeness of $\mathcal{D}$.

To prove that $\mathcal{D}$ is also a topological vector space we have to show that scalar multiplication and addition are continuous.

Concerning scalar multiplication let $(\lambda_n)_{n\geq 1}$ be a sequence of scalars in $\mathbb{R}$ converging to $\lambda \in \mathbb{R}$. Furthermore, let a sequence $(X_n)_{n\geq 1}$ in $\mathcal{D}$ be given which converges with respect to ucp to $X \in \mathcal{D}$. Since $(\lambda_n)_{n\geq 1}$ is a convergent sequence in $\mathbb{R}$ it is bounded by some constant $M > 0$ and we get

$$d(\lambda_n X_n, \lambda X) = \sum_{k\geq 1} \frac{1}{2^k} \|\lambda_n X_n - \lambda X\|_k^* \leq \underbrace{|\lambda_n|}_{\leq M} \underbrace{d(X_n, X)}_{\to 0} + \underbrace{|\lambda_n - \lambda|}_{\to 0} \underbrace{d(X, 0)}_{\leq 1} \to 0.$$

This shows the continuity of scalar multiplication.

Concerning addition let $(X_n)_{n\geq 1}, (Y_n)_{n\geq 1}$ be two sequences in $\mathcal{D}$ converging to $X, Y \in \mathcal{D}$ respectively. We then get

$$d(X_n + Y_n, X + Y) = \sum_{k\geq 1} \frac{1}{2^k} \|X_n - X + Y_n - Y\|_k^* \leq d(X_n, X) + d(Y_n, Y) \to 0.$$

This proves continuity of the addition.

All in all we have proved that $(\mathcal{D}, d)$ is a complete topological vector space which finishes the proof of the proposition. □

Having the complete topological vectors spaces at hand we are ready to introduce the good integrator property. We generalize it in a straightforward manner from Chapter 2.

Definition 3.14 (Good Integrator)**.** *A càdlàg process X is called good integrator for the Banach space E if the map*

$$J_X : (\mathcal{S}, d) \to (\mathcal{D}, d) : H = \sum_{i=0}^{n-1} H_i \mathbb{1}_{(\tau_i, \tau_{i+1}]} \mapsto (H \cdot X)_{\cdot} = \sum_{i=0}^{n-1} H_i (X_{\tau_{i+1}\wedge \cdot} - X_{\tau_i \wedge \cdot}) \tag{3.7}$$

is continuous. We will also say that X has the good integrator property for the Banach space E in this case.

Remark 3.15. *If a càdlàg process X meets our definition of a good integrator for the Banach space E we have a continuous linear map J_X from a dense subset of the complete topological vector space $(\mathcal{L}, d)$ to the complete topological vector space $(\mathcal{D}, d)$. As in the real-valued case we can extend this map continuously to all càglàd processes $\mathcal{L}$.*

After defining the good integrator property, as in the real-valued case, an equivalent condition for Definition 3.14 is given in the following. There again one only needs to work with convergence in probability. On the other hand continuity has to be checked for a family of maps instead of one map.

Theorem 3.16. *Let X be a càdlàg adapted stochastic process.*

Then X is a good integrator for the Banach space E if and only if for all $t \geq 0$ the map

$$I_{X^t} : (\mathcal{S}_u, \|\cdot\|_\infty) \to \mathbb{L}^0(\Omega; E) : H \mapsto (H \cdot X)_t \tag{3.8}$$

where on $\mathbb{L}^0(\Omega; E)$ we use the metric induced by convergence in probability, is continuous.

Proof. Since the spaces considered are topological vector spaces it is enough to prove continuity at the origin.

We start with proving that the good integrator property implies the continuity of the map I_{X^t} for all $t \geq 0$. For this sake, let $(H^n)_{n\geq 1} \subset \mathcal{S}_u$ such that $\|H^n\|_\infty \to 0$. We also get that $H^n \xrightarrow{d} 0$, since

$$d(H^n, 0) = \sum_{m\geq 1} \frac{1}{2^m}\mathbb{E}[\|H^n\|_m^* \wedge 1] \leq \|H^n\|_\infty \to 0.$$

So, by the assumption it follows $J_X(H^n) \xrightarrow{d} 0$ which in particular translates to $I_{X^t}(H^n) \xrightarrow{\mathbb{P}} 0$.

Let us now prove the converse. For this sake, let $(H^n)_{n\geq 1} \subset \mathcal{S}$ such that $H^n \xrightarrow{d} 0$. Let $c > 0$, $t \geq 0$ and $\epsilon > 0$. By assumption there exists a $\delta > 0$ such that we have the implication

$$\|H^n\|_\infty \leq \delta \implies \mathbb{P}[\|(H^n \cdot X)_t\| > c] < \epsilon. \tag{3.9}$$

Next, for $n \in \mathbb{N}$ define the stopping times

$$\begin{aligned} \tau^n &:= \inf\{s \geq 0 \mid \|H_s^n\| \geq \delta\} \\ \sigma^n &:= \inf\{s \geq 0 \mid \|(H^n \mathbb{1}_{[0,\tau^n]} \cdot X)_s\| > c\}. \end{aligned}$$

Having these stopping times at hand note that on the set $\{\sigma^n \leq t\}$ it holds

$$\{\|(H^n \mathbb{1}_{[0,\tau^n]} \cdot X)\|_t^*\} = \{\|(H^n \mathbb{1}_{[0,\tau^n \wedge \sigma^n]} \cdot X)_t\| > c\}.$$

Also using that $\{\tau^n \leq t\} = \{\|H^n\|_t^* > c\}$ we can start estimating the required expression by

$$\begin{aligned} \mathbb{P}[\|(H^n \cdot X)\|_t^* > c] &\leq \mathbb{P}[\|(H^n \mathbb{1}_{[0,\tau^n]} \cdot X)\|_t^*, \tau^n > t] + \mathbb{P}[\tau^n \leq t] \\ &\leq \mathbb{P}[\sigma^n \leq t, \|(H^n \mathbb{1}_{[0,\tau^n \wedge \sigma^n]} \cdot X)_t\| > c] + \mathbb{P}[\tau^n \leq t] \\ &\leq \mathbb{P}[\|H^n \mathbb{1}_{[0,\tau^n \wedge \sigma^n]} \cdot X)_t\| > c] + \mathbb{P}[\tau^n \leq t] \\ &\overset{(3.9)}{\leq} \epsilon + \mathbb{P}[\tau^n \leq t] = \epsilon + \mathbb{P}[\|H^n\|_t^* > c] \leq 2\epsilon \end{aligned}$$

for large n because for large n the set $\{\|H^n\|_t^* > c\}$ has low probability since $H^n \xrightarrow{d} 0$. This finishes the proof. □

Having defined the good integrator property in a Banach space setting and proven the equivalent condition one has to check for this property, let us look at an example. This example is precisely the same as in the real-valued case.

Example 3.17. *Let A be a process of finite variation. Then we get for $H \in \mathcal{S}_u$*

$$\begin{aligned} \|(H\cdot A)_t\| &\leq \sum_{i=0}^{n-1}\|H_i(A_{\tau_{i+1}\wedge t}-A_{\tau_i\wedge t})\| \\ &\leq \|H\|_\infty \sum_{i=0}^{n-1}|A_{\tau_{i+1}\wedge t}-A_{\tau_i\wedge t}| \leq \|H\|_\infty \mathit{Var}(A)_t. \end{aligned}$$

From this inequality it follows that A is a good integrator for any separable Banach space E.

Of course, the second example in Chapter 2 of the $\mathbb{L}^2-$ martingales does not work here since we do not have some orthogonality in general Banach spaces. Straight forwardly, this only works up to Hilbert spaces.

However, remember that in the real-valued case the crucial part towards a stochastic integral for semimartingales was obtaining via the Burkholder inequality that martingales are good integrators. From this one can also prove the good integrator property for all local martingales and then by summing a local martingale and a finite variation process, again semimartingales are good integrators. We formulate the following Lemma concerning this by using a stopping argument telling that if all martingales are good integrators for a Banach space E, so are all local martingales.

Lemma 3.18. *Suppose all martingales are good integrators for the Banach space E. Then also all local martingales are good integrators for the Banach space E.*

Proof. Let M be a local martingale and (τ^n) a sequence of stopping times such that $\tau^n \leq \tau^{n+1}$, $\tau^n \to \infty$ and for each $n \in \mathbb{N}$ it holds that M^{τ^n} is a martingale. Fix $t \geq 0$ and let $(H^m)_{m\geq 1} \subset \mathcal{S}_u$ such that $\|H^m\|_\infty \to 0$. We then get for $\epsilon > 0$

$$\begin{aligned} \mathbb{P}[\|(H^m\cdot M)_t\| > c] &\leq \mathbb{P}[\|(H^m\cdot M)_t\| > c, \tau^n > t] + \mathbb{P}[\tau^n \leq t] \\ &\leq \mathbb{P}[\|(H^m\cdot M^{\tau^n})_t\| > c] + \mathbb{P}[\tau^n \leq t] \\ &\overset{n \text{ large}}{\leq} \mathbb{P}[\|(H^m\cdot M^{\tau^n})_t\| > c] + \epsilon \overset{m \text{ large}}{\leq} 2\epsilon. \end{aligned}$$

This shows that M is a good integrator by applying Theorem 3.16. □

Having this Lemma and the Burkholder inequality for Banach spaces in mind, it is not a long proof to show the following main result of this section.

Theorem 3.19. *Suppose the Banach space E is a Burkholder space. Then all semimartingales are good integrators for E.*

Proof. The only thing left to prove is that martingales are good integrators for E since by the previous discussions it will then follow that all semimartingales are good integrators.

So, let M be a martingale. Since E is a Burkholder space we know that Theorem 3.9, i.e. the Burkholder Inequality (3.5) holds true. If we take $(H^n) \subset \mathcal{S}_u$ such that $\|H^n\|_\infty \to 0$, the inequality can be applied to get for $t \geq 0,\ c > 0$ and $\epsilon > 0$ that

$$\mathbb{P}[\|(H^n \cdot M)_t\| > c] \leq (10+8)C_E^2 \frac{\|H^n\|_\infty^2}{c} \mathbb{E}[|M_t|] \to 0.$$

Therefore, M is a good integrator for E and the proof is finished. □

Having the theorem we close this section by a final example in which the theorem is applicable.

Example 3.20. *We already know that all martingale type* 2 *spaces are Burkholder spaces. Theorem 3.19 tells that all semimartingales are good integrators for martingale type* 2 *spaces.*

3.3 Good Integrator Property for UMD Spaces

In this last section of the Chapter we consider a separable Banach space E which in addition has the UMD property. It is introduced in Section 8.4 and more details can be found there. The idea of this section and approach was to combine Protter's idea of the good integrator property and the theory of stochastic integrals in a Banach space setting as outlined in [vNVW07]. However, it did not work out as hoped yet, so this part is still open. Nevertheless, the approach and results which were worked out are presented since they themselves are interesting.

We will use the Sections 8.2, 8.3, 8.4 and 8.5 and translate it to this special case. The Hilbert space H in those chapters will simply be $\mathbb{R}$ here.

As a first step we establish a simple Lemma for Banach spaces.

Lemma 3.21. *For a Banach space X the space $\mathcal{L}(\mathbb{R}, X)$ of linear operators from $\mathbb{R}$ to X can be identified with X itself.*

Proof. We define the isomorphism

$$\Psi : \mathcal{L}(\mathbb{R}; X) \to X : \ T \mapsto T(1). \tag{3.10}$$

It is easy to see that this defines an isomorphism since for $x \in \mathbb{R}$, $T(x) = xT(1)$ and therefore

$$\|T\| = \sup_{x\in\mathbb{R}:|x|=1} \|T(x)\| = \max(\|T(1)\|, \|T(-1)\|) = \|T(1)\| = \|\Psi(T)\|.$$

□

As a next step we also see that the space of radonifying operators (see also Section 8.2) simplifies to E itself.

Lemma 3.22. *For a separable Banach space X we have that $\gamma(\mathbb{R}, X) \simeq X$ and the norm agrees with the norm on X itself.*

Proof. In the separable case an operator T is $\gamma-$radonifying for a separable Hilbert space H, i.e. $T \in \gamma(H, X)$ if and only if

$$\mathbb{E}\left[\|\sum_{n\geq 1} \gamma_n T h_n\|^2\right] < \infty$$

for an orthonormal basis $(h_n)_{n\geq 1}$ of H and a sequence of independent standard normal random variables $(\gamma_n)_{n\geq 1}$.

In our case $H = \mathbb{R}$, so for an operator $T \in \mathcal{L}(\mathbb{R}, X)$, which we identify by our previous lemma with an unique $x \in E$, we have for a standard normal random variable γ

$$\mathbb{E}[\|\gamma x\|^2] = \|x\|^2 \mathbb{E}[\gamma^2] = \|x\|^2 < \infty.$$

□

Next, we look how an $H-$cylindrical Brownian motion looks like if $H = \mathbb{R}$. One might expect that we would end up with the usual $\mathbb{R}-$valued Brownian motion which is almost the case.

From the Definitions 8.39 and 8.40 of $H-$isonormal processes and $H-$cylindrical Brownian motion we see that an $\mathbb{R}-$cylindrical Brownian motion is an $\mathbb{L}^2(\mathbb{R}_+; \mathbb{R}) = \mathbb{L}^2(\mathbb{R}_+)-$ isonormal process, i.e. a bounded linear map $W : \mathbb{L}^2(\mathbb{R}_+) \to \mathbb{L}^2(\Omega)$ such that

- $\forall f \in \mathbb{L}^2(\mathbb{R}_+) : Wf$ *is Gaussian,*
- $\forall f_1,\ f_2 \in \mathbb{L}^2(\mathbb{R}_+) : \mathbb{E}[Wf_1 Wf_2] = \langle f_1 , f_2 \rangle_{\mathbb{L}^2(\mathbb{R}_+)} = \int_0^\infty f_1(t) f_2(t)\ dt.$

Similar as in Section 8.3 we can define for $x \in \mathbb{R}$

$$W(t)x := W(\mathbb{1}_{(0,t)}x), \tag{3.11}$$

which leads to a Brownian motion $(W(t)x)_{t\geq 0}$ for every $x \in \mathbb{R}$. It is standard if $|x| = 1$. In the following we will write $W(t)$ instead of $W(t)1$ which is a standard Brownian Motion.

The stochastic integral for deterministic integrands from Section 8.3 for an integrand,

$$\Phi : \mathbb{R}_+ \to E : t \mapsto \sum_{i=0}^{n-1} 1_{(t_i,t_{i+1}]} x_i,$$

where $0 \leq t_0 \leq \ldots \leq t_n$ and $x_0, \ldots x_{n-1} \in E$, is then defined as

$$\int_0^\infty \Phi \; dW = \sum_{i=0}^{n-1} (W(t_{i+1}) - W(t_i)) x_i.$$

Also straight forwardly, we look at the stochastic integral for an elementary adapted process in this situation which looks like

$$\Phi : \mathbb{R}_+ \times \Omega \to E : (t, \omega) \mapsto \sum_{n=1}^{N} \sum_{m=1}^{M} \mathbb{1}_{(t_{n-1},t_n] \times A_{mn}} x_{mn},$$

where $0 \leq t_0 \leq \ldots \leq t_n$, $A_{mn} \in \mathcal{F}_{t_{n-1}}$ for $m = 1, \ldots, M$ and $x_{mn} \in E$ for $n = 1, \ldots, N$ and $m = 1, \ldots, M$. The stochastic integral of Φ w.r.t. W is then

$$\int_0^\infty \Phi \; dW = \sum_{n=1}^{N} \sum_{m=1}^{M} \mathbb{1}_{A_{mn}} (W(t_n) - W(t_{n-1})) x_{mn}.$$

Next, we recall Definition 8.61 in the case of $H = \mathbb{R}$.

Definition 3.23. *$\Phi : \mathbb{R}_+ \times \Omega \to \mathbb{E}$ is $\mathbb{L}^p-$stochastically integrable w.r.t. W if there exists a sequence of finite rank processes Φ_n s.t.*

i) $\Phi_n \to \Phi$ in measure, and

ii) $\exists X \in \mathbb{L}^p(\Omega; E)$ s.t. $\int_0^\infty \Phi_n \; dW \to X$ in $\mathbb{L}^p(\Omega; E)$.

The $\mathbb{L}^p-$stochastic integral of Φ w.r.t. W then is

$$\int_0^\infty \Phi \; dW = \lim_{n\to\infty} \int_0^\infty \Phi_n \; dW \quad \text{in } \mathbb{L}^p(\Omega; E).$$

We now wonder if at least all processes $H \in \mathcal{S}_u$ are $\mathbb{L}^p-$stochastically integrable w.r.t. W. We can give a positive answer to this question.

Lemma 3.24. *Let $H \in \mathcal{S}_u$. Then H is $\mathbb{L}^p-$stochastically integrable w.r.t. W.*

Proof. Write H as

$$H = \sum_{i=0}^{n-1} H_i \mathbb{1}_{(t_i, t_{i+1}]}$$

with $0 \leq t_0 \leq \ldots \leq t_n$ and $H_i \in \mathbb{L}^\infty(\Omega, \mathcal{F}_{t_i}; E)$ for $i = 0, \ldots, n-1$.

Now fix $\epsilon > 0$, then for $i = 0, \ldots, n-1$ and $m \in \mathbb{N}$ we find $h_m^i \in \mathbb{L}^\infty(\Omega, \mathcal{F}_{t_i}; E)$ simple, i.e.

$$h_m^i = \sum_{k=1}^{M_i} \alpha_k^i \mathbb{1}_{A_j^i}$$

with $A_1^i, \ldots, A_{M_i}^i \in \mathcal{F}_{t_i}$ and $\alpha_1^i \ldots, \alpha_{M_i}^i \in E$ such that $\|H_i - h_m^i\|_\infty < \epsilon$. We then define

$$\Phi_m := \sum_{i=0}^{n-1} \mathbb{1}_{(t_i, t_{i+1}]} h_m^i \tag{3.12}$$

which is a finite rank process. With this definition we get

$$\|H - \Phi_m\|_\infty \leq \max_{0 \leq i \leq n-1} \|H_i - h_m^i\|_\infty < \epsilon. \tag{3.13}$$

Next, we define

$$X := (H \cdot X)_{t_n} = \sum_{i=0}^{n-1} H_i(W(t_{i+1}) - W(t_i)),$$

which is an element of $\mathbb{L}^p(\Omega;E)$ since $H_i \in \mathbb{L}^\infty(\Omega;E)$. We then estimate

$$\begin{aligned}
\Big\| \int_0^T \Phi_m \, dW - X \Big\|_{\mathbb{L}^p(\Omega;E)} &= \Big\| \sum_{i=0}^{n-1} (h_m^i - H_i)(W(t_{i+1}) - W(t_i)) \Big\|_{\mathbb{L}^p(\Omega;E)} \\
&\leq \sum_{i=0}^{n-1} \|(h_m^i - H_i)(W(t_{i+1}) - W(t_i))\|_{\mathbb{L}^p(\Omega;E)} \\
&= \sum_{i=0}^{n-1} \left(\mathbb{E}[\|h_m^i - H_i\|^p |W(t_{i-1}) - W(t_i)|^p] \right)^{1/p} \\
&\leq \sum_{i=0}^{n-1} \|h_m^i - H_i\|_{\mathbb{L}^\infty(\Omega;E)} \|W(t_{i+1}) - W(t_i)\|_p \\
&\leq \|\Phi_m - H\|_\infty \sum_{i=0}^{n-1} \|W(t_{i+1}) - W(t_i)\|_p \\
&< \epsilon \sum_{i=0}^{n-1} \|W(t_{i+1}) - W(t_i)\|_p .
\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, this and the estimate 3.13 prove that for the sequence $(\Phi_m)_{m\in\mathbb{N}}$ we get $\Phi_m \to H$ in measure and $\int_0^\infty \, dW \to X$ in $\mathbb{L}^p(\Omega;E)$. Therefore, H is stochastically integrable w.r.t. W and its integral X coincides as expected with the martingale transform $X = (H \cdot W)$. □

Knowing that a strategy $H \in \mathcal{S}_u$ is $\mathbb{L}^p$−stochastically integrable we know by Theorem 8.64 that there exists $R \in \mathbb{L}^p(\Omega; \gamma(\mathbb{L}^2(\mathbb{R}_+), E)$ such that for all $f \in \mathbb{L}^2(\mathbb{R}_+)$, $x^* \in E^*$ we have

$$\langle Rf \, , x^* \rangle = \int_0^\infty \langle H(t) f(t) \, , x^* \rangle \, dt, \quad \textit{in } \mathbb{L}^p(\Omega) \tag{3.14}$$

and

$$\mathbb{E}\left[\Big\| \int_0^\infty H \, dW \Big\|^p \right] \simeq_{p,E} \mathbb{E}\left[\|R\|^p_{\gamma(\mathbb{L}^2(\mathbb{R}_+),E)} \right]. \tag{3.15}$$

Here, $\simeq_{p,E}$ means that there exist two constants dependent only on p and E such the the left expression can be estimated by the right expression when multiplying with those constants.

In the following we try to find this R for H. We guess and define

$$R : \Omega \to \gamma(\mathbb{L}^2(\mathbb{R}_+), E) : \omega \mapsto \left(f \mapsto \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} f(t) \, dt H_i(\omega) \right). \tag{3.16}$$

The first question is if it is well defined (up to nullsets). For this sake, let ω be in a set where we have $\|H_i(\omega)\| < \infty$ for $i = 0, \dots, n-1$. Looking at how we can also write $R(\omega)$ namely like

$$R(\omega) : \mathbb{L}^2(\mathbb{R}_+) \to E : f \mapsto \sum_{i=0}^{n-1} \langle f \, , \mathbb{1}_{(t_i,t_{i+1}]} \rangle_{L^2(\mathbb{R}_+)} H_i(\omega) \tag{3.17}$$

we see that $R(\omega)$ is actually a finite rank operator (see Section 8.2). Then by Lemma 8.33 we have that $R(\omega) \in \gamma(\mathbb{L}^2(\mathbb{R}_+), E)$ and

$$\|R(\omega)\|^2_{\gamma(\mathbb{L}^2(\mathbb{R}_+),E)} = \mathbb{E}\left[\Big\| \sum_{i=0}^{n-1} \gamma_i H_i(\omega) \Big\|^2 \right], \tag{3.18}$$

where $\gamma_0, \dots, \gamma_{n-1}$ are independent standard normal random variables.

We now have that R is well defined and can use the last identity to show that R belongs to $\mathbb{L}^2(\Omega; \gamma(\mathbb{L}^2(\mathbb{R}_+), E))$. This we can see via

$$\begin{aligned}
\|R\|^2_{\mathbb{L}^2(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+),E))} &= \mathbb{E}[\|R\|^2_{\gamma(\mathbb{L}^2(\mathbb{R}_+),E)}] \\
&= \int_\Omega \mathbb{E}\left[\Big\| \sum_{i=0}^{n-1} \gamma_i H_i(\omega) \Big\|^2 \right] \, d\mathbb{P}(\omega) \\
&\leq \mathbb{E}\left[\int_\Omega \Big(\sum_{i=0}^{n-1} \|\gamma_i H_i(\omega)\| \Big)^2 \, d\mathbb{P}(\omega) \right] \\
&\leq \mathbb{E}\left[\int_\Omega \Big(\max_{0\leq j\leq n-1} \|H_j(\omega)\| \Big| \sum_{i=0}^{n-1} \gamma_i \Big| \Big)^2 \, d\mathbb{P}(\omega) \right] \\
&\leq \|H\|^2_\infty \mathbb{E}\left[\Big(\sum_{i=0}^{n-1} |\gamma_i| \Big)^2 \right] \leq \|H\|^2_\infty n^2
\end{aligned}$$

To check if R is the one we were looking for we have to check if it fulfills identity (3.14). For this sake let $f \in \mathbb{L}^2(\mathbb{R}_+)$ and $x^* \in \mathbb{E}^*$. We get

$$\begin{aligned}
\int_0^\infty \langle H(t) f(t) \, , x^* \rangle \, dt &= \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \langle H_i f(t) \, , x^* \rangle \, dt \\
&= \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} x^*(H_i f(t)) \, dt = \sum_{i=0}^{n-1} x^*(H_i) \int_{t_i}^{t_{i+1}} f(t) \, dt
\end{aligned}$$

and by the definition of R we have

$$\langle Rf\,,x^*\rangle = \langle\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} f(t)\ dtH_i\,,x^*\rangle = \sum_{i=0}^{n-1} x^*(H_i)\int_{t_i}^{t_{i+1}} f(t)\ dt.$$

This means the two sides coincide and since R is uniquely determined we have guessed it right.

Now, we can ask the question whether we get that W is a good integrator for E. This accounts to proving continuity for the map J_W or I_{X^t} for every $t \geq 0$. The estimate (3.15) gives us a constant $K_{2,E}$ depending on E such that

$$\mathbb{E}\left[\|\int_0^\infty H\ dW\|^2\right] \leq K_{2,E}^2\mathbb{E}\left[\|R\|^2_{\gamma(\mathbb{L}^2(\mathbb{R}_+),E)}\right] \leq K_{2,E}^2\|H\|_\infty^2 n^2. \tag{3.19}$$

This estimate is unfortunately not very promising since by taking its root we see that the norm we would like to estimate grows with n where n accounts to the number of jumps of H. If we take now any sequence $(H^m)_{m\in\mathbb{N}} \subset \mathcal{S}_u$ such that $\|H^m\|_\infty \to 0$ we could run into troubles, because probably the number of jumps tends to infinity and if it does, maybe too fast.

This is why this section still remains open.

4 Banach Space-valued Integrators and Real-valued Integrands

In this very short chapter compared to Chapter 3 we exchange the role of the integrator and integrand. This means that now the integrator will take values in a Banach space while the integrand will be real-valued. An inequality similar to the already seen Burkholder inequality in this setting can be achieved quite fast and is an application of already known theorems.

In the following, consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ being a right-continuous filtration.

The Banach space under consideration, named E, is assumed to be separable and a Burkholder type inequality will be given for that space.

As a first step towards this we state [Pis11, Proposition 8.10].

Proposition 4.1. *Let E be an UMD_p space and $(H_n)_{n\geq 0}$ be adapted to a filtration $(\mathcal{A}_n)_{n\geq 0}$. Let furthermore $(X_n)_{n\geq 0}$ be an $E-$valued martingale. Then we have for some constant $C > 0$ the inequality*

$$\sup_{n\geq 0} \Big\| H_0 X_0 + \sum_{i=1}^{n} H_{i-1}(X_i - X_{i-1}) \Big\|_{\mathbb{L}^p(\Omega;E)} \leq C \sup_{n\geq 0} \|H_n\|_\infty \sup_{n\geq 0} \|X_n\|_{\mathbb{L}^p(\Omega;E)}. \qquad (4.1)$$

Already in inequality (4.1) one can see a good bound for the discrete stochastic integral of H with respect to the Banach space-valued martingale X.

We also state [Pis11, Corollary 8.14] which will be even more useful.

Corollary 4.2. *Let E be UMD_p for some $p \in (1, \infty)$. Then there exists a constant $C > 0$ such that for all martingales $(X_n)_{n\geq 0}$ which are bounded in $\mathbb{L}^1(\Omega; E)$ and $(H_n)_{n\geq 0}$ adapted scalar random variables with $\|H_n\|_{\mathbb{L}^\infty} \leq 1$, we have*

$$\sup_{\lambda>0} \lambda \mathbb{P}\left[\sup_{n\geq 0} \Big\| \sum_{i=0}^{n-1} H_i (X_{i+1} - X_i) \Big\| > \lambda \right] \leq C \sup_{n\geq 0} \|X_n\|_{\mathbb{L}^1(\Omega;E)}.$$

Corollary 4.2 is exactly the Burkholder inequality in the case of $UMD-$ valued martingales which are bounded in $\mathbb{L}^1(\Omega; E)$. Actually we can even follow from the corollary already the following theorem.

Theorem 4.3. *Let E be an UMD_p space for some $p \in (1, \infty)$. Define*

$$\mathcal{S}_u := \left\{ H = H_0 \mathbb{1}_0 + \sum_{i=0}^{n-1} H_i \mathbb{1}_{(t_i, t_{i+1}]} \mid H_i \in \mathbb{L}^\infty,\ 0 = t_0 \leq \ldots \leq t_n \right\}.$$

There exists a constant $C_E > 0$ such that for all $H \in \mathcal{S}_u$ and all $E-$valued martingales $(X_t)_{t \geq 0}$ we have

$$\lambda \mathbb{P}\left[(\|(H \cdot X)\|)^*_t > \lambda \right] \leq C_E \max_{0 \leq i \leq n-1} \|H_i\|_{\mathbb{L}^\infty} \mathbb{E}[\|X_t\|], \quad \forall \lambda > 0,\ \forall t \geq 0.$$

The constant C_E is only dependent on the space E, so the constant is the same for all martingales.

Proof. Let $t > 0$ and a strategy $H \in \mathcal{S}_u$ be arbitrary such that $t_n \leq t$. By looking at

$$\frac{H}{\max_{0 \leq i \leq n-1} \|H_i\|_{\mathbb{L}^\infty}}$$

instead of H we can assume $\|H_i\|_{\mathbb{L}^\infty} \leq 1$ for all $0 \leq i \leq n-1$.

Instead of X we can consider the discrete time martingale with finite time horizon $(X_{t_i})_{i=0}^n$.

For some $\lambda > 0$ Corollary 4.2 can be applied to the finite sequence $(H_i)_{i=0}^{n-1}$ and our discrete finite time martingale and translates precisely into

$$\lambda \mathbb{P}\left[(\|(H \cdot X)\|)^*_{t_n} > \lambda \right] \leq C_E \mathbb{E}[\|X_{t_n}\|] \leq C_E \mathbb{E}[\|X_t\|]$$

with $C_E = C$ from the corollary. This finishes the prove where we note that the second inequality holds since if X is a martingale, then $\|X\|$ is a submartingale, therefore its expectation is non-decreasing. □

5 Nikisin-Yan

This and the next chapter are devoted to two main results which will play a crucial role in Chapter 7. There the two main results of this thesis besides the ones in Chapter 3 will be presented. We start with the paper of Jia-an Yan, see [Yan80], and look at two theorems of it.

For the rest of this chapter fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In the following we cite the statement and rewrite the proof of [Yan80, Theorem 2].

Theorem 5.1 (Yan)**.** *Let $K \subset \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ be convex and $0 \in K$. Then the following are equivalent:*

i) $\forall \eta \in \mathbb{L}^1_{\geq 0},\ \eta \neq 0\ \exists c > 0\ s.t.\ c\eta \notin \overline{K - \mathbb{L}^\infty_{\geq 0}}^{\mathbb{L}^1}$

ii) $\forall A \in \mathcal{F},\ \mathbb{P}[A] > 0\ \exists c > 0\ s.t.\ c\mathbb{1}_A \notin \overline{K - \mathbb{L}^\infty_{\geq 0}}^{\mathbb{L}^1}$

iii) $\exists$ *bounded* $Z > 0,\ \mathbb{P} - a.s.\ s.t.\ \sup_{\zeta \in K} \mathbb{E}[\zeta Z] < \infty.$

Proof. $i) \implies ii)$: This one should be clear since for every $A \in \mathcal{F}$ with $\mathbb{P}[A] > 0$, the function $\mathbb{1}_A$ is just a specific non-vanishing function in $\mathbb{L}^1_{\geq 0}$.

$iii) \implies i)$: Assume by contradiction that there exists some $\eta \in \mathbb{L}_{\geq 0}$ with $\eta \neq 0$ such that

$$\forall n \in \mathbb{N}: \ n\eta \in \overline{K - \mathbb{L}^\infty_{\geq 0}}^{\mathbb{L}^1}.$$

We then get

$$n\eta = \zeta_n - \xi_n + \delta_n, \quad \|\delta_n\|_{\mathbb{L}^1} \leq \frac{1}{n}$$

for some $\zeta_n \in K$ and $\xi_n \in \mathbb{L}^\infty_{\geq 0}$. Take now the $Z > 0$ which exists by assumption of $iii)$ and observe

$$\mathbb{E}[Z\zeta_n] = \mathbb{E}[Zn\eta] + \mathbb{E}[Z\xi_n] - \mathbb{E}[Z\delta_n] \geq n \underbrace{\mathbb{E}[Z\eta]}_{>0} - \frac{1}{n} \to \infty.$$

This contradicts the assumption $\sup_{\zeta \in K} \mathbb{E}[\zeta Z] < \infty$.

ii) $\implies$ *iii)* : Let $A \in \mathcal{F}$, $\mathbb{P}[A] > 0$. Then there exists some $c > 0$ such that $c\mathbb{1}_A \notin \overline{K - \mathbb{L}^\infty_{\geq 0}}^{\mathbb{L}^1}$. Now we notice that $\overline{K - \mathbb{L}^\infty_{\geq 0}}^{\mathbb{L}^1}$ is a closed and convex set and the topological dual satisfies $(\mathbb{L}^1)^* = \mathbb{L}^\infty$ by the Riesz representation theorem. So we can apply the Hahn-Banach theorem (see Theorem 5.2) to get a bounded Y such that

$$\sup_{\zeta \in K,\ \eta \in \mathbb{L}^\infty_{\geq 0}} \mathbb{E}[Y(\zeta - \eta)] < c\mathbb{E}[Y \mathbb{1}_A],$$

where c is a new constant adapted by the constants from the Hahn-Banach theorem.

Replacing now η by $n\eta$, the inequality $-n\mathbb{E}[Y\eta] \leq c\mathbb{E}[Y\mathbb{1}_A]$ must hold for all $\eta \in \mathbb{L}^\infty_{\geq 0}$ when we choose $\zeta = 0 \in K$. This would be contradicted if Y would be negative on a set with positive probability. Therefore, we conclude $Y \geq 0$, $\mathbb{P} - a.s.$

Next we set

$$\mathcal{H} := \{Y \in \mathbb{L}^\infty_{\geq 0} \mid \sup_{\zeta \in K} \mathbb{E}[Y\zeta] < \infty\},$$

which is non-empty by what we have just seen. We also set

$$C := \{\{Y = 0\} \mid Y \in \mathcal{H}\}$$

and will show that it is closed under countable intersections. For this sake consider $\{Y_n\}_{n \geq 1} \subset \mathcal{H}$ and notice that for $(b_n)_{n\geq 1}$ such that $b_n > 0$ we have

$$\bigcap_{n \geq 1} \{Y_n = 0\} = \left\{ \sum_{n \geq 1} b_n Y_n = 0 \right\}.$$

Now define $c_n := \sup_{\zeta \in K} \mathbb{E}[Y_n \zeta]$ and $d_n := \|Y_n\|_{\mathbb{L}^\infty}$ for $n \in \mathbb{N}$ and choose $(b_n)_{n \geq 1}$ positive real numbers such that

$$\sum_{n \geq 1} b_n c_n < \infty \text{ and } \sum_{n \geq 1} b_n d_n < \infty.$$

Then set $Y := \sum_{n \geq 1} b_n Y_n$ in order to prove the claim that C is closed under countable intersections.

By this we have the existence of $Y \in \mathcal{H}$ with $\mathbb{P}[Y = 0] = \inf_{S \in C} \mathbb{P}[S]$. Set $A := \{Y = 0\}$ and assume $\mathbb{P}[A] > 0$.

By what we have shown we can separate $c\mathbb{1}_A$ from $\overline{K - \mathbb{L}^\infty_{\geq 0}}^{\mathbb{L}^1}$ by some bounded $\tilde{Y} \geq 0$ and some $c > 0$. In particular since $0 \in K$ it holds that

$$0 < c\mathbb{E}[\tilde{Y}\mathbb{1}_A] = c\mathbb{E}[\tilde{Y}\mathbb{1}_{\{Y=0\}}].$$

From this conclude $Y + \tilde{Y} \in \mathcal{H}$ and $\{Y + \tilde{Y} = 0\} = \{Y = 0\} \cap \{\tilde{Y} = 0\}$. If $\mathbb{P}[Y = 0] > 0$ would hold then also

$$\mathbb{P}[\{\tilde{Y} = 0\} \cap \{Y = 0\}] < \mathbb{P}[Y = 0].$$

This however contradicts the minimality of Y.

Thus we conclude $\mathbb{P}[Y = 0] = 0 \implies \mathbb{P}[Y > 0] = 1$ to finish the proof. □

For completeness we state the formulation of the geometric Hahn-Banach theorem used in the proof of the previous theorem. This is point *ii)* in [BKW17, Satz 5.2.5] which we reformulate in the next Theorem.

Theorem 5.2 (geometric Hahn-Banach)**.** *Let F be a topological vector space which is locally convex. Take $A, B \subset F$ two disjoint, non-empty and convex sets. If A is compact and B closed, there exists some $f \in F^*$, $\gamma_1, \gamma_2 \in \mathbb{R}$ such that*

$$f(x) \leq \gamma_1 < \gamma_2 \leq f(y), \quad \forall x \in A, y \in B.$$

Remark 5.3. *In the proof of Theorem 5.1 the set A of the Hahn-Banach theorem corresponds to $\{c\mathbb{1}_A\}$ and the set B to $\overline{K - \mathbb{L}^\infty_{\geq 0}}^{\mathbb{L}^1}$. The element $f \in \mathcal{F}^*$ corresponds to Y via the Riesz-representation theorem through the operator*

$$T : (\mathbb{L}^1)^* \to \mathbb{L}^\infty : f \mapsto (g \mapsto \mathbb{E}[fg]) \, .$$

Having looked at the second theorem of Yan's paper we turn our attention to the first one in it. We cite the formulation of the theorem in the following which to be exact is [Yan80, Theorem 1]. After that we will reformulate a proof of the statement which is an application of Theorem 2 in Yan's paper, or Theorem 5.1 here.

Theorem 5.4 (Nikisin-Yan)**.** *Let $K \subset \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ be convex. Suppose*

$$\forall \epsilon > 0 \; \exists c > 0 \; s.t. \; \mathbb{P}[\zeta > c] < \epsilon, \; \forall \zeta \in K.$$

Then there exists a $Z \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P})$ with $Z > 0$, $\mathbb{P} - a.s.$ and $\sup_{\zeta \in K} \mathbb{E}[\zeta Z] < \infty$.

Proof. By picking an element $k \in K$ and looking at $K - k \subset \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ which is still convex and contains 0 in addition, we can assume without loss of generality that $0 \in K$.

Let us show that the assumption *ii)* of Theorem 5.1 is satisfied. For this sake take some $A \in \mathcal{F}$ with $\mathbb{P}[A] > 0$. Define $\epsilon := \mathbb{P}[A]/2$, then there exists some constant $c > 0$ such that

$$\mathbb{P}[\zeta > c/2] < \epsilon, \; \forall \zeta \in K.$$

Define $X := 2c\mathbb{1}_A$ and let $\zeta \in K$ be arbitrary. Now let us claim it holds

$$A \subset \{X > \zeta + c/2\} \cup \{\zeta > c/2\}.$$

To prove this claim let $\omega \in A$ be arbitrary. Let us first assume $\zeta(\omega) \leq c/2$. Then it holds $\zeta(\omega) + c/2 < X(\omega)$ and therefore $\omega \in \{X > \zeta + c/2\}$. In the other case when $\zeta(\omega) > c/2$ it is immediate that ω belongs to $\{\zeta > c/2\}$. This proves the claim.

Using the claim we get

$$\mathbb{P}[A] \leq \mathbb{P}[X > \zeta + c/2] + \mathbb{P}[\zeta > c/2] \leq \mathbb{P}[X > \zeta + c/2] + \frac{\mathbb{P}[A]}{2},$$

leading to

$$\mathbb{P}[X > \zeta + c/2] \geq \frac{\mathbb{P}[A]}{2}.$$

Since ζ was arbitrary we have

$$\mathbb{P}[2c\mathbb{1}_A - \zeta > c/2] \geq \frac{\mathbb{P}[A]}{2} > 0, \ \forall \zeta \in K.$$

This tells that $2c\mathbb{1}_A$ cannot be approximated in $\mathbb{L}^1$ by any function $\zeta \in K$ minus a nonnegative bounded function, which finishes the proof. □

Let us close this chapter by a remark concerning how the Nikisin-Yan theorem will be used later on.

Remark 5.5. *The Nikisin-Yan thereom tells that if one has a convex subset of* $\mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ *which is bounded in probability from above there exists an equivalent measure* $\mathbb{Q}$ *for which the expectation of random variables in* K *is uniformly bounded from above. To get the equivalent measure one just takes* Z *one gets from the theorem as density, i.e. for* $A \in \mathcal{F}$ *set*

$$\mathbb{Q}[A] = \frac{1}{\mathbb{E}[Z]}\mathbb{E}[\mathbb{1}_A Z].$$

6 Girsanov-Meyer

In this short chapter we revisit and prove another result which will be essential for the results in the next chapter. It is the Girsanov-Meyer theorem which is well known in stochastic analysis, however we still want to give a proof here for completeness. It allows one to change the measure to an equivalent one while still keeping semimartingale properties and also shows the semimartingale's precise decomposition into a local martingale and a finite variation process. It is remarkable that its proof is mainly an application of the definition of quadratic variation (see also Chapter 2). This is one of many examples in stochastic analysis showing the power of this definition.

As usual we consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where $\mathbb{F}$ is a right-continuous filtration.

Before turning the attention to the Girsanov-Meyer theorem two additional lemmas are needed. The first one tells how the conditional expectation after a change of measure looks like compared to the previous measure.

Lemma 6.1. *Let $\mathbb{Q}$ be an equivalent probability measure to $\mathbb{P}$ and denote its density by ξ, i.e. $\xi = \frac{d\mathbb{Q}}{d\mathbb{P}}$.*

Then, for any random variable X integrable w.r.t. $\mathbb{P}$ and $\mathbb{Q}$, the conditional expectation with respect to a sub $\sigma-$ algebra $\mathcal{G} \subset \mathcal{F}$ is given by

$$\mathbb{E}_{\mathbb{Q}}[X \mid \mathcal{G}] = \frac{\mathbb{E}[X\xi \mid \mathcal{G}]}{\mathbb{E}[\xi \mid \mathcal{G}]}. \tag{6.1}$$

Proof. The right hand side of (6.1) is $\mathcal{G}-$ measurable and we have

$$\mathbb{E}\left[\xi \frac{\mathbb{E}[X\xi \mid \mathcal{G}]}{\mathbb{E}[\xi \mid \mathcal{G}]} \mid \mathcal{G}\right] = \mathbb{E}[X\xi \mid \mathcal{G}].$$

Next, take some $A \in \mathcal{G}$ arbitrarily and look at

$$\mathbb{E}_{\mathbb{Q}}\left[\mathbb{1}_A \frac{\mathbb{E}[X\xi \mid \mathcal{G}]}{\mathbb{E}[\xi \mid \mathcal{G}]}\right] = \mathbb{E}\left[\mathbb{1}_A \frac{\mathbb{E}[X\xi \mid \mathcal{G}]}{\mathbb{E}[\xi \mid \mathcal{G}]}\xi\right] = \mathbb{E}[\mathbb{1}_A \xi X] = \mathbb{E}_{\mathbb{Q}}[\mathbb{1}_A X].$$

By the definition of conditional expectation this equality exactly corresponds to

$$\mathbb{E}_{\mathbb{Q}}[X \mid \mathcal{G}] = \frac{\mathbb{E}[X\xi \mid \mathcal{G}]}{\mathbb{E}[\xi \mid \mathcal{G}]},$$

which proves the lemma. □

The second lemma given before the Girsanov-Meyer theorem concerns martingales and measure changes. It characterizes exactly how a martingale after a change of measure looks like. One just has to multiply by the density.

Lemma 6.2. *Let $\mathbb{Q}$ be an equivalent probability measure to $\mathbb{P}$. Denote the density process by Z, that is*

$$Z_t := \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_t\right].$$

Then a stochastic process M is a $\mathbb{Q}-$ martingale if and only if MZ is a $\mathbb{P}-$ martingale.

Proof. First note that MZ is adapted if and only if M is adapted since Z is adapted by the definition of the conditional expectation. Furthermore,

$$\mathbb{E}_{\mathbb{Q}}[|M_t|] = \mathbb{E}[|M_t|Z_t] = \mathbb{E}[|M_t Z_t|],$$

so M is $\mathbb{Q}-$ integrable if and only if MZ is $\mathbb{P}-$ integrable. Apply the previous Lemma to get

$$\mathbb{E}_{\mathbb{Q}}[M_t \mid \mathcal{F}_s] = \frac{\mathbb{E}[M_t \frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_s]}{\mathbb{E}[\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_s]} = \frac{1}{Z_s}\mathbb{E}[M_t Z_t \mid \mathcal{F}_s]. \tag{6.2}$$

This equality can be reordered into

$$Z_s \mathbb{E}_{\mathbb{Q}}[M_t \mid \mathcal{F}_s] = \mathbb{E}[M_t Z_t \mid \mathcal{F}_s]. \tag{6.3}$$

If now M is a $\mathbb{Q}-$martingale then

$$\mathbb{E}[M_t Z_t \mid \mathcal{F}_s] \overset{(6.3)}{=} Z_s \mathbb{E}_{\mathbb{Q}}[M_t \mid \mathcal{F}_s] = Z_s M_s,$$

which shows that MZ is a $\mathbb{P}-$martingale.

If conversely MZ is a $\mathbb{P}-$martingale, we have

$$\mathbb{E}_{\mathbb{Q}}[M_t \mid \mathcal{F}_s] \overset{(6.2)}{=} \frac{1}{Z_s}\mathbb{E}[M_t Z_t \mid \mathcal{F}_s] = M_s.$$

This shows that M is a $\mathbb{Q}-$ martingale and the proof is finished. □

By stopping arguments this Lemma can also be extended to local martingales.

With all this previous work we are in the position to give a proof of the Girsanov-Meyer theorem.

Theorem 6.3 (Girsanov-Meyer)**.** *Let X be a semimartingale under $\mathbb{P}$ and decompose it into $X = M + A$ with M being a local $\mathbb{P}-$ martingale and A a process of finite variation. Let $\mathbb{Q}$ be an equivalent probability measure with respect to $\mathbb{P}$.*

Then X is also a semimartingale under $\mathbb{Q}$ and has the decomposition $X = N + L$ with

$$L_t = M_t - \int_0^t \frac{1}{Z_s}\, d[Z, M]_s,$$

and $N_t = X_t - L_t$.

Remark 6.4. *As a reminder, the integral on the right-hand side denotes the pathwise Lebesgue-Stieltjes integral with integrator a finite variation process.*

Proof. We have that M and Z are $\mathbb{P}-$ local martingales. Therefore also $Z_- \cdot M + M_- \cdot Z$ is a local $\mathbb{P}-$ martingale. For this fact we can refer to [Low10]. Applying integration by parts which is in this case just the definition of quadratic covariation, we see that also

$$ZM - [Z, M] = Z_- \cdot M + M_- \cdot Z$$

is a $\mathbb{P}-$ local martingale. Apply Lemma 6.2 to get that

$$M - \frac{1}{Z}[Z, M]$$

is a $\mathbb{Q}-$ local martingale.

Again, we just look at the definition of quadratic variation to get

$$\frac{1}{Z}[Z, M] = \frac{1}{Z_-} \cdot [Z, M] + [Z, M]_- \cdot \frac{1}{Z} + \left[\frac{1}{Z}, [Z, M]\right]. \tag{6.4}$$

For a process càdlàg Y we denote the jump at time t by $\Delta Y_t := Y_t - Y_{t-}$ and with this we calculate the last term in the right-hand side of (6.4) to get

$$\left[\frac{1}{Z}, [Z, M]\right] = \sum_{0<s\leq t} \Delta\big(\frac{1}{Z_s}\big)\Delta[Z, M]_s.$$

It also holds that

$$\frac{1}{Z} \cdot [Z, M] - \frac{1}{Z_-} \cdot [Z, M] = \sum_{0<s\leq t} \Delta\big(\frac{1}{Z_s}\big)\Delta[Z, M]_s.$$

We plug this in into (6.4) to get

$$\frac{1}{Z}[Z,M] = \frac{1}{Z} \cdot [Z,M] + [Z,M]_- \cdot \frac{1}{Z},$$

where the last part is a $\mathbb{Q}-$ local martingale. Therefore by reordering also

$$\frac{1}{Z}[Z,M] - \frac{1}{Z} \cdot [Z,M]$$

is a $\mathbb{Q}-$ local martingale. Finally also

$$M - \frac{1}{Z}[M,Z]$$

is a $\mathbb{Q}-$ local martingale, so by adding up the last two lines we arrive at the point that

$$M - \frac{1}{Z} \cdot [Z,M]$$

is a $\mathbb{Q}-$ local martingale. All in all this tells that X can be written as as

$$X = \left(M - \frac{1}{Z} \cdot [Z,M]\right) + \left(\frac{1}{Z} \cdot [Z,M] + A\right),$$

where the term in the left brackets is a $\mathbb{Q}-$ local martingale and the term in the right brackets a process of finite variation. This proves the statement of the theorem. □

7 Bichteler-Dellacherie and Doob-Meyer, Generalized

In this Chapter two of the main results are presented. One is the famous Bichteler-Dellacherie theorem and the second one the well known Doob-Meyer decomposition. In both we try to drop some path regularity, to be precise we drop the assumption of càdlàg paths.

The ideas concerning the Bichteler-Dellacherie theorem try to generalize a proof presented by Christophe Stricker in [Str84]. For the proof the two previous chapters in which we proved the Nikisin-Yan and the Girsanov-Meyer theorem will be crucial.

Since we drop the càdlàg assumption for paths in that theorem one has to pay attention with stopping times because in general stopped processes will not be measurable with respect to any stopping time. However, we work with stopping times taking at most countably many values. There we do not have such problems. We state the following lemma for this fact, cited from [Sch18, Lemma 3.41]. We also fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where we assume $\mathbb{F}$ to be right-continuous in the following.

Lemma 7.1. *Let $(S, \mathcal{S})$ be a measurable space, $X : T \times \Omega \to S$ a stochastic process and $\tau : \Omega \to T$ an $\mathbb{F}-$stopping time.*

Then $X_\tau : \Omega \to S$, defined by $X_\tau(\omega) := X_{\tau(\omega)}(\omega)$ for every $\omega \in \Omega$, is $\mathcal{F}_\tau-$measurable under each of these conditions

a) $\tau(\Omega) \subset T$ is countable and X is $\mathbb{F}$adapted;

b) X is $\mathbb{F}-$ progressive.

Having this Lemma at hand we are almost at the point to formulate our version of the Bichteler-Dellacherie theorem. We still need an important theorem concerning supermartingales. For those we can win a bit of regularity by changing to a version of the supermartingale. This helps us to get from an countable index set to an uncountable one. For this sake we cite [Low09, Theorem 4]. A proof of this statement can be found there.

Theorem 7.2. *Let X be a martingale, submartingale or supermartingale. Then it has a version Y which has left and right limits everywhere such that there is a countable set $S \subset \mathbb{R}_+$ for which Y_t is right-continuous at every $t \notin S$.*

To use this theorem we need to connect the process we will be looking at with supermartingales. For this sake we cite the following theorem by Rao in [Low12, Theorem 1], where a proof can be found as well.

Theorem 7.3 (Rao). *A process X is a quasimartingale if and only if it decomposes as*

$$X = Y - Z$$

for supermartingales Y and Z.

Remark 7.4. *For completeness we give a definition of a quasimartingale. Given an integrable process X it is called a quasimartingale if its mean variation is finite for each t on $[0, t]$, i.e. if*

$$m\text{-}var(X)_t := \sup_{0 \le t_1 \le \ldots \le t_n \le t,\ n \in \mathbb{N}} \mathbb{E}\left[\sum_{i=1}^{n-1} |\mathbb{E}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}]|\right] < \infty, \quad \forall t \ge 0.$$

Before however, we need another Lemma concerning weak convergence in $\mathbb{L}^2$ which will also be important in the proof. We will in the following use the arrow $\rightharpoonup$ to denote weak convergence.

Lemma 7.5. *Let $M_1^n,\ M_1 \in \mathbb{L}^2$. Set $M_t^n := \mathbb{E}[M_1^n \mid \mathcal{F}_t]$ and $M_t := \mathbb{E}[M_1 \mid \mathcal{F}_t]$ for some $t < 1$. Assume $M_1^n \rightharpoonup M_1$ in $\mathbb{L}^2$. Then we have*

$$M_t^n \rightharpoonup M_t \text{ in } \mathbb{L}^2.$$

Proof. Let $Y \in \mathbb{L}^2$ and note that also $\mathbb{E}[Y \mid \mathcal{F}_t] \in \mathbb{L}^2$ by Jensen's inequality. We can use the assumption of weak convergence to get

$$\mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_t] M_1^n] \to \mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_t] M_1]. \tag{7.1}$$

For the left side of this expression one can use to the tower property of conditional expectation to rewrite it as

$$\mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_t] M_1^n] = \mathbb{E}[\mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_t] M_1^n \mid \mathcal{F}_t]] = \mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_t]\mathbb{E}[M_1^n \mid \mathcal{F}_t]] = \mathbb{E}[Y M_t^n].$$

The right term of (7.1) can be rewritten as

$$\begin{aligned}\mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_t] M_1] = \mathbb{E}[\mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_t] M_1 \mid \mathcal{F}_t]] &= \mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_t] M_t] = \\ &= \mathbb{E}[\mathbb{E}[Y M_t \mid \mathcal{F}_t]] = \mathbb{E}[Y M_t].\end{aligned}$$

Knowing how both sides of (7.1) can be rewritten we deduce the convergence of

$$\mathbb{E}[Y M_t^n] \to \mathbb{E}[Y M_t].$$

Since Y was arbitrary it holds $M_t^n \rightharpoonup M_t$ in $\mathbb{L}^2$. □

Now we are ready to formulate our theorem. Note that for a stochastic process with càdlàg paths the assumption of the boundedness in probability corresponds to being a good integrator. In the proof the goal of decomposing X into a local martingale and a finite variation process is achieved by changing the measure via the Nikisin-Yan theorem repeatedly until we can decompose X as desired for an equivalent measure. Then we can change back to the original measure by the Girsanov-Meyer theorem while still having a decomposition for X at hand on a level of versions.

Theorem 7.6. *Let X be a stochastic process such that the convex set*

$$\{(H \cdot X)_t \mid \|H\|_\infty \leq 1,\ H \in \mathcal{S}_u \text{ jumping at deterministic times}\}$$

is bounded in probability for all $t \geq 0$. Then there exists a version $\tilde{X}$ of X, a càdlàg local martingale M and a process of finite variation A such that $\tilde{X} = M + A$.

Written out this means for every $t \geq 0$ it holds

$$X_t = M_t + A_t,\ a.s.$$

Before we start with the proof let us look at a quick remark concerning the proof's strategy.

Remark 7.7. *The goal is to bound the discrete stochastic integral of uniformly by 1 bounded simple integrals against X in $\mathbb{L}^2$ with respect to some measure $\mathbb{Q}$. For this sake, step by step we bound more and more sets in probability and apply Nikisin-Yan several times to change the measure. On the way we also show that X has finite mean variation with respect to that changed measure, i.e. it is a quasimartingale. Once we achieve all that we apply Theorem 7.3 and Theorem 7.2 to change to a version on which we have càdlàg almost everywhere. We are only left with a countable index set on which we will define a discrete $\mathbb{L}^2$ martingale, use weak compactness of $\mathbb{L}^2-$balls and achieve a decomposition on this countable index set into a martingale and a process. For that process we first show finite variation on that countable index set and generalize this to the full index set by using the càdlàg properties we got by switching to a version.*

Proof. First fix $t = 1$ and prove the statement on $[0, 1]$. Then concatenate time intervals to prove it for $t \in [0, \infty)$. Also by just subtracting assume without loss of generality $X_0 = 0$.

As a start we claim that the (not necessarily convex) set

$$\{|(H \cdot X)|_1^* \mid \|H\|_\infty \leq 1 \text{ jumping at deterministic times}\} \tag{7.2}$$

is bounded in probability.

To prove this assume by contradiction that there exists some $\epsilon > 0$ and a sequence $(H_n)_{n\geq 1} \subset \mathcal{S}_u$ with $\|H^n\|_\infty \leq 1$ jumping at deterministic times numbered as $t_1^n, \dots, t_{m_n}^n$ such that

$$\mathbb{P}[\ |(H^n \cdot X)|_1^* \geq n] \geq \epsilon.$$

Defining stopping times like

$$\tau^n := \min\{t_i^n \mid |(H \cdot X)_{t_i^n}| \geq n,\ i = 1, \dots, m_n\},$$

we get for process stopped at τ^n

$$\mathbb{P}[|(H^n \cdot X)_1 \,|^{\tau^n} \geq n] = \mathbb{P}[|(H^n \mathbb{1}_{[0,\tau^n]} \cdot X)_1 \,| \geq n] \geq \epsilon.$$

This however contradicts the assumption in the theorem and therefore proves claim (7.2).

With this we also get that $|X|_1^*$ is an almost surely finitely valued random variable. In order to see this take the strategy $H \equiv 1$, then we get $|(H \cdot X)|_1^* = |X|_1^*$. By claim (7.2) it holds that

$$\lim_{c\to\infty} \sup_H \mathbb{P}[[|(H \cdot X)|_1^* > c] = 0.$$

This would be contradicted if there would be a set with positive probability, where $|X|_1^*$ is not finitely valued, proving $|X|_1^* < \infty,\ a.s.$

As a next step for a deterministic grid $\{0 = t_0, \dots, t_n\}$ define the strategy

$$H := \sum_{i=1}^{n-1} X_{t_i} \mathbb{1}_{(t_i, t_{i+1}]}.$$

One can write

$$\sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 = X_{t_n}^2 - 2(H \cdot X)_{t_n},$$

by expanding both sides and using $X_0 = 0$. Since by the definition of H it holds $|H|_1^* \leq |X|_1^*$ we have by the proved claim (7.2) that the convex hull of

$$\{[X, X]^\Pi \mid \text{ for any partition with deterministic times } \Pi\} \tag{7.3}$$

is bounded in probability. For a partition $\Pi = \{t_0 \leq \ldots \leq t_n\}$ in the claim we have the notation of the sampled process of the quadratic variation, i.e.

$$[X,X]^{\Pi} = \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2.$$

With this we are in the position to apply the Nikisin-Yan on the convex hull of (7.3). The theorem gives (see also Remark 5.5) the existence of an equivalent probability measure $\mathbb{Q} \sim \mathbb{P}$ such that

$$\sup_{\Pi} \mathbb{E}_{\mathbb{Q}} \left[[X,X]_1^{\Pi} \right] + \mathbb{E}_{\mathbb{Q}} \left[(|X|_1^*)^2 \right] + \sup_{\|H\|_\infty \leq 1} \mathbb{E}_{\mathbb{Q}} \left[(H \cdot X)_1 \right] \leq U < \infty. \tag{7.4}$$

Having done this first measure change we claim that the total mean variation (we bound the mean variation uniformly in $t \leq 1$) of X with respect to $\mathbb{Q}$ is finite. This means we claim

$$m - var(X)_1 = \sup_{\pi} \mathbb{E}_{\mathbb{Q}} \left[\sum_{i=1}^{n-1} |\, \mathbb{E}_{\mathbb{Q}}[X_{t_{i+1}} - X_{t_i} | \mathcal{F}_{t_i}] \,| \right] \leq U < \infty. \tag{7.5}$$

To see this define a strategy as

$$H := \sum_{i=1}^{n-1} \operatorname{sign} \left(\mathbb{E}_{\mathbb{Q}}[X_{t_{i+1}} - X_{t_i} \mid F_{t_i}] \right) \mathbb{1}_{(t_i, t_{i+1}]}.$$

Having this strategy at hand we calculate its stochastic integral with respect to X, i.e.

$$\mathbb{E}_{\mathbb{Q}}[(H \cdot X)_1] = \sum_{i=1}^{n-1} \mathbb{E}_{\mathbb{Q}}[\operatorname{sign} \left(\mathbb{E}_{\mathbb{Q}}[X_{t_{i+1}} - X_{t_i} \mid F_{t_i}] \right) (X_{t_{i+1}} - X_{t_i})] =$$
$$\sum_{i=1}^{n-1} \mathbb{E}_{\mathbb{Q}}[\operatorname{sign} \left(\mathbb{E}_{\mathbb{Q}}[X_{t_{i+1}} - X_{t_i} \mid F_{t_i}] \right) \mathbb{E}_{\mathbb{Q}}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}]] =$$
$$\sum_{i=1}^{n-1} \mathbb{E}_{\mathbb{Q}}[|\, \mathbb{E}_{\mathbb{Q}}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}] \,|] = \mathbb{E}_{\mathbb{Q}} \left[\sum_{i=1}^{n-1} |\, \mathbb{E}_{\mathbb{Q}}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}] \right].$$

The last term is precisely the term over which the supremum of partitions is taken over in the definition of the mean variation. Since the supremum of stochastic integrals of strategies which are bounded by 1 with respect to X is bounded we get that the total mean variation of X is finite. This means claim (7.5) is proved and X is a quasimartingale.

As next step take a strategy $H = \sum_{i=0}^{n-1} H_i \mathbb{1}_{(t_i, t_{i+1}]}$ uniformly bounded by 1 and deterministic grid $\Pi = \{0 = t_0 < \ldots < t_n\}$. Then we can define a discrete martingale via

$$M_{t_j} := \sum_{i=0}^{j-1} H_i(X_{t_{i+1}} - X_{t_i}) - A_{t_i},$$

where $A_{t_i} := H_i \mathbb{E}_{\mathbb{Q}}[X_{t_{i+1}} - X_{t_i} | \mathcal{F}_{t_i}]$. Apply Doob's maximal inequality on this discrete martingale to get

$$\mathbb{E}_{\mathbb{Q}}[(|M|_{t_n}^*)^2] \leq 4\mathbb{E}_{\mathbb{Q}}[M_{t_n}^2] \leq 16U. \tag{7.6}$$

The second inequality follows by using orthogonality from

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}[M_{t_n}^2] &= \sum_{i=0}^{n-1} H_i^2 \mathbb{E}_{\mathbb{Q}}\big[(X_{t_{i+1}} - X_{t_i} - \mathbb{E}_{\mathbb{Q}}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}])^2\big] \overset{(a+b)^2 \leq 2(a^2+b^2)}{\leq} \\ &\leq 2\sum_{i=0}^{n-1} \left(\mathbb{E}_{\mathbb{Q}}\big[(X_{t_{i+1}} - X_{t_i})^2\big] + \mathbb{E}_{\mathbb{Q}}\big[(\mathbb{E}_{\mathbb{Q}}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}])^2\big]\right) \leq \\ &\leq 4\sum_{i=0}^{n-1} \mathbb{E}_{\mathbb{Q}}\big[(X_{t_{i+1}} - X_{t_i})^2\big] \leq 4E_{\mathbb{Q}}\big[[X,X]_1^{\Pi}\big] \leq 4U. \end{aligned}$$

Note that also $\mathbb{E}_{\mathbb{Q}}[|A_t|] \leq$ m-var(X) holds and therefore with estimate (7.6) we get by using $|(H \cdot X)| \leq |M_t| + |A_t|$ and Hölder's inequality

$$\mathbb{E}_{\mathbb{Q}}\left[\sup_{t\in\Pi} |(H \cdot X)|_t\right] \leq \left(\mathbb{E}_{\mathbb{Q}}\left[\sup_{t\in\Pi} |M_t|^2\right]\right)^{\frac{1}{2}} + \text{m-var}_1(X) \leq 4\sqrt{U} + \text{m-var}_1(X) < \infty.$$

In this estimate the right side does neither depend on H nor n, therefore we obtain that actually

$$\sup_{\|H\|_\infty \leq 1} \mathbb{E}_{\mathbb{Q}}[|(H \cdot X)|_1^*] < \infty. \tag{7.7}$$

This also shows that the convex hull of

$$\{|(H \cdot X)|_1^* \mid \|H\|_\infty \leq 1\} \tag{7.8}$$

is bounded in probability $\mathbb{Q}$.

Keeping this fact in mind take again a strategy $H = \sum_{i=0}^{n-1} H_i \mathbb{1}_{(t_i,t_{i+1}]}$ uniformly bounded by 1 and deterministic grid $\Pi = \{0 = t_0 < \ldots < t_n\}$ and define another strategy via

$$K := \sum_{j=0}^{n-1} H_j \left(\sum_{i=0}^{j-1} H_i (X_{t_{i+1}} - X_{t_i}) \right) \mathbb{1}_{(t_j, t_{j+1}]}.$$

Then $K \leq |(H \cdot X)|_1^*$ holds and $(H \cdot X)^2$ can be decomposed into

$$(H \cdot X)^2 = \sum_{i=0}^{n-1} H_i^2 (X_{t_{i+1}} - X_{t_i})^2 + 2(K \cdot X).$$

Now look at

$$\sup_{\|H\|_\infty \leq 1} \sup_{\|K\|_\infty \leq |(H \cdot X)|_1^*} \mathbb{Q}[|(K \cdot X)|_1 \geq c] \leq$$

$$\sup_{\|H\|_\infty \leq 1} \sup_{\|K\|_\infty \leq |(H \cdot X)|_1^*} (\mathbb{Q}[|(K \cdot X)|_1 \geq c,\ |(H \cdot X)|_1^* < b] + \mathbb{Q}[|(H \cdot X)|_1^* \geq b]) \leq$$

$$\sup_{\|K\|_\infty \leq b} \mathbb{Q}[|(K \cdot X)|_1 \geq c] + \frac{\epsilon}{2}$$

where b is chosen large enough for given ϵ using that the convex hull of (7.8) is bounded in probability $\mathbb{Q}$. With this we obtain that the set

$$\operatorname{conv}\left(\{(H \cdot X)^2 \mid \|H\|_\infty \leq 1\}\right) \tag{7.9}$$

is bounded in probability $\mathbb{Q}$ as the sum of two convex sets which are both bounded in probability $\mathbb{Q}$, the first one being the convex hull of $\{\sum_{i=0}^{n-1} H_i^2 (X_{t_{i+1}} - X_{t_i})^2\}$.

With all these preliminary results, especially that the set in (7.9) is bounded in probability, we can apply the Nikisin-Yan Theorem (i.e. Theorem 5.4) again to obtain (after renaming) $\mathbb{Q} \sim \mathbb{P}$ for which we have

$$\sup_{\|H\|_\infty \leq 1,\ \Pi \ partition} \mathbb{E}_{\mathbb{Q}}[(H \cdot X)_1^2 + [X, X]_1^\Pi + (H \cdot X)_1] < \infty. \tag{7.10}$$

Now, we will use two theorems stated earlier in this Chapter to win a bit of regularity of X. First of all we can apply Theorem 7.3 by Rao to write $X = Y - Z$ with Y and Z being supermartingales. Next, we can switch to versions of Y and Z which we again call Y and Z for which we have that they are càdlàg on $[0,1] \setminus S$, with S being an at most countable set by Theorem 7.2. Therefore, there exists a version of X which we call again X which is càdlàg on $[0,1] \setminus S$.

We now set $\widehat{\Pi} := S \cup ([0,1] \cap \mathbb{Q})$ which is still countable.

Using the estimate 7.10 and taking any partition Π we get that the discrete time martingale

$$M_t^\Pi := \sum_{t_i<t} \big((X_{t_{i+1}} - X_{t_i}) - \mathbb{E}_\mathbb{Q}[X_{t_{i+1}} - X_{t_i}|\mathcal{F}_{t_i}]\big)$$

for $t \in \Pi$ is well defined. Furthermore, it is a square integrable martingale with respect to its corresponding discrete natural filtration. Totally we obtain by (7.10)

$$\sup_{\Pi,\ partition} \mathbb{E}_\mathbb{Q}\left[(M_1^\Pi)^2\right] < \infty. \tag{7.11}$$

Having this estimate take an enumeration $(t_n)_{n\geq 1}$ of $\widehat{\Pi}$ with $t_0 = 0$ and define

$$\Pi^n := \bigcup_{i=0}^{n} t_i, \quad n \geq 1. \tag{7.12}$$

It then holds $\Pi^1 \subset \ldots \subset \Pi^n \subset \Pi^{n+1} \subset \ldots \subset \widehat{\Pi}$ and

$$\bigcup_{n\geq 1} \Pi^n = \widehat{\Pi}.$$

Using now the estimate (7.11) we get

$$\sup_{n\geq 1} \mathbb{E}_\mathbb{Q}[(M_1^{\Pi^n})^2] \leq \sup_{\Pi,\ partition} \mathbb{E}_\mathbb{Q}[(M_1^\Pi)^2] < \infty. \tag{7.13}$$

Since balls in L^2 are weakly compact there exists a convergent subsequence, i.e. a subsequence of our original sequence of partitions which after renaming we call again $(\Pi^n)_{n\geq 1}$ for which it holds that

$$M_1^{\Pi^n} \rightharpoonup M_1$$

in $\mathbb{L}^2$ for some $M_1 \in \mathbb{L}^2$.

By our construction we still have $\bigcup_{n\geq 1} \Pi^n = \widehat{\Pi}$. The continuous time martingale with càdlàg trajectories generated by M_1 will be denoted by M, i.e. we have $M_t = \mathbb{E}_\mathbb{Q}[M_1|\mathcal{F}_t]$ a.s. for $t \in [0, 1]$. Furthermore, we define $A := X - M$.

By applying Lemma 7.5 we get $\mathbb{E}[M_1^{\Pi^k} \mid \mathcal{F}_t] \rightharpoonup \mathbb{E}[M_1 \mid \mathcal{F}_t]$ in $\mathbb{L}^2$ and therefore also $M_t^{\Pi^k} \rightharpoonup M_t$ whenever $t \in \widehat{\Pi}$.

Let now $\sigma_n = \{0 = s_0 < \ldots < s_n\}$ be a partition with points in $\widehat{\Pi}$ and take $Y \in \mathbb{L}^2(\mathbb{Q})$, s.t. $\|Y\|_2 = 1$. For large m we have $\sigma_n \subset \Pi^m$. We then can use weak convergence and estimate

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[Y\sum_{i=0}^{n-1}|A_{s_{i+1}}-A_{s_i}|] &= \sum_{i=0}^{n-1}\lim_{m\to\infty}\mathbb{E}_{\mathbb{Q}}[Y|X_{s_{i+1}}-M^{\Pi^m}_{s_{i+1}}-X_{s_i}+M^{\Pi^m}_{s_i}|]= \\
&= \sum_{i=0}^{n-1}\lim_{m\to\infty}\mathbb{E}[Y|\sum_{s_i\le t_j,<s_{i+1}:\ t_j\in\Pi^m}\mathbb{E}_{\mathbb{Q}}[X_{t_{j+1}}-X_{t_j}\mid\mathcal{F}_{t_j}]|]\le \\
&\le \sup_{m\ge1}\mathbb{E}_{\mathbb{Q}}[Y\sum_{t_j\in\Pi^m}|\mathbb{E}_{\mathbb{Q}}[X_{t_{j+1}}-X_{t_j}\mid\mathcal{F}_{t_j}]|]\le \\
&\le \sup_{m\ge1}\left(\mathbb{E}_{\mathbb{Q}}[(\sum_{t_j\in\Pi^m}|\mathbb{E}_{\mathbb{Q}}[X_{t_{j+1}}-X_{t_j}\mid\mathcal{F}_{t_j}]|)^2]\right)^{1/2}.
\end{aligned}$$

To show that the last expression is finite we define a strategy with

$$\begin{aligned}
H_j &:= \operatorname{sign}\left(\mathbb{E}_{\mathbb{Q}}[X_{t_{j+1}}-X_{t_j}\mid\mathcal{F}_{t_j}]\right), \\
H &:= \sum_{t_j\in\Pi^m}H_j\mathbb{1}_{(t_j,t_{j+1}]}.
\end{aligned}$$

With this we can write by adding a zero to the stochastic integral

$$\begin{aligned}
&(H\cdot X)= \\
&=\sum_{t_j\in\Pi^m}H_j(X_{t_{j+1}}-X_{t_j}-\mathbb{E}_{\mathbb{Q}}[X_{t_{j+1}}-X_{t_j}\mid\mathcal{F}_{t_j}])+H_j\mathbb{E}_{\mathbb{Q}}[X_{t_{j+1}}-X_{t_j}\mid\mathcal{F}_{t_j}].
\end{aligned}$$

The last expression can be reordered into

$$\begin{aligned}
&\sum_{t_j\in\Pi^m}|\mathbb{E}_{\mathbb{Q}}[X_{t_{j+1}}-X_{j_i}\mid\mathcal{F}_{t_j}]|= \\
&=(H\cdot X)-\sum_{t_j\in\Pi^m}H_i\left(X_{t_{j+1}}-X_{t_j}-\mathbb{E}_{\mathbb{Q}}[X_{t_{j+1}}-X_{t_j}\mid\mathcal{F}_{t_j}]\right).
\end{aligned}$$

By squaring and using $(a+b)^2\le 2(a^2+b^2)$ we obtain (cross terms in second sum gets zero after conditioning) for some suitable constant $C>0$ that

$$\begin{aligned}
&\sup_{m\in\mathbb{N}}\mathbb{E}_{\mathbb{Q}}\left[\left(\sum_{t_j\in\Pi^m}|\mathbb{E}_{\mathbb{Q}}[X_{t_{j+1}}-X_{j_i}\mid\mathcal{F}_{t_j}]|\right)^2\right]\le \\
&C\left(\sup_{\|H\|_\infty\le1}\mathbb{E}_{\mathbb{Q}}[(H\cdot X)^2]+\sup_{\pi}\mathbb{E}_{\mathbb{Q}}[[X,X]^\pi]\right)<\infty.
\end{aligned}$$

Therefore we have shown that

$$\mathbb{E}_{\mathbb{Q}}\big[Y\sum_{i=0}^{n-1}|A_{s_{i+1}}-A_{s_i}|\big] \leq C^{1/2}\left(\sup_{\|H\|_\infty\leq 1}\mathbb{E}_{\mathbb{Q}}[(H\cdot X)^2]+\sup_{\pi}\mathbb{E}_{\mathbb{Q}}[[X,X]^\pi]\right)^{1/2} < \infty$$

for all partitions with points $\widehat{\Pi}$ and $Y\in\mathbb{L}^2$ s.t. $\|Y\|_2=1$. With this we can conclude that A has finite variation, when only calculated on points in $\widehat{\Pi}$.

Now we are almost done because X and therefore also A are càdlàg on $[0,1]\setminus\widehat{\Pi}$. Take some arbitrary points $\{s_0,\dots,s_n\}\subset[0,1]\setminus\widehat{\Pi}$. Since $\widehat{\Pi}$ is dense, we can take sequences $(t_i^m)_{m\in\mathbb{N}}$ which are in $\widehat{\Pi}$ such that $t_i^m\to s_i$ for $i=0,\dots,n$. Hence for some arbitrary $Y\in\mathbb{L}^2$ with norm equal to 1 we can estimate by Fatou's Lemma

$$\mathbb{E}[Y\sum_{i=0}^{n-1}|A_{s_{i+1}}-A_{s_i}|] \leq \liminf_{m\to\infty}\mathbb{E}[Y\sum_{i=0}^{n-1}|A_{t_{i+1}^m}-A_{t_i^m}|]\leq C.$$

With this we proved the finite variation property for A, i.e. we found a version of X which we can write as

$$M+A$$

with M a càdlàg $\mathbb{L}^2(\mathbb{Q})-$martingale and A a process of finite variation.

In order to change back to the original measure $\mathbb{P}$ we just have to apply the Girsanov-Meyer Theorem. □

Having done our version of the Bichteler-Dellacherie theorem we also want to tackle a second very famous theorem in stochastic analysis, namely the Doob-Meyer decomposition and do a proof similar to the one done just now. In the Doob-Meyer decomposition one tries to decompose a supermartingale instead of a semimartingale. A proof can be deduced by almost the same techniques used in the end of the last theorem.

With this we can formulate our version of the Doob-Meyer theorem.

Theorem 7.8. *Let S be a non-negative supermartingale on $[0,1]$.*

i) Let S additionally be bounded. Then there exists a càdlàg martingale M and a predictable, increasing finite variation process A such that $S=M-A$. The decomoposition is unique up to versions of A and up to indistinguishibility of M.

ii) Let S be in a way such that there exists a sequence of stopping times τ_n which stops S at the level n, i.e.

$$S^{\tau_n}=S\wedge n+(S_{\tau_n}-n)\mathbb{1}_{[\tau_n,1)},$$

in particular $\mathbb{E}[S_{\tau_n}] \leq \mathbb{E}[S_0]$ *is assumed. Then* S *can be decomposed into a càdlàg local martingale* M *and an increasing, predictable process* A *such that* $S = M - A$*. Again the composition is unique up to versions of* A *and up to indistinguishibility of* M*.*

Proof. We first prove i).

Let S be a super-martingale such that $0 \leq S \leq c$ with $c \geq 0$. We get for a partition Π a discrete Doob-Meyer decomposition $S^\Pi = M^\Pi - A^\Pi$ with M^Π being a martingale and A^Π an increasing process. Furthermore, by Theorem 3.8 we have the estimate

$$\sup_\Pi \mathbb{E}\left[(M_1^\Pi)^2\right] \leq 2c\mathbb{E}[S_0] < \infty. \tag{7.14}$$

Now apply Theorem 7.2 to get a version $\tilde{S}$ of S such that it is has left and right limits everywhere and there exists a countable set $P \subset [0,1]$ such that $\tilde{S}_t$ is right-continuous at every $t \notin P$.

As a next step define $\widehat{\Pi} := P \cup (\mathbb{Q} \cap [0,1])$, which is a countable set. As in the previous theorem take an enumeration $(t_n)_{n\geq 1}$ of $\widehat{\Pi}$ and $t_0 = 0$ and define

$$\Pi^n := \bigcup_{i=0}^n t_i, \quad n \geq 1.$$

We then get $\bigcup_{n\geq 1} \Pi^n = \widehat{\Pi}$ and have a nested sequence of time points as in the previous theorem. Using again weak compactness in $\mathbb{L}^2$ and the estimate (7.14) we get on a subsequence which we rename as the original one

$$M_1^{\Pi^n} \rightharpoonup M_1 \tag{7.15}$$

with $M_1 \in \mathbb{L}^2$. Denote by M the càdlàg $\mathbb{L}^2$ martingale generated by M_1 and set $A := M - \tilde{S}$.

Now for all $n \geq 1$ we have that A^{Π^n} is an increasing process and since $\bigcup_{n\geq 1} \Pi^n = \widehat{\Pi}$ we get that A is an increasing process when only compared at points in $\widehat{\Pi}$.

But we also have that $\tilde{S}$ is actually right-continuous on $[0,1] \setminus \widehat{\Pi}$ (actually even $[0,1] \setminus N$).

With this we can show that A is actually an increasing process in general. Therefore, let $s < t$ with $s,\ t \in [0,1] \setminus \widehat{\Pi}$. $\widehat{\Pi}$ is dense in $[0,1]$, so let $(s_n)_{n\geq 1}$, $(t_n)_{n\geq 1}$ be two sequences in $\widehat{\Pi} \setminus N$ such that $s_n \leq t_n$ for all $n \geq 1$. Since A is an increasing process on $\widehat{\Pi}$ we get

$$A_{s_n} \leq A_{t_n}, \quad \forall n \geq 1.$$

Furthermore, A is right-continuous at s and t so we get

$$\lim_{n\to\infty} A_{s_n} = A_s$$
$$\lim_{n\to\infty} A_{t_n} = A_t$$

and therefore finally $A_s \leq A_t$.

So we proved that A is an increasing process and therefore $\tilde{S}$ admits a decomposition $\tilde{S} = M - A$ into a càdlàg $\mathbb{L}^2$ martingale minus an increasing process. For the original process we therefore get the equality also on an almost sure level, i.e.

$$\forall t \in [0,1]: \; S_t = M_t - A_t, \; a.s. \tag{7.16}$$

Since A is an increasing process, it is of finite variation as well.

Let now N be a càdlàg square integrable martingale and take an arbitrary sequence $(\sigma_n)_{n\geq 1}$ of partitions tending to the identity and chosen in $\widehat{\Pi}$, then we have (using weak convergence)

$$\lim_{n\to\infty} \mathbb{E}(N_-{}^{\sigma_n} \cdot A^{\sigma_n})_1] = \lim_{n\to\infty} \mathbb{E}[\sum_{i=0}^{n-1} N_{t_i}(S_{t_{i+1}} - S_{t_i})] = \tag{7.17}$$

$$\lim_{n\to\infty} \mathbb{E}[N_1 \sum_{i=0}^{n-1} \mathbb{E}[S_{t_{i+1}} - S_{t_i} \mid \mathcal{F}_{t_i}]] = \mathbb{E}[N_1 A_1], \tag{7.18}$$

since

$$\mathbb{E}[\sum_{i=0}^{n-1} N_{t_i}(S_{t_{i+1}} - S_{t_i})] \overset{\text{first } \mathbb{E}[|\mathcal{F}_{t_i}]}{=}$$

$$\sum_{i=0}^{n-1} \big(\mathbb{E}[(N_{t_i} - N_1)\mathbb{E}[S_{t_{i+1}} - S_{t_i}|\mathcal{F}_{t_i}]] + \mathbb{E}[N_1 \mathbb{E}[S_{t_{i+1}} - S_{t_i}|\mathcal{F}_{t_i}]]\big) =$$

$$\mathbb{E}[\sum_{i=0}^{n-1} N_1 \mathbb{E}[S_{t_{i+1}} - S_{t_i}|\mathcal{F}_{t_i}]].$$

So what we proved is that for an $\mathbb{L}^2$ càdlàg martingale N we have

$$\lim_{n\to\infty} \mathbb{E}[(N_-{}^{\sigma_n} \cdot A^{\sigma_n})_1] = \mathbb{E}[N_1 A_1] \tag{7.19}$$

which is naturality. Doing the same for any $t < 1$ we get an equation like

$$\lim_{n\to\infty} \mathbb{E}[(N_-^{\sigma_n} \cdot A^{\sigma_n})_t] = \mathbb{E}[N_t A_t]. \tag{7.20}$$

Assume now you have another decomposition of $S = M^2 - A^2$ with M^2 being a càdlàg $\mathbb{L}^2$ martingale and A^2 an increasing process. We then get $M - A = M^2 - A^2$ and therefore $M - M^2 = A^2 - A$ being a càdlàg $\mathbb{L}^2$ martingale. Using this in equation (7.20) we obtain

$$\mathbb{E}[(A^2 - A)_t^2] = \lim_{n\to\infty} \mathbb{E}[((A^2 - A)_-^{\sigma_n} \cdot (A^2 - A)^{\sigma_n})_t] = 0.$$

So we have that A^2 and A are versions from each other, meaning A is unique up to version. We get the same for M but since it is càdlàg this extends to indistinguishability. Predictability of A follows by general facts on natural processes, which we proved for A in (7.19) already.

Now we are left with proving *ii*).

For this sake let S be a non-negative super-martingale such that there exists a sequence of stopping times τ_n with

$$S^{\tau_n} = S \wedge n + (S_{\tau_n} - n)\mathbb{1}_{[\tau_n,1)},$$

like in our assumption, in particular $\mathbb{E}[S_{\tau_n}] \leq \mathbb{E}[S_0]$.

First of all, $S \wedge n$ is a bounded, non-negative supermartingale, so by *i*) we get a decomposition. The second term can be written as

$$(S_{\tau_n} - n - \mathbb{E}[S_{\tau_n} - n])\mathbb{1}_{[\tau_n,1)} + \mathbb{E}[S_{\tau_n} - n]\mathbb{1}_{[\tau_n,1)}. \tag{7.21}$$

Here, the first part in the sum is a martingale and the second sum is by our assumption a decreasing process if $n \geq \mathbb{E}[S_0]$, so for large enough n.

Therefore, for large enough n we have a decomposition into a càdlàg martingale M^n minus a non-decreasing finite variation process A^n, i.e.

$$S^{\tau_n} = M^n - A^n.$$

The decomposition is also unique up to versions w.r.t. A^n and indistinguishability w.r.t. M^n. This means that for $n \geq m$ we have

$$(M^n)^{\tau_m} = M^m, \quad \text{and} \quad (A^n)^{\tau_m} = A^m.$$

With this we can define a càdlàg local martingale M and an increasing finite variation process A such that

$$S = M - A,$$

where the decomposition is unique up to version w.r.t. A and unique w.r.t. indistinguishability w.r.t. M.

□

8 Appendix

In a rather extensive appendix a guided tour through probability theory, martingale theory and stochastic integration in Banach spaces is given. The first two sections are rather important since they build the foundation of the generalization of real-valued stochastic processes to stochastic processes with values in Banach spaces. Most of the remaining sections are not too relevant for the sake of reading this thesis, however they summarise already existing very general stochastic integration theory and played a crucial role for gaining ideas of this thesis, especially for Chapters 3 and 4.

Most results and notes are based on and taken from [vVW15] by Jan van Neerven, Mark Veraar and Lutz Weis, [vNVW07] by van Neerven, Veraar and Weis, [Pis11] by Gilles Pisier, [vN07] by van Neerven and [Pro72].

If a definition or a theorem is cited the reference will be included in brackets right next to it.

8.1 Probability Theory in Banach spaces

Here, we will introduce some basic concepts and definitions of probability theory in Banach spaces in order to talk about expectation, martingales, etc. It will be based on [Pro72].

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete (i.e. $\mathcal{A}$ contains all $\mathbb{P}$ nullsets) probability space and $(E, \mathcal{B})$ be a measurable space. Here E is a Banach space and $\mathcal{B}$ denotes the $\sigma-$algebra of all Borel subsets of E generated by the open sets in E. We start with some definitions.

Definition 8.1. *[Pro72, Definition 1.10] A map $X : \Omega \to E$ is a random variable with values in E, if for all Borel sets $B \in \mathcal{B}$ we have $X^{-1}(B) \in \mathcal{A}$.*

Given a $\sigma-$algebra $\mathcal{G} \subset \mathcal{A}$, a random variable is called measurable with respect to $\mathcal{G}$ if $X^{-1}(B) \in \mathcal{G}$ for all $B \in \mathcal{B}$.

Definition 8.2. *[Pro72, Definition 1.11] $X : \Omega \to E$ is called a finitely valued random variable or a simple random variable, if there are finitely many disjoint sets $B_1, \dots, B_n \in \mathcal{B}$ such that X is constant on B_i for all $i = 1, \dots, n$ and 0 on $\Omega \setminus (\cup_{i=1}^n B_i)$.*

One can define this for a general measure instead of a probability measure but since we will only need it for probability spaces we define it in this setting.

Definition 8.3. *[Pro72, Definition 1.12] $X : \Omega \to E$ is said to be a $\mathbb{P}-$almost separably valued random variable if there exists a set $B_0 \in \mathcal{B}$ such that $\mathbb{P}[B_0] = 0$ and $X(\Omega \setminus B_0)$ is separable.*

Now we can also define strong measurability which is the crucial concept of measurability in Banach spaces.

Definition 8.4 (strong or Bochner random variable)**.** *[Pro72, Definition 1.13] $X : \Omega \to E$ is said to be a strong random variable, or also called a Bochner random variable, if there exists a sequence $(X_n)_{n\geq 1}$ of simple random variables which converges almost surely to X, i.e.*

$$\lim_{n\to\infty} \|X_n - X\| = 0, \quad \mathbb{P} - a.s.$$

Of course, if there are strong measurable random variables, there have to be weak ones as well.

Definition 8.5 (weak or Pettis random variable)**.** *[Pro72, Definition 1.14] $X : \Omega \to E$ is called a weak random variable, also called a Pettis random variable, if for all $x^* \in E^*$ the functions $x^*(X)$ are real-valued random variables.*

There is the following connection between weak and strong random variables.

Theorem 8.6. *[Pro72, Theorem 1.2] $X : \Omega \to E$ is a strong random variable if and only if it is a weak random variable and $\mathbb{P}-$almost separably valued.*

In the case when E is a separable Banach space the $\sigma-$algebra generated by the set of all spherical neighbourhoods of E is equal to the Borel $\sigma-$algebra $\mathcal{B}$. We then have that all definitions of being a random variable in some way are equivalent, so we can talk about Banach space valued random variables or $E-$valued random variables. Therefore, when we only say Banach space valued random variables we always implicitly assume the Banach space to be separable.

Definition 8.7. *[Pro72, Definition 1.18] Let X, Y be two $E-$valued random variables on the same probability space. X and Y are said to be equivalent if*

$$\mathbb{P}[\{\omega : X(\omega) \in B\} \triangle \{\omega : Y(\omega) \in B\}] = 0.$$

In the case when E is separable, equivalence means that X and Y are equal with probability one, i.e.

$$\mathbb{P}[\{\omega : X(\omega) \neq Y(\omega)\}] = 0.$$

Next, we cite a definition concerning some means of convergence of Banach space valued random variables.

Definition 8.8. *[Pro72, Definition 1.20] Let $(X_n)_{n\geq 1}$ be a sequence of $E-$valued random variables and X be a $E-$valued random variable. We say $(X_n)_{n\geq 1}$ converges to X in Ω*

i) strongly almost surely if there exists a $\mathbb{P}-$nullset $B \in \mathcal{B}$ such that

$$\lim_{n\to\infty} \|X_n(\omega) - X(\omega)\| = 0, \quad \forall \omega \in \Omega \setminus B;$$

ii) weakly almost surely if there exists a $\mathbb{P}-$nullset $B \in \mathcal{B}$ such that

$$\lim_{n\to\infty} x^*(X_n(\omega)) = x^*(X(\omega)); \quad \forall x^* \in E^*, \ \omega \in \Omega \setminus B;$$

iii) in probability if

$$\forall \epsilon > 0 : \ \lim_{n\to\infty} \mathbb{P}^*[\{\omega : \|X_n(\omega) - X(\omega)\| > \epsilon] = 0,$$

where $\mathbb{P}^*$ denotes the outer measure of $\mathbb{P}$.

Now we can turn our attention to the important definition of the Bochner integral. As usual, first it is defined for simple functions and then generalized to more general functions via an approximation.

Definition 8.9. *[Pro72, Definition 1.23] A simple random variable X is called Bochner integrable if and only if $\|X\| \in \mathbb{L}^1(\Omega)$. We then define*

$$\int_\Omega X(\omega)\ d\mathbb{P}(\omega) := \sum_{i=1}^{\infty} x_i \mathbb{P}[B_i \cap \Omega],$$

where $X(\omega) = x_i$ on $B_i \in \mathcal{B}$, $i \in \mathbb{N}$.

The integral over a measurable set $A \in \mathcal{B}$ is then defined via

$$\int_B X(\omega)\ d\mathbb{P}(\omega) := \int_\Omega X(\omega)\mathbb{1}_B(\omega)\ d\mathbb{P}(\omega).$$

Concerning the notation we will mostly omit the ω. From the definition we get the immediate inequality

$$\| \int_\Omega X(\omega)\ d\mathbb{P}(\omega) \| \leq \int_\Omega \|X(\omega)\|\ d\mathbb{P}(\omega). \tag{8.1}$$

Definition 8.10 (Bochner integral). *[Pro72, Definition 1.24] X is said to be Bochner integrable if there exists a sequence of simple random variables $(X_n)_{n\geq 1}$ converging* *to X almost surely such that*

$$\lim_{n\to\infty}\int_\Omega \|X_n(\omega)-X(\omega)\|\ d\mathbb{P}(\omega)=0.$$

We then define

$$\int_\Omega X(\omega)\ d\mathbb{P}(\omega):=\lim_{n\to\infty}\int_\Omega X_n(\omega)\ d\mathbb{P}(\omega).$$

We can define the integral as a limit because by inequality (8.1) the integrals of the simple random variables form a Cauchy sequence obtaining a limit in E.

Remark 8.11. *For a Bochner integrable random variable X we have for all $x^*\in E^*$*

$$x^*\left(\int_\Omega X\ d\mathbb{P}\right)=\int_\Omega x^*(X)\ d\mathbb{P}$$

Next we can define the expected value of a strong random variable.

Definition 8.12. *[Pro72, Definition 1.25] The expected value of a strong random variable X which is Bochner integrable is defined as*

$$\mathbb{E}[X]:=\int_\Omega X(\omega)\ d\mathbb{P}(\omega).$$

This expectation is often also referred to as strong expectation.

We state the following theorem giving a necessary and sufficient condition for a random variable to be integrable.

Theorem 8.13. *[Pro72, Theorem 1.8] $\mathbb{E}[X]$ exists for $X:\Omega\to E$ if and only if X is a strong random variable and $\mathbb{E}[\|X\|]<\infty$.*

Similar to the real-valued case we define $\mathbb{L}^1(\Omega;E)$ as the set of all $E-$valued random variables which are Bochner integrable. However we consider elements as equivalence classes w.r.t. the then becoming a norm an that space

$$\|X\|_{\mathbb{L}^1(\Omega;E)}:=\int_\Omega \|X(\omega)\|\ d\mathbb{P}(\omega)=\mathbb{E}[\|X\|].$$

$\mathbb{L}^1(\Omega;E)$ becomes a Banach space with that norm.

For $p\in(1,\infty)$ we do it similarly and define $\mathbb{L}^p(\Omega;E)$ as the set of all $E-$valued random variables X (identified up to almost sure equivalence) such that

$$\lim_{n\to\infty}\int_\Omega \|X_n(\omega)-X(\omega)\|^p\ d\mathbb{P}(\omega)=0$$

for some sequence $(X_n)_{n\geq 1}$ of simple random variables. They become Banach spaces when equipped with the norm

$$\|X\|_{\mathbb{L}^p(\Omega;\mathcal{X})} := \left(\int_\Omega \|X(\omega)\|^p \, d\mathbb{P}(\omega)\right)^{1/p}.$$

The last space we define is $\mathbb{L}^\infty(\Omega; E)$ which is the set of all strong random variables X (identified up to almost sure equivalence) such that $\|X\| \in \mathbb{L}^\infty(\Omega)$.

There is the following well known theorem as in the real-valued case.

Theorem 8.14. *For $p \in [1,\infty]$ the spaces $\mathbb{L}^p(\Omega; E)$ and $\mathbb{L}^\infty(\Omega; E)$ are Banach spaces and the simple random variables are dense in $\mathbb{L}^p(\Omega; E)$.*

The usual properties like linearity, monotonicity, etc. hold for the expectation, but we also state the form of dominated convergence in this case.

Theorem 8.15. *[Pro72, page 23, 4.] If a sequence $\{X_n\}_{n\geq 1} \subset \mathbb{L}^1(\Omega; E)$ converges almost surely to a random variable X and there exists a nonnegative random variable $Y \in \mathbb{L}^1(\Omega)$ such that $\|X_n\| \leq Y$ a.s. for all $n \in \mathbb{N}$, then we have $X \in \mathbb{L}^1(\Omega; E)$ and*

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X].$$

We also state the following theorem which allows exchanging of expectation and linear operators.

Theorem 8.16. *[Pro72, Theorem 1.9] Let L be a closed linear operator on E onto itself. Suppose $\mathbb{E}[X]$ and $\mathbb{E}[L(X)]$ exist, then we have $L(\mathbb{E}[X]) = \mathbb{E}[L(X)]$.*

Remark 8.17. *Given two Banach spaces E, F and a linear operator $T : E \to F$ we call T closed if for all sequences $\{x_n\}_{n\geq 1} \subset E$ such that $x_n \to x \in E$ and $Tx_n \to y \in F$ it follows that $Tx = y$.*

In the following we will define the conditional expectation of Banach space valued random variables. For this sake fix a sub $\sigma-$algebra $\mathcal{G} \subset \mathcal{B}$.

Definition 8.18. *[Pro72, Definition 1.27] Given a random variable $X \in \mathbb{L}^1(\Omega; E)$ we call a strong random variable $\mathbb{E}[X \mid \mathcal{G}]$ the conditional expectation of X relative to $\mathcal{G}$ if it satisfied the following two conditions:*

i) $\mathbb{E}[X \mid \mathcal{G}]$ is measurable w.r.t. $\mathcal{G}$ and an element of $\mathbb{L}^1(\Omega; E)$;

ii) for every $A \in \mathcal{G}$ it holds

$$\int_A \mathbb{E}[X \mid \mathcal{G}](\omega) \, d\mathbb{P}(\omega) = \int_A X(\omega) \, d\mathbb{P}(\omega).$$

Also for the conditional expectation the usual properties hold, but we state again two theorems which might be useful including a dominated convergence theorem.

Theorem 8.19. *[Pro72, page 24, 5.] Let L be a bounded linear operator from E onto itself. Then it holds that*

$$L(\mathbb{E}[X \mid \mathcal{G}]) = \mathbb{E}[L(X) \mid \mathcal{G}], \quad a.s.$$

Theorem 8.20. *[Pro72, page 24, 6.] Let $X_n \to X$ strongly a.s. and assume there exists a nonnegative random variable $Y \in \mathbb{L}^1(\Omega)$ such that $\|X_n\| \leq Y$ a.s., then we have*

$$\lim_{n\to\infty} \mathbb{E}[X_n \mid \mathcal{G}] = \mathbb{E}[X \mid \mathcal{G}].$$

The next topic we introduce are Banach space-valued processes. We again consider a measurable space $(E, \mathcal{B})$ with E being a separable Banach space and $\mathcal{B}$ the $\sigma-$algebra of Borel subsets of E, but we also consider a measurable space $(T, \mathcal{T})$, where T is a subset of the extended real line $\overline{\mathbb{R}}$ (think of $\mathbb{R}_{\geq 0} \cup \infty$ or $\overline{\mathbb{N}}$ and $\mathcal{T}$ is the $\sigma-$algebra of Borel subsets of T).

Definition 8.21. *An $E-$valued stochastic process is a map $X : T \times \Omega \to E$ such that for every $t \in T$, $X(t, \cdot)$ is an $E-$valued random variable.*

Given a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ an $E-$valued stochastic process is adapted if for all $t \in T$, $X_t(\cdot) := X(t, \cdot)$ is F_t-measurable.

We call two stochastic processes X, Y versions of each other if $X(t, \cdot) = Y(t, \cdot)$ *a.s.* $\forall t \in T$. For technicalities we also state the following definition.

Definition 8.22. *[Pro72, Definition 1.36] An $E-$valued stochastic process X is said to be separable if there exists a countable subset $S \subset T$ and $B \in \mathcal{B}$ with $\mathbb{P}[B] = 0$, such that for all events $\{\omega : X(t, \omega) \in F;\ t \in I \cap T\}$ where $F \subset E$ is closed and I an open interval we have that for the symmetric difference it holds*

$$\{\omega : X(t, \omega) \in F,\ t \in I \cap S\} \triangle \{\omega : X(t, \omega) \in F,\ t \in I \cap T\} \subset B.$$

Definition 8.23. *[Pro72, Definition 1.37] A measurable map $X : (T \times \Omega, \mathcal{T} \otimes \mathcal{F}) \to (E, \mathcal{B})$ is called a measurable $E-$valued stochastic process.*

There is also an analogue to Fubini's theorem.

Theorem 8.24 (Fubini)**.** *[Pro72, Theorem 1.11] Let X be Bochner integrable on $T \times \Omega$. Then the functions $Y(\omega) := \int_T X(t, \omega)\ d\lambda(t)$ and $Z(t) := \int_\Omega X(t, \omega)\ d\mathbb{P}(\omega)$ are defined almost everywhere on Ω and T, respectively, and it holds that*

$$\int_{T\times\Omega} X(t, \omega)\ d(\lambda \times \mathbb{P})(t, \omega) = \int_T Z(t)\ d\lambda(t) = \int_\Omega Y(\omega)\ d\mathbb{P}(\omega).$$

We also state some concepts of continuity.

Definition 8.25. *[Pro72, Definition 1.38] Let X be a separable $E-$valued stochastic process and $t_0 \in T$. Define*

$$\Omega_{t_0} := \{\omega : X(t_0, \omega) = \lim_{t \to t_0} X(t, \omega)\},$$

where the limit is taken in strong (respectively weak) sense. X is said to be strongly (respectively weakly) continuous a.s. at t_0 if $\mathbb{P}[\Omega_{t_0}] = 0$. X is said to be strongly (respectively weakly) continuous a.s. on T if it is strongly (respectively weakly) continuous a.s. at every $t \in T$.

As last point we define martingales. Therefore, T is either an interval of the extended real line or the extended real line. Furthermore let $(\mathcal{F}_t)_{t\in T}$ be an increasing sequence of sub $\sigma-$algebras of $\mathcal{F}$. Let X be an $E-$valued stochastic process with values in a separable Banach space E.

Definition 8.26 (Martingales)**.** *[Pro72, Definition 1.38] An $E-$valued stochastic process X which is adapted to $(\mathcal{F}_t)_{t\in T}$ is said to be a Banach space valued martingale if*

i) $\mathbb{E}[\|X_t\|] < \infty, \ \forall t \in T$ *and*

ii) $\forall s, \ t \in T, \ s.t. \ s < t$ *we have* $\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s$ *a.s.*

If in addition, it holds for all $t \in T$ that $\mathbb{E}[\|X_t\|^p] < \infty$ we call X an $E-$valued $\mathbb{L}^p-$martingale.

Having martingales at hand we will define the discrete stochastic integral which is just the martingale transform.

Definition 8.27. *Let $X = (X_i)_{i=1}^n$ an $E-$valued martingale and $H = (H_i)_{i=1}^n$ a predictable and real-valued process. The martingale transform of X by H is defined as $(H \cdot X) := ((H \cdot X)_i)_{i=1}^n$ via*

$$(H \cdot X)_i := \sum_{k=1}^{i} H_k(X_k - X_{k-1}), \quad k = 1, \dots, n,$$

where we set $X_0 = 0$.

One can also show Doob inequalities for Banach space valued martingales and a convergence theorem. For this sake we fix a finite time $E-$valued martingale, i.e. $X = (X_i)_{i=1}^n$. We will define the running maximum via

$$X_n^* : \Omega \to \mathbb{R}_+ : \omega \to \max_{1\leq i\leq n} \|X_i(\omega)\|.$$

We can formulate the respective Doob theorems in the setting which mostly follow from the real-valued case. For details we can refer to [Pis11, Chapter 1] and [vN07, Chapter 11].

Theorem 8.28 (Doob inequalities)**.** *[vN07, Theorem 11.20] For* $\lambda > 0$ *we have*

$$\mathbb{P}[X_n^* > \lambda] \leq \frac{1}{\lambda}\mathbb{E}[\|X_n\|].$$

If $p \in (1,\infty)$ *and* $X_n \in \mathbb{L}^p(\Omega; E)$ *we have* $X_n^* \in \mathbb{L}^p(\Omega; E)$ *and*

$$\|X_n^*\|_p \leq \frac{p}{p-1}\|X_n\|_p.$$

Theorem 8.29 (convergence theorems)**.** *[vN07, Theorem 11.22] Let* $(X_n)_{n\in\mathbb{N}}$ *be an* $\mathbb{L}^1-$*bounded* $E-$*valued martingale. For an* $E-$*valued random variable* X *the following are equivalent:*

i) $\forall x^* \in E^* : x^*(X_n) \to x^*(X),$ *a.s.*

ii) $\forall x^* \in E^* : x^*(X_n) \to x^*(X),$ *in probability*

iii) $X_n \to X$ *a.s.*

iv) $X_n \to X$ *in probability.*

If for some $p \in [1,\infty)$ *in addition, we have* $X \in \mathbb{L}^p(\Omega; E)$ *then it holds* $X_n \in \mathbb{L}^p(\Omega; E)$ *for all* $n \in \mathbb{N}$ *and we have* $X_n \to X$ *in* $\mathbb{L}^p(\Omega; E)$*.*

8.2 Radonifying Operators

In the next sections we will give a tour through already existing stochastic integration theory in Banach spaces. For this sake one needs to talk about radonifying operators which we will introduce in the following via [vNVW07], [vN07] and [vVW15]. For the interested reader the most detailed reference are the lecture notes [vN07].

In order to come to radonifying operators we will first introduce summing operators. In the following H denotes a Hilbert space and $(\gamma_n)_{n\geq 1}$ a Gaussian sequence, i.e. a sequence of independent standard Gaussian random variables (i.e. i.i.d. $\mathcal{N}(0,1)$). We start with the definition of $\gamma-$summing operators.

Definition 8.30. *[vN07, Definition 5.1] A linear operator* $T : H \to E$ *is said to be* $\gamma-$*summing if for some* $p \in [1,\infty)$ *we have*

$$\|T\|_{\gamma_p^\infty(H,E)} := \sup\left(\mathbb{E}\left[\Big\|\sum_{i=1}^n \gamma_i T h_i\Big\|^p\right]\right)^{1/p} < \infty$$

where the supremum is taken over all finite orthonormal systems $\{h_1, \dots, h_n\}$.

For $p = 2$, *we set* $\|T\|_{\gamma^\infty(H,E)} := \|T\|_{\gamma_p^\infty(H,E)}$.

The space $\gamma^\infty(H, E)$ *is the space of all* $\gamma-$*summing operators.*

Remark 8.31. *If an operator is* $\gamma-$*summing for a* $p \in [1, \infty)$ *this is actually already equivalent to being* $\gamma-$*summing for all* $p \in [1, \infty)$. *A detailed proof of this can be found in [vN07] as well.*

Theorem 8.32. *[vN07, Proposition 5.2]* $\gamma^\infty(H, E)$ *is a Banach space.*

Now we can turn towards radonifying operators. We denote by $H \otimes X$ the linear space of all finite rank operators from H to X. That means $T \in H \otimes X$ if $T : H \to X$ is a linear and bounded operator whose range is finite-dimensional. If the dimension is n it is said to have rank n. Such an operator can be written as

$$Th = \sum_{i=1}^{n} \langle h\,, v_i \rangle u_i, \quad \forall h \in H$$

for some $v_i \in H,\ u_i \in X,\ i = 1, \dots, n$. For the map $\langle \cdot\,, v_i \rangle u_i$ we also write $v_i \otimes u_i$. Then the operator looks like

$$T = \sum_{i=1}^{n} v_i \otimes u_i.$$

The $v_1, \dots, v_n$ can also be assumed to be orthonormal by a Gram-Schmidt process argument. It follows immediately that a finite rank operator T is an element of $\gamma^\infty(H, E)$. Actually we have

Lemma 8.33. *[vN07, Lemma 5.7] Let* $T = \sum_{i=1}^{n} v_i \otimes u_i$ *be a finite rank operator. Then we have for all* $p \in (1, \infty)$

$$\|T\|^p_{\gamma_p^\infty(H,E)} = \mathbb{E}\left[\Big\| \sum_{i=1}^{n} \gamma_i u_i \Big\|^p \right].$$

We also define the space of radonifying operators.

Definition 8.34. *[vN07, Definition 5.8] The Banach space* $\gamma(H, X)$ *is defined as the completion of* $H \otimes X$ *in* $\gamma^\infty(H, E)$.

$\gamma(H, X)$ is a Banach space by definition and for $T \in \gamma(H, E)$ we write for the norm

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

We can also extend this definition towards $p \in [1, \infty)$ (however, for $p = 2$ we will always write just $\gamma(H, X)$) and call the space $\gamma^p(H, X)$ and the norm respectively if we use the Kahane-Khintchine inequality which tells that for all $0 < p,\ q < \infty$ there exists $\kappa_{q,p} \geq 0$ such that

$$\left(\mathbb{E}\left[\|\sum_{n=1}^{N}\gamma_n x_n\|^q\right]\right)^{1/q} \leq \kappa_{q,p}\left(\mathbb{E}\left[\|\sum_{n=1}^{N}\gamma_n x_n\|^p\right]\right)^{1/p}. \tag{8.2}$$

This tells us that the norms are equal and $\gamma^p(H,E) = \gamma^q(H,E)$.

The identity on $H \otimes X$ extends to an injective and contractive embedding of $\gamma(H,X)$ into $\mathcal{L}(H,X)$, the space of all bounded linear operators from H to X. So we can identify $\gamma(H,X)$ as a linear subspace of $\mathcal{L}(H,X)$. With this we call a bounded operator $T \in \mathcal{L}(H,X)$ $\gamma - radonifying$ if it belongs to $\gamma(H,X)$. In the case when H would be separable, T would be $\gamma - radonifying$ if there exists an orthonormal basis $(h_n)_{n\geq 1}$ in H such that $\sum_{n\geq 1}\gamma_n T h_n$ converges in $\mathbb{L}^2(\Omega;X)$. Then its norm would look like (in the case of $p = 2$)

$$\|T\|_{\gamma(H,X)} = \left(\mathbb{E}\left[\|\sum_{n\geq 1}\gamma_n T h_n\|^2\right]\right)^{1/2}.$$

Actually it will turn out that it is enough to consider separable Hilbert spaces in the following.

We have the following characterisation concerning measurability.

Theorem 8.35. *[vN07][Proposition 5.14] Let $(A, \mathcal{A}, \mu)$ be a $\sigma-$finite measure space and H be a separable Hilbert space. Let $\Phi : A \to \gamma(H,E)$ and define $\Phi h : A \to E, \zeta \mapsto (\Phi h)(\zeta) := \Phi(\zeta)h$. Then the following are equivalent*

i) Φ is strongly $\mu-$measurable;

ii) Φh is strongly $\mu-$measurable for all $h \in H$.

Remark 8.36. *(see also [vNVW07, Proposition 2.6]) For a $\sigma-$finite measure space (S, Σ, μ) we can associate a map*
$f \in \mathbb{L}^p(\mu; \gamma^p(H,E))$ for $p \in [1,\infty)$ with the mapping

$$H \to \mathbb{L}^p(\mu;E) : \tilde{h} \mapsto f(\cdot)\tilde{h}$$

to define an isometric isomorphism

$$F_{\gamma^p} : \mathbb{L}^p(\mu;\gamma^p(H,E)) \to \gamma^p(H, \mathbb{L}^p(\mu;E)) : f \mapsto (F_{\gamma^p}(f)(h)(s) := f(s)h)$$

and see that the following Banach spaces are isomorphic

$$\gamma^p(H, \mathbb{L}^p(\mu;E)) \simeq \mathbb{L}^p(\mu;\gamma^p(H;E)).$$

For this one can apply Fubinis theorem and use the Kahane-Khintchine inequality. (For $p = 2$ one can prove it with equalities, otherwise two inequalities.)

8.3 Wiener integral (deterministic integrands)

All vector spaces in the following are real, H and $\mathcal{H}$ stand for fixed Hilbert spaces. We also fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

By a little abuse of notation for a Banach space E we will write $\langle x\,, x^*\rangle := x^*(x)$ for $x \in E,\ x^* \in E^*$. With $\mathcal{L}(H, E)$ we will denote the space of bounded linear operators from H to E and we will identify its adjoint operator as an element of $\mathcal{L}(E^*, H)$ by the identification $H \simeq H^*$.

Definition 8.37. *[vN07, Definition 6.1] An $E-$valued stochastic process $(X_i)_{i\in I}$ is called Gaussian if for all $n \geq 1$ and for all finitely many indices $i_1, \dots, i_n \in I$ we have that the E^n-valued random variable $(X_{i_1}, \dots, X_{i_n})$ is Gaussian.*

Here, an $E-$valued random variable X is Gaussian if for all $x^* \in E^*$ the random variables $x^*(X)$ are Gaussian.

Definition 8.38. *[vN07, Definition 6.8] Two processes $X = (X_i)_{i\in I},\ Y = (Y_i)_{i\in I}$ are called versions of each other if for all $i \in I$ we have $X_i = \tilde{X}_i,\ a.s.$*

Definition 8.39. *[vNVW07, Definition 2.1] An $\mathcal{H}-$isonormal process is a bounded linear map $W : \mathcal{H} \to \mathbb{L}^2(\Omega)$ such that:*

i) $\forall h \in \mathcal{H} : Wh$ *is Gaussian,*

ii) $\forall h_1,\ h_2 \in \mathcal{H} :\ \mathbb{E}[Wh_1 \cdot Wh_2] = \langle h_1\,, h_2\rangle.$

For every Hilbert space it is known that there actually exists an isonormal process.

The random variables $Wh,\ h \in \mathcal{H}$ are jointly normal distributed, since every linear combination

$$\sum_{i=1}^{n} c_i W h_i = W\left(\sum_{i=1}^{n} c_i h_i\right)$$

is normally distributed. So we have that if $h_1, \dots, h_n$ are orthogonal then $Wh_1, \dots, Wh_n$ are independent.

Definition 8.39 allows to generalize Brownian motions.

Definition 8.40. *[vN07, Definition 2.2] An $H-$cylindrical Brownian motion is an $\mathbb{L}^2(\mathbb{R}_+; H)-$isonormal process.*

We will denote an $H-$cylindrical Brownian motion by W_H and sometimes just by W. For $h \in H$ we define

$$W_H(t)h := W_H(\mathbb{1}_{(0,t)} \otimes h).$$

We then have that for each $h \in H$ $(W_H(t)h)_{t\geq 0}$ is a Brownian motion.

After this, we turn our attention to defining a stochastic integral for suitable $\Phi : \mathbb{R}_+ \to \mathcal{L}(H, E)$ w.r.t. W_H.

Definition 8.41. *see also [vNVW07, Section 4.1] Let W be an $H-$cylindrical Brownian motion. $\Phi : \mathbb{R}_+ \to H \otimes E$ is called an elementary function if it is a linear combination of functions like $\mathbb{1}_{(s,t]} \otimes (h \otimes x)$ where $0 \leq s < t < \infty$, $h \in H$ and $x \in E$. For a part of this elementary function we define the stochastic integral with respect to W as*

$$\int_0^\infty \mathbb{1}_{(s,t]} \otimes (h \otimes x)\, dW := W(\mathbb{1}_{(s,t]} \otimes h) \otimes x = (W(t)h - W(s)h) \otimes x \tag{8.3}$$

and extend this definition for elementary functions by linearity.

Remark 8.42. *With this we have a stochastic integral for maps $\Phi : \mathbb{R}_+ \to H \otimes E$ and get back a map looking like $\mathbb{L}^2(\Omega) \to E$.*

Any step function $\Phi : \mathbb{R}_+ \to \mathcal{L}(H, E)$ uniquely defines a bounded operator $R_\Phi \in \mathcal{L}(\mathbb{L}^2(\mathbb{R}_+; H), E)$ via

$$R_\Phi f := \int_0^\infty \Phi(t) f(t)\, dt, \quad f \in \mathbb{L}^2(\mathbb{R}_+; H).$$

Theorem 8.43 (Itô isometry)**.** *[vN07, Theorem 6.14] For all finite rank step functions $\Phi : \mathbb{R}_+ \to \mathcal{L}(H, E)$ we have $R_\Phi \in \gamma(\mathbb{L}^2(\mathbb{R}_+; H), E)$, $\int_0^T \Phi\, dW$ is a Gaussian random variable and*

$$\mathbb{E}\left[\left\| \int_0^\infty \Phi\, dW \right\|^2\right] = \|R_\Phi\|^2_{\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E)}.$$

With this isometry we can define the linear map

$$J^W : R_\Phi \mapsto \int_0^\infty \Phi\, dW.$$

This map uniquely extends to the isometric embedding

$$J^W : \gamma(\mathbb{L}^2(\mathbb{R}_+; H), E) \to \mathbb{L}^2(\Omega; E).$$

This allows to make the following definition.

Definition 8.44. *For an operator $R \in \gamma(\mathbb{L}^2(\mathbb{R}_+; H), E)$ we define the stochastic integral as $J^W(R)$.*

We still need to recognise the $\mathcal{L}(H, E)-$valued functions which the operator represents.

We define for $\Phi : \mathbb{R}_+ \to \mathcal{L}(H, E)$ and $x^* \in E^*$

$$\Phi^* x^* : \mathbb{R}_+ \to H : t \mapsto (\Phi^* x^*)(t) := \Phi^*(t) x^*$$

where $\Phi^*(t) := (\Phi(t))^*$.

Definition 8.45. *[vN07, Definition 6.15] $\Phi : \mathbb{R}_+ \to \mathcal{L}(H, E)$ is stochastically integrable w.r.t. W if there exists a sequence of finite rank step functions $\Phi_n : \mathbb{R}_+ \to \mathcal{L}(H, E)$ such that*

i) $\forall h \in H : \ \Phi_n h \to \Phi h$ in measure, and

ii) $\exists X$ an $E-$valued random variable, s.t. $\lim_{n\to\infty} \int_0^\infty \Phi_n \, dW = X$ in probability.

The stochastic integral of a stochastically integrable function Φ is then defined as the limit in probability

$$\int_0^\infty \Phi \, dW := \lim_{n\to\infty} \int_0^\infty \Phi_n \, dW.$$

Remark 8.46. *The condition ii) in the definition of convergence in probability is equivalent to convergence in $\mathbb{L}^p(\Omega, E)$ for some or equivalently all $p \in [1, \infty)$.*

To get a final characterization of stochastically integrable processes we state another definition concerning measurability suitable for this situation.

Definition 8.47. *$\Phi : \mathbb{R}_+ \to \mathcal{L}(H, E)$ is called scalarly measurable if the function $\Phi^* x^*$ is strongly measurable for all $x^* \in E^*$.*

Φ is called $H-$strongly measurable if for all $h \in H$ Φh is strongly measurable. If clear from context we will just call Φ strongly measurable.

With this we have the following theorem characterizing the definition of a process being stochastically integrable.

Theorem 8.48. *[vN07, Theorem 6.17] Let $\Phi : \mathbb{R}_+ \to \mathcal{L}(H, E)$ be $H-$strongly measurable. Then the following are equivalent:*

i) Φ is stochastically integrable w.r.t. W;

ii) $\Phi^* x^* \in \mathbb{L}^2(\mathbb{R}_+; H)$ for all $x^* \in E^*$ and there exists an $E-$valued random variable X such that for all $x^* \in E^*$ it holds

$$\langle X \, , x^* \rangle = \int_0^\infty \Phi^* x^* \, dW; \quad a.s.;$$

iii) $\Phi^* x^* \in \mathbb{L}^2(\mathbb{R}_+; H)\ \forall x^* \in E^*$, *and there exists an operator* $R \in \gamma(\mathbb{L}^2(\mathbb{R}_+; H), E)$, *s.t.* $\forall f \in \mathbb{L}^2(\mathbb{R}_+; H),\ x^* \in E^*$ *it holds that*

$$\langle Rf\,, x^*\rangle = \int_0^\infty \langle \Phi(t)f(t)\,, x^*\rangle\; dt.$$

If these equivalent conditions are satisfied, X and R are uniquely determined and we have $X = \int_0^\infty \Phi\; dW$ *almost surely and*

$$\mathbb{E}\left[\Big\| \int_0^\infty \Phi\; dW \Big\|^2\right] = \|R\|^2_{\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E)} = \|X\|^2_{\mathbb{L}^2(\Omega);E}.$$

With this we end the section by the following definition.

Definition 8.49. *In the situation of the theorem we say that* Φ *represents the operator* R.

8.4 UMD spaces

UMD spaces are essentially the spaces in which the stochastic integration theory developed in [vNVW07], [vVW15] and [vN07] works. First we remember the definition of martingale difference sequences which was also needed in Chapter 3.

Definition 8.50. *Let* $(M_n)_{n\in\mathbb{N}}$ *be an* $E-$*valued martingale. The sequence* $(d_n)_{n\in\mathbb{N}}$ *defined by* $d_n := M_n - M_{n-1}$ *where we set* $M_0 := 0$ *is called the martingale difference sequence associated with* $(M_n)_{n\in\mathbb{N}}$.

It is called an $\mathbb{L}^p-$*martingale difference sequence if* $(M_n)_{n\in\mathbb{N}}$ *is an* $\mathbb{L}^p-$*martingale.*

Having this, UMD spaces are defined as follows.

Definition 8.51. *[vVW15, Definition 5.2] A Banach space E is called an UMD space if for some* $p \in (1,\infty)$ *there is a constant* $\beta \geq 0$ *such that for all* $E-$*valued* $\mathbb{L}^p-$*martingale difference sequences* $(d_n)_{n\geq 1}$ *and all signs* $(\epsilon_n)_{n\geq 1}$ *one has*

$$\mathbb{E}\left[\Big\| \sum_{n=1}^N \epsilon_n d_n \Big\|^p\right] \leq \beta^p \mathbb{E}\left[\Big\| \sum_{n=1}^N d_n \Big\|^p\right], \quad \forall N \geq 1. \tag{8.4}$$

The least admissible constant will be denoted by $\beta_{p,E}$.

Remark 8.52. *Actually the existence of a* $p \in (1,\infty)$ *in the definition is already equivalent that it holds for all* $p \in (1,\infty)$. *See also [vN07, Section 12.2] for details.*

Some examples of UMD spaces are Hilbert spaces and $\mathbb{L}^p(\mu)$ for $p \in (1,\infty)$. Also E is an UMD space if and only if its dual E^* is an UMD space.

We end this section by a reverse estimate of martingale differences in UMD spaces. Namely one can apply the inequality in the definition of UMD spaces to the martingale difference sequence $(\epsilon_n d_n)_{n\geq 1}$ and get

$$\mathbb{E}\left[\|\sum_{n=1}^{N} d_n\|^p\right] \leq \beta_{p,E}^p \mathbb{E}\left[\|\sum_{n=1}^{N} \epsilon_n d_n\|^p\right], \quad \forall N \geq 1. \tag{8.5}$$

8.5 Itô integral

Here we will cite some introductions into stochastic integration using UMD spaces. For this sake we fix $p \in (1, \infty)$.

In many of the theorems a certain decoupling technique is necessary which gives an independent copy of the cylindrical Brownian motion on another probability space. With respect to the copy of the cylindrical Brownian motion one can estimate its $\mathbb{L}^p-$norms path-by-path w.r.t. Ω.

In order to formulate this in a theorem we consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and independent copies $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. We identify the filtrations $\mathcal{F}$ and $\tilde{\mathcal{F}}$ on $\Omega \times \tilde{\Omega}$ with $\mathcal{F} \times \{\varnothing, \tilde{\Omega}\}$ and $\{\varnothing, \Omega\} \times \tilde{\mathcal{F}}$. Similarly random variables ξ and $\tilde{\xi}$ on Ω and $\tilde{\Omega}$ are identified with random variables on $\Omega \times \tilde{\Omega}$ as $\xi(\omega, \tilde{\omega}) := \xi(\omega)$ and $\tilde{\xi}(\omega, \tilde{\omega}) := \tilde{\xi}(\tilde{\omega})$.

Theorem 8.53. *[vVW15, Theorem 5.4] Let E be an UMD space and let $p \in (1, \infty)$. Let $(\eta_n)_{n\geq 1}$ be an $\mathcal{F}-$adapted sequence of centered random variables in $\mathbb{L}^p(\Omega)$ such that for each $n \geq 1$, η_n is independent of $\mathcal{F}_{n-1}$. Let $(\tilde{\eta}_n)_{n\geq 1}$ be an independent $\tilde{\mathcal{F}}-$adapted copy of this sequence in $\mathbb{L}^p(\tilde{\Omega})$. Finally let $(\nu_n)_{n\geq 1}$ be an $\mathcal{F}-$predictable sequence in $\mathbb{L}^\infty(\Omega; E)$. Then, for all $N \geq 1$,*

$$\frac{1}{\beta_{p,E}^p}\mathbb{E}\left[\tilde{\mathbb{E}}\left[\|\sum_{n=1}^{N} \nu_n \tilde{\eta}_n\|^p\right]\right] \leq \mathbb{E}\left[\tilde{\mathbb{E}}\left[\|\sum_{n=1}^{N} \nu_n \eta_n\|^p\right]\right] \leq \beta_{p,E}^p \mathbb{E}\left[\tilde{\mathbb{E}}\left[\|\sum_{n=1}^{N} \nu_n \tilde{\eta}_n\|^p\right]\right].$$

Having looked at this theorem one starts defining a stochastic integral as always for some kind of elementary process. In this setting those are introduced in the following definition.

Definition 8.54. *[vNVW07, Section 2.4] A process $\Phi : \mathbb{R}_+ \times \Omega \to \mathcal{L}(H, E)$ is said to be elementary adapated to the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ if it is of the form*

$$\Phi(t, \omega) = \sum_{n=1}^{N} \sum_{m=1}^{M} \mathbb{1}_{(t_{n-1}, t_n] \times A_{mn}}(t, \omega) \sum_{k=1}^{K} h_k \otimes x_{kmn}, \tag{8.6}$$

where $0 \leq t_0 < \ldots < t_n$ and the sets $A_{1n}, \ldots, A_{Mn} \in \mathcal{F}_{t_{n-1}}$ are disjoint for each n and the vectors $h_1, \ldots, h_K \in H$ are orthonormal. It is also called a finite rank *adapted process w.r.t.* $\mathbb{F}$.

In the following we also fix an $H-$cylindrical Brownian motion W on $(\Omega, \mathcal{F}, \mathbb{P})$ adapted to $\mathbb{F}$ in the sense that the increments $W(t)h - W(s)h$ are independent of $\mathcal{F}_s$ for all $h \in H$ and $s < t$.

Now we can define the stochastic integral w.r.t. W of elementary adapted functions, also called random simple functions as in [vN07, Section 13.2].

Definition 8.55. *For Φ as above we define the stochastic integral w.r.t. W via*

$$\int_0^\infty \mathbb{1}_{(s,t]\times F}(h \otimes x)\ dW := \mathbb{1}_F W(\mathbb{1}_{(s,t]} \otimes h) \otimes x$$

and extend it via linearity.

Remark 8.56. *Written out the stochastic integral for Φ looks like*

$$\int_0^\infty \Phi(t)\ dW(t) = \sum_{m=1}^{M} \sum_{n=1}^{N} \mathbb{1}_{A_{mn}} \sum_{k=1}^{K} (W(t_n)h_k - W(t_{n-1})h_k) x_{kmn}.$$

We get that $\int_0^\infty \Phi\ dW \in \mathbb{L}^p(\Omega, \mathcal{F}_\infty; E)$ for all $p \in [1, \infty)$ and

$$\mathbb{E}\left[\int_0^\infty \Phi(t)\ dW(t)\right] = 0.$$

For $\omega \in \Omega$ the trajectory $t \mapsto \Phi_\omega(t) := \Phi(t, \omega)$ is a finite rank step function and therefore defines an element $R_{\Phi_\omega} \in \gamma(\mathbb{L}^2(\mathbb{R}_+; H), E)$. Therefore, we obtain a random variable

$$R_\Phi : \Omega \to \gamma(\mathbb{L}^2(\mathbb{R}_+; H), E).$$

With this at hand one can prove the following Itô isomorphism.

Theorem 8.57 (Itô isomorphism). *[vVW15, Theorem 5.5] Let E be an UMD space and $p \in (1, \infty)$. For all adapted elementary processes $\Phi : \mathbb{R}_+ \times \Omega \to H \otimes E$ we have*

$$\frac{1}{\beta_{p,E}} \|\Phi\|_{\mathbb{L}^p(\Omega;\gamma^p(\mathbb{L}^2(\mathbb{R}_+;H),E))} \leq \left\| \int_0^\infty \Phi\ dW \right\|_{\mathbb{L}^p(\Omega;E)} \leq \beta_{p,E} \|\Phi\|_{\mathbb{L}^p(\Omega;\gamma^p(\mathbb{L}^2(\mathbb{R}_+;H),E))}. \tag{8.7}$$

Remark 8.58. *The isomorphism can be used to obtain an equivalence of norms.*

One just applies Doob's inequality for $p \in (1,\infty)$ to obtain

$$\frac{1}{\beta_{p,E}}\|\Phi\|_{\mathbb{L}^p(\Omega;\gamma^p(\mathbb{L}^2(\mathbb{R}_+;H),E))} \leq \left(\mathbb{E}\left[\sup_{t\geq 0}\left\|\int_0^t \Phi\, dW\right\|^p\right]\right)^{1/p} \leq \tag{8.8}$$

$$\leq \frac{p}{p-1}\beta_{p,E}\|\Phi\|_{\mathbb{L}^p(\Omega;\gamma^p(\mathbb{L}^2(\mathbb{R}_+;H),E))}. \tag{8.9}$$

To make the step towards more general processes we will define adapted processes in this setting.

Definition 8.59. *[vN07, Definition 13.3] A random variable $R \in \mathbb{L}^p(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E))$ is called adapted if it is in the closure in $\mathbb{L}^p(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E))$ of the finite rank adapted processes.*

This closed subspace we will denote by $\mathbb{L}^p_{\mathbb{F}}(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E))$.

Remark 8.60. *We see how the stochastic integral defines an isomorphic embedding*

$$I : \mathbb{L}^p_{\mathbb{F}}(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E)) \to \mathbb{L}^p(\Omega;E). \tag{8.10}$$

The stochastic integral defines also an isomorphic embedding of $\mathbb{L}^p_{\mathbb{F}}(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E))$ into $\mathbb{L}^p(\Omega;C_b(\mathbb{R}_+;E))$.

With this one can make the following definition of a process being stochastically integrable.

Definition 8.61. *[vN07, Definition 13.4] Fix $p \in (1,\infty)$. $\Phi : \mathbb{R}_+ \times \Omega \to \mathcal{L}(H,E)$ is called $\mathbb{L}^p-$stochastically integrable w.r.t. W if there exists a sequence of finite rank adapted step processes Φ_n s.t.*

i) $\forall h \in H : \Phi_n h \to \Phi h$ in measure, and

ii) $\exists X \in \mathbb{L}^p(\Omega;E)$ such that $\int_0^\infty \Phi_n\, dW \to X$ in $\mathbb{L}^p(\Omega;E)$.

The $\mathbb{L}^p-$stochastic integral of Φ w.r.t. W is then defined by

$$\int_0^\infty \Phi\, dW := \lim_{n\to\infty}\int_0^\infty \Phi_n\, dW \quad \textit{in } \mathbb{L}^p(\Omega;E).$$

One can also look at the integral process which turns out to be a martingale again and for which we can as usually estimate its running maximum.

Theorem 8.62. *[vN07, Theorem 13.5] Let $p \in (1,\infty)$. If Φ is $\mathbb{L}^p-$stochastically integrable w.r.t. W, then the stochastic integral process*

$$\left(\int_0^t \Phi\, dW\right)_{t\geq 0}$$

is an $E-$*valued martingale which has a continuous version satisfying*

$$\mathbb{E}\left[\sup_{t\geq 0}\|\int_0^t \Phi\, dW\|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[\|\int_0^\infty \Phi\, dW\|^p\right].$$

We will define strongly measurable processes in this setting and give a characterization if they are $\mathbb{L}^p$ stochastically integrable with respect to a cylindrical Brownian motion.

Definition 8.63. *[vN07, Definition 13.6]* $\Phi : \mathbb{R}_+ \times \Omega \to \mathcal{L}(H, E)$ *is called an* $H-$*strongly measurable process if for all* $h \in H$ *the map*

$$\Phi h : \mathbb{R}_+ \times \Omega \to E, (t, \omega) \mapsto \Phi h(t, \omega) := \Phi(t, \omega)h$$

is strongly measurable.

Theorem 8.64. *[vN07, Theorem 13.7] For an* $H-$*strongly measurable adapted process* Φ *the following are equivalent:*

i) Φ *is* $\mathbb{L}^p-$*stochastically integrable w.r.t.* W*;*

ii) $\Phi^* x^* \in \mathbb{L}^p(\Omega; \mathbb{L}^2(\mathbb{R}_+; H))$ *for all* $x^* \in E^*$ *and there exists* $X \in \mathbb{L}^p(\Omega; E)$ *s.t. for all* $x^* \in E^*$

$$\langle X\,, x^*\rangle = \int_0^\infty \Phi^* x^*\, dW \quad \text{in } \mathbb{L}^p(\Omega).$$

iii) $\Phi^* x^* \in \mathbb{L}^p(\Omega; \mathbb{L}^2(\mathbb{R}_+; H))$ *for all* $x^* \in E^*$ *and there exists* $R \in \mathbb{L}^p(\Omega; \gamma(\mathbb{L}^2(\mathbb{R}_+; H), E))$ *s.t. for all* $f \in \mathbb{L}^2(\mathbb{R}_+; H),\ x^* \in E^*$

$$\langle Rf\,, x^*\rangle = \int_0^\infty \langle \Phi(t)f(t)\,, x^*\rangle\, dt \quad \text{in } \mathbb{L}^p(\Omega).$$

In this case, X *and* R *are uniquely determined and we have* $X = \int_0^\infty \Phi\, dW$ *and*

$$\mathbb{E}\left[\|\int_0^\infty \Phi\, dW\|^p\right] \simeq_{p,E} \mathbb{E}\left[\|R\|^p_{\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E)}\right].$$

Moreover, $R \in \mathbb{L}^p_{\mathbb{F}}(\Omega; \gamma(\mathbb{L}^2(\mathbb{R}_+; H), E))$*, which means* R *is adapted.*

With this we can even characterize processes in $\mathbb{L}^p(\Omega, \mathcal{F}^W_\infty; E)$ via their expectation and a stochastic integral.

Theorem 8.65. *[vN07, Theorem 13.9] For* $X \in \mathbb{L}^p(\Omega, \mathcal{F}^W_\infty; E)$ *there exists a unique* $R \in \mathbb{L}^p_{\mathbb{F}^W}(\Omega; \gamma(\mathbb{L}^2(\mathbb{R}_+; H), E))$ *such that*

$$X = \mathbb{E}[X] + J^W(R).$$

With this we obtain an isomorphism of Banach spaces.

$$J^W : \mathbb{L}^p_{\mathbb{F}^W}(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E)) \simeq \mathbb{L}^p_0(\Omega,\mathcal{F}^W_\infty;E),$$

where $\mathbb{L}^p_0(\Omega,\mathcal{F}^W_\infty;E)$ denotes the closed subspace of $\mathbb{L}^p(\Omega,\mathcal{F}^W_\infty;E)$ having all elements with expectation 0.

Next, we have a theorem which tells when a process is integrable.

Theorem 8.66. *Let $p \in [0,\infty)$. For $\Phi \in \mathbb{L}^p(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E))$ the following are equivalent:*

i) $\Phi \in \mathbb{L}^p_{\mathbb{F}}(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E))$,

ii) $\langle \Phi(\mathbb{1}_{[0,t]}f)\,, x^*\rangle \in \mathbb{L}^p(\Omega)$ *is* $\mathcal{F}_t-$*measurable* $\forall t \in \mathbb{R}_+,\ f \in \mathbb{L}^2(\mathbb{R}_+;H),\ x^* \in E^*$.

This tells that if $\Phi : \mathbb{R}_+ \times \Omega \to \mathcal{L}(H,E)$ is $H-$strongly measurable and adapted, adapted here meaning that for all $h \in H$ the process $\Phi h : \mathbb{R}_+ \times \Omega \to E$ is adapted, then $\Phi \in \mathbb{L}^p(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E))$ already implies $\Phi \in \mathbb{L}^p_{\mathbb{F}}(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E))$.

Looking again at the integral process it can be shown is that the stochastic integral $I : \Phi \mapsto \int_0^{\cdot} \Phi\, dW$ uniquely extends to a continuous linear mapping

$$I : \mathbb{L}^0_{\mathbb{F}}(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E)) \to \mathbb{L}^0(\Omega;C_b(\mathbb{R}_+;E))) \tag{8.11}$$

The integral process is defined as follows.

Definition 8.67. *For $\Phi \in \mathbb{L}^0_{\mathbb{F}}(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),E))$ we define $I\Phi$ to be the stochastic integral of Φ w.r.t. W and define the integral process*

$$\int_0^t \Phi\, dW := (I\phi)(t), \quad t \geq 0. \tag{8.12}$$

We state the following fundamental theorem.

Theorem 8.68. *[vVW15, Theorem 5.8] Let E be an UMD Banach space. Let $\Phi : \mathbb{R}_+ \times \Omega \to \mathcal{L}(H,E)$ be an $H-$strongly measurable adapted process such that $\Phi^* x^* \in \mathbb{L}^0(\Omega;\mathbb{L}^2(\mathbb{R}_+;H))$ for all $x^* \in E^*$. Let $\zeta : \mathbb{R}_+ \times \Omega \to E$ be a process whose paths are almost surely bounded. If for all $x^* \in E^*$, a.s. it holds that*

$$\int_0^t \Phi^* x^*\, dW = \langle \zeta_t\,, x^*\rangle, \quad t \geq 0,$$

then Φ represents an element in $\mathbb{L}^0_{\mathbb{F}}(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H))$, and a.s. it holds that

$$\int_0^t \Phi\, dW = \zeta_t, \quad t \geq 0.$$

Moreover, ζ is a local martingale with continuous paths a.s.

We also state a theorem which tells that UMD spaces are useful and shows why this stochastic integration theory works there.

Theorem 8.69. *Let E be a Banach space and suppose for a $p \in (1,\infty)$ there exist constants c_p and C_p such that for all adapted elementary processes $\Phi : \mathbb{R}_+ \times \Omega \to E$*

$$\frac{1}{c_p}\|\Phi\|_{\mathbb{L}^p(\Omega;\gamma^p(\mathbb{L}^2(\mathbb{R}_+),E))} \leq \Big\| \int_0^\infty \Phi \; dW \Big\|_{\mathbb{L}^p(\Omega;E)} \leq C_p \|\Phi\|_{\mathbb{L}^p(\Omega;\gamma^p(\mathbb{L}^2(\mathbb{R}_+),E))}.$$

Then E is an UMD space with constant $\beta_{p,E} \leq c_p C_p$.

We want to finish this section by stating a version of Itô's formula in this setting. For this sake take Banach spaces $X,\ Y,\ Z$ and let $(h_n)_{n\geq 1}$ be an orthonormal basis of H. Let $R \in \gamma(H,X),\ S \in \gamma(H,Y)$ and $T \in \mathcal{L}(X, \mathcal{L}(Y,Z))$, then the sum

$$\mathrm{tr}_{R,S} T := \sum_{n\geq 1} (TRh_n)(Sh_n)$$

converges in Z and does not depend on the choice of the orthonormal basis. It also holds that

$$\|\mathrm{tr}_{R,S} T\| \leq \|T\| \|R\|_{\gamma(H,X)} \|S\|_{\gamma(H,Y)}.$$

If $X = Y$, we also write $\mathrm{tr}_R := \mathrm{tr}_{R,R}$.

Theorem 8.70 (Itô formula)**.** *[vVW15, Proposition 5.11] Let X and Y be UMD spaces. Assume $f : \mathbb{R}_+ \times X \to Y$ is $C^{1,2}$ on every bounded interval. Let $\phi : \mathbb{R}_+ \times \Omega \to \mathcal{L}(H,X)$ be $H-$strongly measurable and adapted and assume that ϕ locally defines an element of $\mathbb{L}^0(\Omega;\gamma(\mathbb{L}^2(\mathbb{R}_+;H),X)) \cap \mathbb{L}^0(\Omega;\mathbb{L}^2(\mathbb{R}_+;\gamma(H,X)))$. Let $\psi : \mathbb{R}_+ \times \Omega \to X$ be strongly measurable and adapted with locally integrable paths a.s. Let $\xi : \Omega \to X$ be strongly $\mathcal{F}_0-$measurable. Define $\zeta : \mathbb{R}_+ \times \Omega \to X$ by*

$$\zeta = \xi + \int_0^\cdot \psi_s \; ds + \int_0^\cdot \phi_s \; dW_s.$$

Then $s \mapsto D_2 f(s,\zeta_s)\phi_s$ is stochastically integrable and a.s. we have for all $t \geq 0$,

$$\begin{aligned} f(t,\zeta_t) - f(0,\zeta_0) &= \int_0^t D_1 f(s,\zeta_s) \; ds + \int_0^t D_2 f(s,\zeta_s)\psi_s \; ds \\ &\quad + \int_0^t D_2 f(s,\zeta_s)\phi_s \; dW_s + \frac{1}{2}\int_0^t tr_{\phi_s}(D_2^2 f(s,\zeta_s)) \; ds. \end{aligned}$$

The first two integrals and the last one are a.s. defined as Bochner integrals.

Beyond the Itô formula one can define some sort of quadratic variation and also arrive at Burkholder-Davis-Gundy inequalities. However, we omit this here and refer the interested reader to the references used in this section.

8.6 Stochastic Integration in Martingale Type 2 Spaces

As last section we look into martingale type 2 spaces and how they play a role in stochastic integration. As the sections before we will again base most of this tour on [vVW15], [vNVW07] and [Pro72].

We begin with the definition of a Rademacher sequence.

Definition 8.71. *A Rademacher sequence is a sequence $(r_n)_{n\geq 1}$ of independent random variables, s.t. $\mathbb{P}[r_n = 1] = \mathbb{P}[r_n = -1] = \frac{1}{2}$.*

With this we define generally type p spaces for $p \in [1, 2]$.

Definition 8.72. *[vVW15, Definition 4.1] Let $p \in [1, 2]$. A Banach space E has type p if there exists a constant $\tau \geq 0$ such that for all finite sequences $(x_n)_{n=1}^N$ in E it holds that*

$$\mathbb{E}\left[\Big\|\sum_{n=1}^{N} r_n x_n\Big\|^p\right] \leq \tau^p \sum_{n=1}^{N} \|x_n\|^p.$$

The least admissible constant is denoted by $\tau_{p,E}$.

In many proofs some randomisation identity is important which we want to sketch here. For a sequence of independent symmetric random variables $(\xi_n)_{n\geq 1}$ in $\mathbb{L}^p(\Omega; E)$ and an independent Rademacher sequence $(\tilde{r}_n)_{n\geq 1}$ on another probability space $(\tilde{\Omega}, \tilde{\mathbb{P}})$, we have

$$\mathbb{E}\left[\Big\|\sum_{n=1}^{N} \xi_n\Big\|^p\right] = \mathbb{E}\left[\tilde{\mathbb{E}}\left[\Big\|\sum_{n=1}^{n} \tilde{r}_n \xi_n\Big\|^p\right]\right], \quad \forall N \geq 1.$$

Noticing that for fixed $\tilde{\omega} \in \tilde{\Omega}$ $(\xi_n)_{n\geq 1}$ has the same distribution as $(\tilde{r}_n(\tilde{\omega})\xi_n)_{n\geq 1}$ we can use Fubinis theorem to prove this by looking at

$$\begin{aligned}
\mathbb{E}\left[\tilde{\mathbb{E}}\left[\|\sum_{n=1}^{n}\tilde{r}_n\xi_n\|^p\right]\right] &= \int_\Omega\int_{\tilde{\Omega}}\|\sum_{n=1}^{N}\tilde{r}_n(\tilde{\omega})\xi_n(\omega)\|^p\ d\tilde{\mathbb{P}}(\tilde{\omega})\ d\mathbb{P}(\omega) \overset{Fubini}{=} \\
&= \int_{\tilde{\Omega}}\int_\Omega\|\sum_{n=1}^{N}\tilde{r}_n(\tilde{\omega})\xi_n(\omega)\|^p\ d\mathbb{P}(\omega)\ d\tilde{\mathbb{P}}(\tilde{\omega}) = \\
&= \int_{\tilde{\Omega}}\mathbb{E}\left[\|\sum_{n=1}^{N}\xi_n\|^p\right]\ d\tilde{\mathbb{P}}(\tilde{\omega}) = \mathbb{E}\left[\|\sum_{n=1}^{N}\xi_n\|^p\right].
\end{aligned}$$

In a next step one can also define a stochastic integral for simple functions in this section as in Section 8.3.

With this we have a stochastic integral for maps $\phi:\mathbb{R}_+\to H\otimes E$ and get back a map looking like $\mathbb{L}^2(\Omega)\to E$.

Now we turn to also defining a stochastic integral for random integrands. The important part of the definition is for $p=2$.

Definition 8.73. *[vVW15, Definition 4.4] Let $p\in[1,2]$ and E a Banach space. It has martingale type p if there exists $\mu\geq 0$ such that for all finite $E-$valued martingale difference sequences $(d_n)_{n=1}^N$ one has*

$$\mathbb{E}\left[\|\sum_{n=1}^{N}d_n\|^p\right]\leq\mu^p\sum_{n=1}^{N}\mathbb{E}[\|d_n\|^p]. \tag{8.13}$$

The least admissible constant is denoted by $\mu_{p,E}$.

Remark 8.74. *Since every Gaussian sequence is a martingale difference sequence, we get that every Banach space with martingale type p also has type p.*

In the following we see an $H-$cylindrical Brownian motion W as given and denote by $(\mathcal{F}_t)_{t\geq 0}$ its generated filtration, i.e. $\mathcal{F}_t=\sigma\left(W(f)\mid f\in\mathbb{L}^2([0,t];H)\right)$.

An integral w.r.t. an $H-$cylindrical Brownian motion for adapted elementary processes can also be defined as in Section 8.5.

For the integral we get the following inequality.

Theorem 8.75. *[vVW15, Theorem 4.6] Let the Banach space E have type 2 and let ϕ be an adapted elementary process. Then we have*

$$\mathbb{E}\left[\|\int_0^\infty\phi\ dW\|\right]^2\leq\mu_{p,E}^2\mathbb{E}\left[\int_0^\infty\|\phi_t\|_{\gamma(H,E)}^2\ dt\right].$$

Remark 8.76. *With Doob's maximal inequality, the result improves to*

$$\mathbb{E}\left[\sup_{t\geq 0}\Big\|\int_0^t \phi\, dW\Big\|^2\right] \leq 4\mu_{p,E}^2\mathbb{E}\left[\int_0^\infty \|\phi_t\|_{\gamma(H,E)}^2\, dt\right].$$

As usual the stochastic integral can be extended by a density argument. It can be extended to all progressively measurable processes $\phi : \mathbb{R}_+ \times \Omega \to \gamma(H,E)$ that satisfy

$$\mathbb{E}\left[\int_0^\infty \|\phi_t\|_{\gamma(H,E)}^2\, dt\right] < \infty.$$

Then the process $t \mapsto \int_0^t \phi\, dW$ turns out to be a continuous martingale. With this result, the usual stopping techniques apply to extend the stochastic integral towards all progressive measurable processes satisfying

$$\int_0^\infty \|\phi_t\|_{\gamma(H,E)}^2\, dt < \infty, \quad a.s.$$

We close with the following Burkholder inequality.

Theorem 8.77. *[vVW15, Theorem 4.7] Let E have martingale type* 2*. Then for all* $p \in (0,\infty)$ *there exists a constant* $C_{p,E}$ *such that for all strongly measurable adapted processes* $\phi : \mathbb{R}_+ \times \Omega \to \gamma(H,E)$ *we have*

$$\mathbb{E}\left[\sup_{t\geq 0}\Big\|\int_0^t \phi\, dW\Big\|^p\right] \leq C_{p,E}^p\|\phi\|_{\mathbb{L}^p(\Omega;\mathbb{L}^2(\mathbb{R}_+,\gamma(H,E)))}^p.$$

For more details on martingale type p spaces we refer the reader especially to [vVW15].

Bibliography

[Bic79] Klaus Bichteler. Stochastic integrators. *Bull. Am. Math. Soc., New Ser.*, 1:761–765, 1979.

[BKW17] Martin Blümlinger, Michael Kaltenbäck, and Harald Woracek. Funktionalanalysis, February 2017.

[BS12] Mathias Beiglboeck and Pietro Siorpaes. A simple proof of the bichteler-dellacherie theorem. 01 2012.

[BSV11] Mathias Beiglböck, Walter Schachermayer, and Bezirgen Veliyev. A direct proof of the bichteler–dellacherie theorem and connections to arbitrage. *Ann. Probab.*, 39(6):2424–2440, 11 2011.

[CT15] Christa Cuchiero and Josef Teichmann. A convergence result for the Emery topology and a variant of the proof of the fundamental theorem of asset pricing. *Finance Stoch.*, 19(4):743–761, 2015.

[Els18] Jürgen Elstrodt. *Maß- und Integrationstheorie*. Heidelberg: Springer Spektrum, 8th enlarged and updated edition edition, 2018.

[It44] Kiyosi It. Stochastic integral. *Proceedings of the Imperial Academy*, 20(8):519 – 524, 1944.

[Kal02] Olav Kallenberg. *Foundations of modern probability. 2nd ed.* New York, NY: Springer, 2nd ed. edition, 2002.

[Low09] George Lowther. Cadlag modifications. Almost Sure website, December 2009.

[Low10] George Lowther. Preservation of the local martingale property. Almost Sure website, March 2010.

[Low12] George Lowther. Rao's quasimartingale decomposition. Almost Sure website, June 2012.

[Mey72] Paul-Andre Meyer. *Martingales and stochastic integrals. I*, volume 284. Springer, Cham, 1972.

[Pis11] Gilles Pisier. Martingales in banach spaces (in connection with type and cotype). lecture notes, February 2011.

[Pro72] Chapter 1 probability theory in banach spaces: An introductory survey. In A.T. Bharucha-Reid, editor, *Random Integral Equations*, volume 96 of *Mathematics in Science and Engineering*, pages 7 – 63. Elsevier, 1972.

[Pro04] Philip E. Protter. *Stochastic integration and differential equations. 2nd ed.*, volume 21. Berlin: Springer, 2nd ed. edition, 2004.

[Sch18] Uwe Schmock. Stochastic analysis for financial and actuarial mathematics, December 2018.

[Str84] Christophe Stricker. Caractérisation des semimartingales. *Séminaire de probabilités de Strasbourg*, 18:148–153, 1984.

[Tei19] Josef Teichmann. Foundations of martingale theory and stochastic calculus from a finance perspective. lecture notes ETH Zurich, October 2019.

[vN07] Jan van Neerven. Stochastic evolution equations. ISEM lecture notes, 2007.

[vNVW07] J. M. A. M. van Neerven, M. C. Veraar, and L. Weis. Stochastic integration in umd banach spaces. *Ann. Probab.*, 35(4):1438–1478, 07 2007.

[vVW15] Jan van Neerven, Mark Veraar, and Lutz Weis. Stochastic integration in Banach spaces – a survey. In *Stochastic analysis: a series of lectures. Centre Interfacultaire Bernoulli, January – June 2012, École Polytechnique Fédérale Lausanne, Switzerland*, pages 297–332. Basel: Birkhäuser/Springer, 2015.

[Yan80] Jia-An Yan. Caractérisation d'une classe d'ensembles convexes de l^1 ou h^1. *Séminaire de probabilités de Strasbourg*, 14:220–222, 1980.