

Progressive Stochastic Processes and an Application to the Itô Integral

Svenja Kaden¹ and Jürgen Potthoff²

¹*Gerling Insurances, Cologne, Germany*

²*Universität Mannheim, D-68131 Mannheim, Germany*

Abstract An elementary proof of the theorem of Chung–Doob–Meyer on the existence of a progressively measurable modification of a measurable adapted process is given. It is shown how this result can be applied to the construction of the Itô integral with respect to a Brownian motion.

Keywords: Itô integral; Adapted stochastic process; Progressively measurable stochastic process; Approximation of stochastic processes.

1. Introduction and Main Results

Most students who begin to learn stochastic integration are primarily interested in Itô's stochastic integral of a stochastic process with respect to a Brownian motion. This is due to the fact that such integrals have an enormous number of applications in many fields, including the natural sciences, engineering, social sciences, economy, finance, and almost all branches of mathematics. On the other hand, in most of the common literature on stochastic integration which focusses on the Itô integral with respect to Brownian motion the details of a certain argument necessary for the construction of the integral (see below) are not given or only hinted at. The purpose of the present paper is to provide the beginner with these details in a rather self-consistent and elementary way.

Consider a standard one-dimensional Brownian motion $B = (B_t, t \in \mathbb{R}_+)$ on some probability space (Ω, \mathcal{A}, P) , and suppose that we are given a stochastic process $X = (X_t, t \in \mathbb{R}_+)$ with values in the extended reals $\overline{\mathbb{R}}$. One is interested to construct the Itô integral $\int_a^b X_t dB_t$, $0 \leq a < b \leq +\infty$, of X with respect to B . The assumptions which are typically made in the literature are that

- (i) X is adapted to the filtration $\mathcal{F} = (\mathcal{F}_t, t \in \mathbb{R}_+)$ generated by the Brownian motion, and
- (ii) X belongs to $\mathcal{L}^2(\Omega \times [a, b], \mathcal{A} \otimes \mathcal{B}([a, b]), P \otimes \lambda)$.

($\mathcal{B}([a, b])$ stands for the Borel σ -algebra over $[a, b]$, and λ for the Lebesgue measure.) Then one tries to construct $\int_a^b X_t dB_t$ as the $\mathcal{L}^2(\Omega)$ -limit of approximations of Riemann-type

$$\sum_k X_{t_k}^n (B_{t_{k+1}} - B_{t_k}), \tag{1.1}$$

where the t_k define a partition of $[a, b]$, and X^n is a suitable approximation to X . (For other approaches, aiming at more general stochastic integrals, we refer the interested reader also to, e.g., [CW 83], [Pr 90], [RY 91], [WW 90] and the references cited there.)

The key to all constructions of the Itô integral is, of course, the Itô isometry. The approximation X^n of X above (in $\mathcal{L}^2(\Omega \times \mathbb{R}_+)$) is usually obtained through a sequence of steps as follows. First one makes X bounded by “cutting it off” from below and above, e.g.,

$$\widehat{X}^n := (X \vee (-n)) \wedge n.$$

Clearly, \widehat{X}^n is still adapted to \mathcal{F} . The next step is to make \widehat{X}^n pathwise continuous. This is most of the time done by smoothing through an integration procedure like

$$X_t^n := n \int_{t-1/n}^t \widehat{X}_s^n ds. \tag{1.2}$$

(We have not been very careful about what happens for t and a near zero, cf. Section 3 for a more careful discussion.) However, here we meet the problem that the last operation does in general *not* preserve the adaptedness of \widehat{X}^n to \mathcal{F} because it involves uncountable many values of the time parameter. But then we cannot use the Itô isometry for the sum in (1.1), and our construction of the Itô integral breaks down.

One remedy is to assume in addition that X is separable, e.g. [Fr 75], but this is not done in most of the literature and therefore generates some incompatibility. Another possibility, which can for example be found in the book by Dynkin [Dy 65], is to make a slightly stronger measurability assumption than adaptedness, namely, to suppose that X is *progressive* with respect to \mathcal{F} :

Definition 1 Let $X = (X_t, t \in \mathbb{R}_+)$ be a stochastic process on a probability space (Ω, \mathcal{A}, P) which is equipped with a filtration $\mathcal{F} = (\mathcal{F}_t, t \in \mathbb{R}_+)$ of sub- σ -algebras \mathcal{F}_t of \mathcal{A} . X is called *progressively measurable with respect to \mathcal{F}* or *\mathcal{F} -progressive*, if for every $t \in \mathbb{R}_+$ the restriction of X to $\Omega \times [0, t]$ is measurable with respect to the product σ -algebra $\mathcal{F}_t \otimes \mathcal{B}([0, t])$.

If the filtration \mathcal{F} is understood from the context, we shall also simply say that X is *progressively measurable* or just X is *progressive*. It follows from a fact in elementary measure theory, that every process X which is progressively measurable with respect to a filtration \mathcal{F} is also adapted to \mathcal{F} . In general, the converse is false.

We return to the construction of the Itô integral of X , and assume for the moment that X is progressive with respect to the filtration \mathcal{F} generated by the Brownian motion B . Then it follows from the “first part” of the theorem of Fubini–Tonelli that X^n as defined by (1.2) is actually adapted to \mathcal{F} . Therefore we can proceed with the construction of the Itô integral in the usual way: X^n is approximated in $\mathcal{L}^2(\Omega \times \mathbb{R}_+)$ by a step function with steps “sticking out into the future”. For this process one defines the Itô integral as in (1.1), shows for it the Itô isometry and finally takes appropriate limits in $\mathcal{L}^2(\Omega)$ to reverse the approximation steps done above. The details can be found in every book on Itô integration and are omitted here.

As mentioned above, our temporary assumption that X is progressive is stronger than the commonly made assumption that X is adapted. The essential clue towards the resolution of this apparent inconsistency is the following theorem of Chung, Doob [CD

64] and Meyer [Me 66]. Before stating the result, we recall that the process X is called *measurable* if the mapping $X : \Omega \times \mathbb{R}_+ \rightarrow \overline{\mathbb{R}}$ is measurable.

Theorem 1 Let X be an $\overline{\mathbb{R}}$ -valued measurable stochastic process adapted to a filtration \mathcal{F} . Then it has a measurable \mathcal{F} -progressive modification.

This theorem has been proved by Chung and Doob in [CD 64] under the additional assumption that X is separable. In the above form, i.e., without the separability assumption, Theorem 1 is stated in the book [Me 66] by Meyer (which is most of the time quoted for this result). The argument in [Me 66] has been judged by Karatzas and Shreve “lengthy and rather demanding” [KS 88, p. 5]. Actually, the proof in [Me 66] has a gap, which has been fixed in [DM 78]. The proof in [DM 78] makes indeed use of rather heavy machinery: the Dunford–Pettis compactness theorem and a version of the Eberlein–Šmulian theorem. In [KS 88] then the progressive modification of X is used to construct its Itô-integral (under the additional assumption that the underlying probability space is complete).

The proof of Theorem 1 in [DM 78] (cf. Theorem IV.30 there) has essentially two parts. The first is to establish that a measurable process X has an approximation by processes of the form

$$X_n(t) = \sum_k H_{n,k} 1_{A_{n,k}}(t), \quad (1.3)$$

where the $A_{n,k}$ are Borel sets which form a partition of the time parameter domain, and the $H_{n,k}$ are random variables. In order to prove this, one combines elementary measure theory with the monotone class theorem. In the second part, one replaces the $H_{n,k}$ (which are close to X on $A_{n,k}$) by values of X , so that the expressions $H_{n,k} 1_{A_{n,k}}$ become progressive processes. For example, if $A_{n,k}$ is an interval of the form $[s_{n,k}, t_{n,k}]$ one may choose $H_{n,k} = X_{s_{n,k}}$, because the fact that $X_{s_{n,k}}$ is $\mathcal{F}_{s_{n,k}}$ -measurable implies that

$$(\omega, t) \mapsto X_{s_{n,k}}(\omega) 1_{[s_{n,k}, t_{n,k}]}(t)$$

is progressive. For a general Borel set $A_{n,k}$ which does not have a minimum the situation becomes more involved. In this case, Dellacherie and Meyer replace $H_{n,k}$ by a weak $\mathcal{L}^1(P)$ -limit of a sequence $(X_{t_i}, i \in \mathbb{N})$, where the $t_i \in A_{n,k}$ decrease to the infimum of $A_{n,k}$. Once one has the new progressive processes, say $(\tilde{X}_n, n \in \mathbb{N})$, one may choose

$$\tilde{X} = \limsup_n \tilde{X}_n$$

to obtain a progressive modification of X .

In the present paper we carry this proof out in detail. However, instead of the weak $\mathcal{L}^1(P)$ -convergence mentioned above, we use weak $\mathcal{L}^2(P)$ -convergence. This has the advantage that we can avoid the Dunford–Pettis and Eberlein–Šmulian theorems and use instead a very simple result, namely the classical “Theorem of Choice”, which, e.g., can be found in [RN 55]. Its proof uses only very elementary Hilbert space theory (the projection theorem and the Riesz representation theorem), and this way, the proof of the desired result becomes in our opinion rather simple (though still somewhat lengthy).

Finally, we combine Theorem 1 with the theory of smoothing in $\mathcal{L}^p(\mathbb{R})$, as it can, e.g., be found in [Zy 88]. This way we obtain a result which can be used as the basis for the usual construction of the Itô-integral (under the usual assumptions on X) as sketched above. In order to formulate the result we introduce the following notion:

Definition 2 An extended real valued stochastic process $X = (X_t, t \in \mathbb{R}_+)$ is called *P*-a.s. (uniformly) sample continuous if there exists a *P*-null set $N \in \mathcal{A}$ so that for all $\omega \in \mathbb{C}N$, the mapping $t \mapsto X_t(\omega)$ from \mathbb{R}_+ into $\overline{\mathbb{R}}$ is (uniformly) continuous.

Theorem 2 Suppose that $X = (X_t, t \in T)$ is an $\overline{\mathbb{R}}$ -valued stochastic process, which belongs to $\mathcal{L}^2(\Omega \times \mathbb{R}_+)$ and which is adapted to $\mathcal{F} = (\mathcal{F}_t, t \in \mathbb{R}_+)$. Then there exists a sequence $(X_n, n \in \mathbb{N})$ of bounded measurable \mathcal{F} -adapted processes which are *P*-a.s. uniformly sample continuous, so that

$$\lim_{n \rightarrow \infty} \int_{\Omega \times \mathbb{R}_+} |X - X_n|^2 d(P \otimes \lambda) = 0.$$

Throughout the paper we avoided to enlarge the various σ -algebras, even though this would have been very convenient here and there. In particular, we do not assume that the underlying probability space is complete, nor that the filtration \mathcal{F} satisfies the “usual” conditions, i.e., that \mathcal{F}_0 contains all *P*-null sets, and that \mathcal{F} is right-continuous. We did this in order to be compatible with the existing literature, especially in applied fields, where the underlying σ -algebra and the filtration generated by the integrator process are often used “as is”.

The organization of the paper is as follows. Sections 2 and 3 contain the proofs of Theorems 1 and 2, respectively. In order to make the paper self-contained and to keep the flow of the arguments of the proofs of the theorems steady, we have deferred a number of results to four appendices. Appendix A contains the powerful version of the monotone class theorem as it is found in [DM 78], Appendix B discusses weak \mathcal{L}^p -convergence versus measurability, Appendix C contains the statement and proof of the Theorem of Choice, and Appendix D gives an account of $\mathcal{L}^p(\mathbb{R})$ -smoothing.

2. Proof of Theorem 1

First we establish in a couple of lemmas that every measurable stochastic process X has an approximation (in a sense which will be clarified below, cf. Lemma 2.4 and Corollary 2.5) by processes of the form (1.3). We begin with a result which is well-known in measure theory.

Lemma 2.1 Let (E, \mathcal{E}) be a measurable space, (M, d) a metric space equipped with its Borel σ -algebra \mathcal{B}_d . Let $(f_n, n \in \mathbb{N})$ be a sequence of measurable mappings from E into M which converges pointwise to f . Then f is measurable.

Remark In the case $M = \mathbb{R}$ one finds this result in every textbook on measure and integration theory. A proof for the general case can be found, e.g., in [La 69, Chapter X, §1].

If M is in addition separable, we can take the Lebesgue approximation of f and obtain a result in the reverse direction:

Lemma 2.2 Let (E, \mathcal{E}) be a measurable space, (M, d) a separable metric space. Assume that f is a measurable mapping from E into M . Then f is a uniform pointwise limit of a sequence of countably valued, measurable step functions.

Proof For given $n \in \mathbb{N}$, we construct a countably valued, measurable step function f_n so that for all $x \in E$, $d(f_n(x), f(x)) < 1/n$, in the following way: Let $\{y_m, m \in \mathbb{N}\}$ be a dense subset in M . For $m, n \in \mathbb{N}$, $B_{n,m}$ denotes the open ball of radius $1/n$ with center y_m : $B_{n,m} = \{y \in M; d(y_m, y) < 1/n\}$. Furthermore, define

$$A_{n,m} := B_{n,m} \setminus \left(\bigcup_{k=1}^{m-1} B_{n,k} \right).$$

Then for every $n \in \mathbb{N}$, $\{A_{n,m}, m \in \mathbb{N}\}$ is a cover of M of pairwise disjoint Borel sets. Now we set

$$f_n(x) := y_m \quad \text{for all } x \in f^{-1}(A_{n,m}).$$

f_n is clearly of the desired type. Let $x \in E$, $n \in \mathbb{N}$. Then $x \in f^{-1}(A_{n,m})$ for some $m \in \mathbb{N}$, and $f_n(x) = y_m \in A_{n,m}$, so that $d(f_n(x), f(x)) < 1/n$. \square

Consider again a general metric space (M, d) and a sequence of countably valued mappings from E into M . Then the union of the images of E under these mappings is still a countable subset of M , and its closure in (M, d) is a separable subspace of (M, d) . Thus we can combine Lemmas 2.1 and 2.2 into the following

Corollary 2.3 Let f be a mapping from a measurable space (E, \mathcal{E}) into a metric space (M, d) equipped with its Borel σ -algebra. Then the following two statements are equivalent:

- (i) f is measurable and takes values in a separable subspace of M ;
- (ii) f is the uniform pointwise limit of a sequence of countably valued, measurable step functions.

Let (Ω, \mathcal{A}, P) be a probability space. \mathcal{R}_P denotes the space of P -equivalence classes $[Y]_P$ of extended real valued random variables Y on (Ω, \mathcal{A}, P) . By $\mathcal{R}_{P,b}$ we denote the space of P -equivalence classes of P -a.s. bounded real valued random variables. We equip $\mathcal{R}_{P,b}$ with the norm $\|\cdot\|_2$ of $L^2(P)$.

Let X be an extended real valued stochastic process on (Ω, \mathcal{A}, P) . For simplicity we shall only consider \mathbb{R}_+ as time parameter domain; — the modification of the arguments of this paper for other domains are straightforward. Throughout we consider X as a mapping

$$\begin{aligned} X : \quad \Omega \times \mathbb{R}_+ &\longrightarrow \overline{\mathbb{R}} \\ (\omega, t) &\longmapsto X_t(\omega). \end{aligned}$$

Also the notation $X(\omega, t)$ for the evaluation $X_t(\omega)$ of X at (ω, t) will be convenient here and there. With X we can associate the following mapping

$$\begin{aligned} \widehat{X} : \quad \mathbb{R}_+ &\longrightarrow \mathcal{R}_P \\ t &\longmapsto \widehat{X}_t := [X_t]_P. \end{aligned}$$

Without loss of generality, we shall consider from now on only processes X which are such that \widehat{X} takes values in $\mathcal{R}_{P,b}$. We call X (P -a.s.) *bounded*, if there is $M \geq 0$ so that for all $t \in \mathbb{R}_+$ and (P -a.e.) $\omega \in \Omega$, $|X(\omega, t)| \leq M$.

Lemma 2.4 Assume that $X = (X_t, t \in \mathbb{R}_+)$ is a measurable stochastic process so that for all $t \in \mathbb{R}_+$, X_t is bounded. Then the following two equivalent statements hold:

- (i) \widehat{X} is measurable and takes its values in a separable subspace of $\mathcal{R}_{P,b}$;
- (ii) \widehat{X} is a uniform pointwise limit of countably valued, measurable step functions.

Proof If we choose $E = \mathbb{R}_+$, $\mathcal{E} = \mathcal{B}(\mathbb{R}_+)$, $M = \mathcal{R}_{P,b}$, and d as defined by $\|\cdot\|_2$ on $\mathcal{R}_{P,b}$, then the equivalence of (i) and (ii) is just the statement of Corollary 2.3.

First we reduce to the case that X is bounded. Assume that we have shown that for every measurable bounded process Y the properties (i) and (ii) hold for \widehat{Y} . Let X be a process as in the hypothesis of the lemma, and for $n \in \mathbb{N}$ set

$$X_t^{(n)}(\omega) := (X_t(\omega) \wedge n) \vee (-n), \quad \omega \in \Omega, t \in \mathbb{R}_+.$$

Then $X^{(n)}$ is bounded, and by our assumption we have that $\widehat{X^{(n)}}$ is measurable and takes values in a separable subspace of $\mathcal{R}_{P,b}$. For every $t \in \mathbb{R}_+$, $X_t^{(n)}$ converges pointwise on Ω to X_t . Since we can bound $|X_t - X_t^{(n)}|^2$ from above by $4|X_t|^2$, the dominated convergence theorem implies that the sequence $(\widehat{X^{(n)}})$, $n \in \mathbb{N}$ of mappings from \mathbb{R}_+ into $(\mathcal{R}_{P,b}, \|\cdot\|_2)$ converges pointwise to \widehat{X} . $\widehat{X^{(n)}}$ being measurable, we obtain from Lemma 1.1 that \widehat{X} is measurable, too. Furthermore, as the limit of separably valued $\widehat{X^{(n)}}$, \widehat{X} is separably valued, too. Thus, from now on we may and do assume that X is bounded.

We denote by \mathcal{H} the set of all bounded processes X so that for \widehat{X} property (i), and consequently also property (ii), holds. We show that \mathcal{H} is a vector space which admits all the properties in the hypothesis of the monotone class theorem (Theorem A.1 in Appendix A). That \mathcal{H} is a vector space with $1 \in \mathcal{H}$ is trivial from (i). Next we show that \mathcal{H} is stable under uniform limits. Assume that $(X_n, n \in \mathbb{N})$ is a sequence in \mathcal{H} which converges uniformly on $\Omega \times \mathbb{R}_+$ to X . Then in particular for every $t \in \mathbb{R}_+$, $(X_n(\cdot, t), n \in \mathbb{N})$ converges uniformly on Ω to $X(\cdot, t)$, and therefore for every $t \in \mathbb{R}_+$, $(\widehat{X_n}(t), n \in \mathbb{N})$ converges in $L^2(P)$ to $\widehat{X}(t)$. By assumption, for every $n \in \mathbb{N}$ the mapping $t \mapsto \widehat{X_n}(t)$ from \mathbb{R}_+ into $(\mathcal{R}_{P,b}, \|\cdot\|_2)$ is measurable, and by Lemma 1.1, the same holds for $t \mapsto \widehat{X}(t)$. Moreover, since for all $n \in \mathbb{N}$, $\widehat{X_n}$ is separably valued, so is \widehat{X} . (Set $M_{00} = \bigcup_n \widehat{X_n}(\mathbb{R}_+)$, which is a separable subspace of (M, d) . Let M_0 be the closure of M_{00} . Then also M_0 is separable. For every $t \in \mathbb{R}_+$ we have $\widehat{X}(t) = \lim_n \widehat{X_n}(t)$, $\widehat{X_n}(t) \in M_{00}$. Hence $\widehat{X}(t) \in M_0$, because M_0 is closed.) Thus \widehat{X} fulfills (i) and we have $X \in \mathcal{H}$, so that \mathcal{H} is closed under uniform limits.

Now assume that $(X_n, n \in \mathbb{N})$ is a uniformly bounded, increasing sequence of positive processes in \mathcal{H} which converges to X . Clearly, X is bounded, too. By the dominated convergence theorem, for every $t \in \mathbb{R}_+$, $\widehat{X_n}(t)$ converges to $\widehat{X}(t)$ in $L^2(P)$. Then we can argue as in the previous case and find that \mathcal{H} is closed under limits of uniformly bounded, increasing sequences.

Let \mathcal{C} denote the set of processes X of the form

$$X_t(\omega) = 1_A(\omega) 1_{[a,b]}(t), \quad \omega \in \Omega, t \in \mathbb{R}_+,$$

with $a, b \in \mathbb{R}_+$, $a < b$, $A \in \mathcal{A}$. Clearly, \mathcal{C} is closed under multiplication, and $\sigma(\mathcal{C}) = \mathcal{A} \otimes \mathcal{B}(\mathbb{R}_+)$. By the monotone class theorem (Theorem A.1 in Appendix A) \mathcal{H} contains all bounded measurable processes. \square

Let X be as in the hypothesis of Lemma 2.4, and let the sequence in statement (ii) of Lemma 2.4 be denoted by $(\widehat{X}_n, n \in \mathbb{N})$. Then for $n \in \mathbb{N}$, \widehat{X}_n is of the following form

$$\widehat{X}_n(t) = \sum_{k=1}^{\infty} [H_{n,k}]_P 1_{A_{n,k}}(t), \quad t \in \mathbb{R}_+,$$

where $(A_{n,k}, k \in \mathbb{N})$ is a sequence of pairwise disjoint Borel subsets which form a partition of \mathbb{R}_+ , and $(H_{n,k}, k \in \mathbb{N})$ is a sequence of random variables. Hence statement (ii) of Lemma 2.4 reads explicitly as follows:

Corollary 2.5 Let X be as in the hypothesis of Lemma 2.4. Then there exists a sequence $(X_n, n \in \mathbb{N})$ of processes of the form

$$X_n(\omega, t) = \sum_{k=1}^{\infty} H_{n,k}(\omega) 1_{A_{n,k}}(t), \quad \omega \in \Omega, t \in \mathbb{R}_+,$$

so that

$$\sup_{t \in \mathbb{R}_+} \left(\mathbb{E} \left(X(\cdot, t) - X_n(\cdot, t) \right)^2 \right)^{1/2} \xrightarrow{n \rightarrow \infty} 0.$$

Lemma 2.6 Let (E_i, \mathcal{E}_i) , $i = 1, 2$, be two measurable spaces and f_i , $i = 1, 2$, be measurable mappings from (E_i, \mathcal{E}_i) to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Set

$$f_i \otimes f_j(x, y) := f_i(x) f_j(y), \quad x \in E_i, y \in E_j.$$

Then $f_1 \otimes f_2$ is $\mathcal{E}_1 \otimes \mathcal{E}_2$ - $\mathcal{B}(\mathbb{R})$ -measurable.

The proof of Lemma 2.6 is an elementary exercise in “turning the prayer-wheel” of measure theory: Begin with indicators, move on to their linear combinations with positive coefficients, take increasing limits to get the statement for all positive measurable functions f_i , $i = 1, 2$, and finally decompose general f_i into positive and negative parts. Details are left to the interested reader.

Given a filtration $\mathcal{F} = (\mathcal{F}_t, t \in \mathbb{R}_+)$ and $t \in \mathbb{R}_+$, we set $\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$.

Lemma 2.7 Let $A \in \mathcal{B}(\mathbb{R}_+)$, $\tau = \inf A$, and let Z be a random variable. Set $Y = Z \otimes 1_A$. Then Y is in the following cases \mathcal{F} -progressive:

- (a) $\tau \in A$ and Z is \mathcal{F}_τ -measurable, or
- (b) $\tau \notin A$ and Z is $\mathcal{F}_{\tau+}$ -measurable.

Proof We only give the (almost trivial) proof for case (b), case (a) is similar (and even easier). If $t \leq \tau$, then $Y(s) = 0$ for all $s \in [0, t]$, and hence the restriction of Y to $\Omega \times [0, t]$ is $\mathcal{F}_t \otimes \mathcal{B}([0, t])$ -measurable. Let $t > \tau$. Then $\mathcal{F}_{\tau+} \subset \mathcal{F}_t$, and therefore Z is \mathcal{F}_t -measurable. Moreover, 1_A restricted to $[0, t]$ is $\mathcal{B}([0, t])$ -measurable. Now we can apply Lemma 2.6 to conclude that the restriction of Y to $\Omega \times [0, t]$ is $\mathcal{F}_t \otimes \mathcal{B}([0, t])$ -measurable. \square

Lemma 2.8 Suppose that X is an \mathcal{F} -adapted process, $t \in \mathbb{R}_+$, and $(t_n, n \in \mathbb{N})$ is a sequence which converges to t . Assume furthermore that $(X_{t_n}, n \in \mathbb{N})$ converges weakly in $\mathcal{L}^p(P)$ to a random variable Z . Then the following statements hold:

- (a) If $(t_n, n \in \mathbb{N})$ is decreasing to t , then Z is P -a.s. equal to an \mathcal{F}_{t+} -measurable random variable.
- (b) If for all $n \in \mathbb{N}$, $\|X_{t_n}\|_p \leq M$, then also $\|Z\|_p \leq M$.

Proof

(a) Let $Z_n := X_{t_n}$, $n \in \mathbb{N}$, so that a comparison with Lemma B.1 in Appendix B shows that Z is P -a.s. equal to a random variable which is measurable with respect to the σ -algebra $\bigcap_n \sigma(X_{t_k}, k \geq n)$. On the other hand, since X is \mathcal{F} -adapted, we have that $\sigma(X_{t_k}, k \geq n) \subset \mathcal{F}_{t_n}$. Finally, we observe that $\bigcap_n \mathcal{F}_{t_n} = \mathcal{F}_{t+}$, because $\mathcal{F}_t \subset \mathcal{F}_s$ for $s \geq t$.

(b) With the notation $\mathcal{B}_1^q := \{Y \in \mathcal{L}^q(P); \|Y\|_q \leq 1\}$, we have

$$\begin{aligned}
\|Z\|_p &= \sup_{Y \in \mathcal{B}_1^q} \mathbb{E}(ZY) \\
&= \sup_{Y \in \mathcal{B}_1^q} \liminf_n \mathbb{E}(X_{t_n} Y) \\
&\leq \liminf_n \sup_{Y \in \mathcal{B}_1^q} \mathbb{E}(X_{t_n} Y) \\
&= \liminf_n \|X_{t_n}\|_p \\
&\leq M.
\end{aligned}$$

\square

Now we are ready for the proof of Theorem 1.

Proof of Theorem 1 We do the proof in several steps.

Step 1 With the following argument we reduce to the case that X is in addition bounded. Suppose that we can show the statement of Theorem 1 for every bounded, measurable and \mathcal{F} -adapted process. Consider $Y = \arctan X$. Then Y is bounded, measurable and \mathcal{F} -adapted, and by our assumption Y has an \mathcal{F} -progressive modification \tilde{Y} . Set $\tilde{X} = \tan \tilde{Y}$. Then \tilde{X} is progressive, too, and for $t \in \mathbb{R}_+$ we have $P(X_t = \tilde{X}_t) = P(\arctan X_t = \tilde{Y}_t) = P(Y_t = \tilde{Y}_t) = 1$. Hence \tilde{X} is also a modification of X . Therefore, we can from now on assume that X is bounded.

Step 2 By Corollary 2.5 we know that there exists a sequence of processes $(X_n, n \in \mathbb{N})$ of the form

$$X_n(t) = \sum_{k=1}^{\infty} H_{n,k} 1_{A_{n,k}}(t), \quad t \in \mathbb{R}_+, \quad (2.1)$$

where the $A_{n,k} \in \mathcal{B}(\mathbb{R}_+)$ form a partition of \mathbb{R}_+ , and $(H_{n,k}, n, k \in \mathbb{N})$ are random variables, such that

$$\lim_{n \rightarrow \infty} \sup_{t \in \mathbb{R}_+} \|X(t) - X_n(t)\|_2 = 0,$$

where $\|\cdot\|_2$ denotes the semi-norm of $\mathcal{L}^2(P)$. By choosing a subsequence — if necessary — we may assume that for every $n \in \mathbb{N}$ we have

$$\sup_{t \in \mathbb{R}_+} \|X(t) - X_n(t)\|_2 \leq 2^{-n}.$$

This entails that for all $n, k \in \mathbb{N}$,

$$\sup_{t \in A_{n,k}} \|X(t) - H_{n,k}\|_2 \leq 2^{-n}.$$

Step 3 We replace the $H_{n,k}$ in (2.1) by other random variables so as to make the resulting process \mathcal{F} -progressive as follows: Let $n, k \in \mathbb{N}$, and define $\tau_{n,k} := \inf A_{n,k}$.

Case 1 $\tau_{n,k} \in A_{n,k}$. Set

$$\tilde{X}_{n,k} := X_{\tau_{n,k}}.$$

Then we have

$$\begin{aligned} \sup_{t \in A_{n,k}} \|X(t) - \tilde{X}_{n,k}\|_2 &\leq \sup_{t \in A_{n,k}} (\|X(t) - H_{n,k}\|_2 + \|H_{n,k} - X_{\tau_{n,k}}\|_2) \\ &\leq 2^{-(n-1)}. \end{aligned}$$

By Lemma 2.7a, $\tilde{X}_{n,k} \otimes 1_{A_{n,k}}$ is \mathcal{F} -progressive.

Case 2 $\tau_{n,k} \notin A_{n,k}$. Choose a sequence $(t_m, m \in \mathbb{N})$ in $A_{n,k}$ which is decreasing to $\tau_{n,k}$. Consider the sequence $(X_{t_m}, m \in \mathbb{N})$ which is bounded in $\mathcal{L}^2(P)$. By Theorem C.2 in Appendix C, we may assume — by selecting another subsequence, if necessary — that $(X_{t_m}, m \in \mathbb{N})$ converges weakly in $\mathcal{L}^2(P)$ to some random variable $\tilde{X}_{n,k}$. By Lemma 2.8a, we may choose $\tilde{X}_{n,k}$ as $\mathcal{F}_{\tau_{n,k}+}$ -measurable. Therefore, the process $\tilde{X}_{n,k} \otimes 1_{A_{n,k}}$ is \mathcal{F} -progressive by Lemma 2.7b. Moreover, by Lemma 2.8b we have

$$\sup_{t \in A_{n,k}} \|X(t) - \tilde{X}_{n,k}\|_2 \leq 2^{-(n-1)},$$

because $(X_{t_m} - H_{n,k}, m \in \mathbb{N})$ converges weakly in $\mathcal{L}^2(P)$ to $\tilde{X}_{n,k} - H_{n,k}$.

Step 4 For $n \in \mathbb{N}$ define

$$\tilde{X}_n := \sum_{k=1}^{\infty} \tilde{X}_{n,k} \otimes 1_{A_{n,k}}.$$

In view of Lemma 2.6 it is clear that \tilde{X}_n is measurable and by Step 3, for every $n \in \mathbb{N}$, \tilde{X}_n is \mathcal{F} -progressive. Moreover, we have

$$\sup_{t \in \mathbb{R}_+} \|X(t) - \tilde{X}_n(t)\|_2 \leq 2^{-(n-1)}.$$

But this implies that for every $t \in \mathbb{R}_+$, $(\tilde{X}_n(t), n \in \mathbb{N})$ converges P -a.s. to $X(t)$, uniformly in $t \in \mathbb{R}_+$. (The argument is the same as proving, e.g., that \mathcal{L}^2 -convergence implies a.e. convergence of a subsequence.)

Step 5 Define

$$\tilde{X} := \limsup_n \tilde{X}_n.$$

Then \tilde{X} is still measurable and \mathcal{F} -progressive, because it is the pointwise lim sup of mappings of this type. On the other hand, for every $t \in \mathbb{R}_+$, $X(t)$ is P -a.s. the limit of $(\tilde{X}_n(t), n \in \mathbb{N})$. Hence we have for every $t \in \mathbb{R}_+$, $P(X(t) = \tilde{X}(t)) = 1$, so that \tilde{X} is a modification of X . \square

3. Proof of Theorem 2

Let X be as in the hypothesis of Theorem 2, and consider a measurable \mathcal{F} -progressive modification \tilde{X} of X which exists according to Theorem 1. Because \tilde{X} is a modification of X , we have that for every $t \in \mathbb{R}_+$,

$$\int_{\Omega} |\tilde{X}_t - X_t|^2 dP = 0.$$

The theorem of Fubini–Tonelli entails that $\tilde{X} \in \mathcal{L}^2(P \otimes \lambda)$ and

$$\|\tilde{X} - X\|_{\mathcal{L}^2(P \otimes \lambda)} = 0.$$

Therefore, it suffices to show the statement for \tilde{X} .

For $n \in \mathbb{N}$, set

$$\tilde{X}_n := (\tilde{X} \wedge n) \vee (-n).$$

Then \tilde{X}_n is bounded by n , and for all $n \in \mathbb{N}$, we have $|\tilde{X}_n| \leq |\tilde{X}|$. Since $|\tilde{X}| \in \mathcal{L}^2(P \otimes \lambda)$, we may apply the dominated convergence theorem, and conclude that $(\tilde{X}_n, n \in \mathbb{N})$ converges to \tilde{X} in $\mathcal{L}^2(P \otimes \lambda)$. Therefore, in the sequel we may assume that \tilde{X} is bounded, say by $M > 0$.

Let $\omega \in \Omega$ and extend $\tilde{X}(\omega, \cdot)$ to \mathbb{R} by setting it identically zero on $(-\infty, 0)$. Let $\varphi = 1_{[0,1]}$, and for $n \in \mathbb{N}$, $\varphi_n(u) = n \varphi(nu)$, $u \in \mathbb{R}$. Put

$$X_n(\omega, t) := (\varphi_n * \tilde{X}(\omega, \cdot))(t), \quad t \in \mathbb{R},$$

and restrict $X_n(\omega, \cdot)$ to \mathbb{R}_+ . According to Lemma D.2 in Appendix D, for every $\omega \in \Omega$, $X_n(\omega, \cdot)$ is uniformly continuous on \mathbb{R}_+ . Lemma D.5 in Appendix D implies that $(X_n(\omega, \cdot), n \in \mathbb{N})$ converges in $\mathcal{L}^2(\mathbb{R}_+)$ to $\tilde{X}(\omega, \cdot)$. It is plain to check that $|X_n(\omega, t)| \leq M$ for all $n \in \mathbb{N}$, $(\omega, t) \in \Omega \times \mathbb{R}_+$. Thus the dominated convergence theorem implies that

$$\|\tilde{X} - X_n\|_{\mathcal{L}^2(P \otimes \lambda)} \rightarrow 0,$$

as $n \rightarrow \infty$.

It remains to show that X_n is \mathcal{F} -adapted. Let $t \in \mathbb{R}_+$. Writing the convolution $\varphi_n * \tilde{X}$ as an integral, we have

$$\begin{aligned} X_n(\omega, t) &= n \int_{t-1/n}^t \tilde{X}(\omega, s) ds \\ &= n \int_0^t 1_{[(t-1/n) \vee 0, t]}(s) \tilde{X}(\omega, s) ds. \end{aligned}$$

\tilde{X} is \mathcal{F} -progressive so that its restriction to $\Omega \times [0, t]$ is $\mathcal{F}_t \otimes \mathcal{B}([0, t])$ -measurable. It is trivial that $1_{[(t-1/n) \vee 0, t]}$ has the same properties, and therefore the integrand is $\mathcal{F}_t \otimes \mathcal{B}([0, t])$ -measurable. But then the (first part of the) theorem of Fubini–Tonelli states that $\omega \mapsto X_n(\omega, t)$ is \mathcal{F}_t -measurable. \square

Appendix A: Monotone Class Theorem

The following powerful version of the monotone class theorem can be found, e.g., in [DM 78]. For the convenience of the reader we give a detailed proof.

Theorem A.1 Let (E, \mathcal{E}) be a measurable space, and \mathcal{H} a vector space of bounded, real valued functions on E which contains 1. Assume that the following holds:

- (i) \mathcal{H} is closed under uniform convergence;
- (ii) \mathcal{H} is closed under limits of uniformly bounded, monotone increasing sequences of nonnegative functions.

If \mathcal{C} is a subset of \mathcal{H} which is closed under multiplication, then \mathcal{H} contains all bounded, $\sigma(\mathcal{C})$ -measurable functions.

Proof Let \mathcal{C}' be the algebra generated by \mathcal{C} and 1, i.e. $f \in \mathcal{C}'$ is of the form

$$f = a_0 + \sum_{k=1}^n a_k f_k$$

for certain $n \in \mathbb{N}$, $a_0, a_1, \dots, a_n \in \mathbb{R}$, $f_1, \dots, f_n \in \mathcal{C}$. Since $1 \in \mathcal{H}$, $\mathcal{C} \subset \mathcal{H}$ and \mathcal{H} is a vector space, we have $\mathcal{C}' \subset \mathcal{H}$.

Let \mathbb{A} denote the set of all algebras \mathcal{T} of functions in \mathcal{H} so that $\mathcal{C}' \subset \mathcal{T}$. \mathbb{A} is partially ordered by inclusion. Consider a chain $\mathcal{T}_1 \subset \mathcal{T}_2 \subset \dots \subset \mathcal{T}_n \subset \dots$ in \mathbb{A} . Then \mathbb{A} contains

also $\bigcup_k \mathcal{T}_k$: If $f, g \in \bigcup_k \mathcal{T}_k$, then we can find k so that $f, g \in \mathcal{T}_k$, and therefore $f \cdot g \in \mathcal{T}_k$, and $f \cdot g \in \bigcup_k \mathcal{T}_k$. Similarly, we can show that $f + ag$ belongs to $\bigcup_k \mathcal{T}_k$ for any $a \in \mathbb{R}$. Hence $\bigcup_k \mathcal{T}_k$ is an algebra, it contains \mathcal{C}' and $\bigcup_k \mathcal{T}_k \subset \mathcal{H}$. Therefore $\bigcup_k \mathcal{T}_k \in \mathbb{A}$, and $\bigcup_k \mathcal{T}_k$ is an upper bound for the chain $\mathcal{T}_1 \subset \mathcal{T}_2 \subset \dots \subset \mathcal{T}_k \subset \dots$. By Zorn's lemma, \mathbb{A} has a maximal element which we denote by \mathcal{T}_0 .

We are now going to prove that \mathcal{T}_0 shares all stability properties of \mathcal{H} . Note that $1 \in \mathcal{T}_0$, because $\mathcal{C}' \subset \mathcal{T}_0$ and $1 \in \mathcal{C}'$.

We show that \mathcal{T}_0 is closed under uniform limits. Assume that $(f_n, n \in \mathbb{N})$ is a sequence of functions in \mathcal{T}_0 which converges uniformly to f . Since $\mathcal{T}_0 \subset \mathcal{H}$ and \mathcal{H} is stable with respect to uniform convergence, we have $f \in \mathcal{H}$. Assume that $f \notin \mathcal{T}_0$. We bring this in contradiction to the maximality of \mathcal{T}_0 as follows: Let \mathcal{T}'_0 be the algebra generated by \mathcal{T}_0 and f so that our assumption implies that \mathcal{T}'_0 is strictly larger than \mathcal{T}_0 . The contradiction follows if we can show that $\mathcal{T}'_0 \in \mathbb{A}$. $\mathcal{C}' \subset \mathcal{T}'_0$ is trivial, and so it remains to show that $\mathcal{T}'_0 \subset \mathcal{H}$. Clearly, the uniform convergence of $(f_n, n \in \mathbb{N})$ to f entails the uniform convergence of $(f_n^p, n \in \mathbb{N})$ to $f^p, p \in \mathbb{N}$. Since $f_n^p \in \mathcal{T}_0$, we have $f_n^p \in \mathcal{H}$, and the stability of \mathcal{H} with respect to uniform convergence implies $f^p \in \mathcal{H}$. Similarly, for every $g \in \mathcal{T}_0$, $g \cdot f_n^p \in \mathcal{T}_0 \subset \mathcal{H}$ and these functions converge uniformly to $g \cdot f^p$, so that $g \cdot f^p \in \mathcal{H}$. Thus we have $\mathcal{T}'_0 \subset \mathcal{H}$ and the proof of the stability of \mathcal{T}_0 with respect to uniform convergence.

Next we prove that \mathcal{T}_0 is closed under taking absolute values. Let $f \in \mathcal{T}_0$ and recall that $|f|$ is bounded, say by $M > 0$. By Weierstraß' theorem the function $x \mapsto |x|$ can be approximated uniformly on $[-M, M]$ by a sequence $(p_n, n \in \mathbb{N})$ of polynomials p_n . Then $p_n \circ f$ belongs to \mathcal{T}_0 and converges uniformly to $|f|$, and therefore $|f|$ belongs to \mathcal{T}_0 . Consequently, \mathcal{T}_0 is also closed under the operations \wedge and \vee .

Now we can show that \mathcal{T}_0 is closed under increasing limits of uniformly bounded sequences of nonnegative functions in \mathcal{T}_0 . Let $(f_n, n \in \mathbb{N})$ be such a sequence. Then it has a limit f in \mathcal{H} . Also $(f_n^p, n \in \mathbb{N})$ is monotone increasing to f^p for all $p \in \mathbb{N}$. Hence we have $f^p \in \mathcal{H}$. For $g \in \mathcal{T}_0$, we decompose $g = g^+ - g^-$ with $g^\pm \geq 0$ in \mathcal{T}_0 , by the previous step. Then $g^\pm \cdot f_n^p$ increase to $g^\pm \cdot f^p$. Therefore we have $g \cdot f^p \in \mathcal{H}$. Thus, as before, the assumption $f \notin \mathcal{T}_0$ leads to a contradiction.

Consider the family

$$\mathcal{D} := \{A \subset E; 1_A \in \mathcal{T}_0\}.$$

Since \mathcal{T}_0 is an algebra it follows that \mathcal{D} is stable with respect to intersections. $1 \in \mathcal{T}_0$ implies that $E \in \mathcal{D}$, and the fact that \mathcal{T}_0 is a vector space entails that \mathcal{D} is closed under formation of complements. Finally, the stability of \mathcal{T}_0 with respect to uniformly bounded, monotone increasing limits of positive functions translates into the stability of \mathcal{D} with respect to countable unions of monotone increasing sequences. Thus \mathcal{D} is an intersection-stable Dynkin system and consequently, a σ -algebra. It is clear, that \mathcal{T}_0 contains all bounded \mathcal{D} - $\mathcal{B}(\mathbb{R})$ -measurable functions.

Finally, we show $\sigma(\mathcal{C}) \subset \mathcal{D}$ to conclude the proof: This implies that every $\sigma(\mathcal{C})$ -measurable function is \mathcal{D} -measurable, hence in \mathcal{T}_0 and consequently in \mathcal{H} . To this end,

we remark that we have $\sigma(\mathcal{C}) \subset \sigma(\mathcal{T}_0)$, since $\mathcal{C} \subset \mathcal{T}_0$, so that it suffices to prove that $\sigma(\mathcal{T}_0) \subset \mathcal{D}$. This in turn follows, if we can show that for every $f \in \mathcal{T}_0$,

$$B := \{x \in E; f(x) \geq 1\} \in \mathcal{D},$$

because then we also have $\{x \in E; f(x) \geq a\} \in \mathcal{D}$ for all $a \in \mathbb{R}$. To prove $B \in \mathcal{D}$, consider

$$1_B(x) = \begin{cases} 1 & \text{if } f(x) \geq 1 \\ 0 & \text{otherwise,} \end{cases}$$

and define

$$g = (f \wedge 1) \vee 0 \in \mathcal{T}_0.$$

Then $(g^n, n \in \mathbb{N})$ decreases to 1_B , or in other words, $1 - g^n$ increases to $1_{\mathcal{C}_B}$. Thus we have $1_{\mathcal{C}_B} \in \mathcal{T}_0$ and hence also $1_B \in \mathcal{T}_0$. Consequently, $B \in \mathcal{D}$ and the proof is finished. \square

Appendix B: Weak \mathcal{L}^P – Convergence and Measurability

Let (Ω, \mathcal{A}, P) be a probability space, and assume that $(Z_n, n \in \mathbb{N})$ is a sequence of random variables in $\mathcal{L}^p(P)$, $p \geq 1$, which converges weakly in $\mathcal{L}^p(P)$ to a random variable Z . For $n \in \mathbb{N}$ set

$$\mathcal{A}_{-n} := \sigma(Z_k, k \geq n),$$

and furthermore define

$$\mathcal{A}_{-\infty} := \bigcap_{n \in \mathbb{N}} \mathcal{A}_{-n}.$$

Lemma B.1 Z is P -a.s. equal to an $\mathcal{A}_{-\infty}$ -measurable random variable.

Proof First we show that for every $n \in \mathbb{N}$, $(Z_k, k \geq n)$ converges weakly in $\mathcal{L}^p(P)$ to $\mathbb{E}(Z | \mathcal{A}_{-n})$. To this end, let $Y \in \mathcal{L}^q(P)$ with $p^{-1} + q^{-1} = 1$, and observe that by Jensen's inequality we have that $\mathbb{E}(Z | \mathcal{A}_{-n}) \in \mathcal{L}^q(P)$. Then we have for every $k \in \mathbb{N}$ with $k \geq n$,

$$\begin{aligned} \mathbb{E}(Y (\mathbb{E}(Z | \mathcal{A}_{-n}) - Z_k)) &= \mathbb{E}(Y \mathbb{E}(Z - Z_k | \mathcal{A}_{-n})) \\ &= \mathbb{E}(\mathbb{E}(Y | \mathcal{A}_{-n}) (Z - Z_k)), \end{aligned}$$

and by hypothesis the last term converges to zero as k tends to infinity.

Consequently, for every $n \in \mathbb{N}$ there exists $N_n \in \mathcal{N}$, where \mathcal{N} is the family of P -null sets in \mathcal{A} , with $Z = \mathbb{E}(Z | \mathcal{A}_{-n})$ on the complement of N_n . Set $N = \bigcup_n N_n \in \mathcal{N}$. Then for all $n \in \mathbb{N}$, $Z = \mathbb{E}(Z | \mathcal{A}_{-n})$ on the complement of the null set N .

Since $Z \in \mathcal{L}^p(P)$ for some $p \geq 1$, we have that Z is P -integrable. Therefore we can now apply Theorem VII.4.3 in [Do 53] with the result that the sequence $(\tilde{Z}_n, n \in -\mathbb{N})$ given by $\tilde{Z}_n := \mathbb{E}(Z | \mathcal{A}_n)$, $n \in -\mathbb{N}$, converges P -a.s. to $\mathbb{E}(Z | \mathcal{A}_{-\infty})$, as n tends to $-\infty$. If we denote the exceptional set for this convergence by M , then we have that $Z = \mathbb{E}(Z | \mathcal{A}_{-\infty})$ on the complement of the P -null set $N \cup M$. \square

Remarks The theorem of Doob we used in the last step of the proof is a very simple consequence of Doob's second martingale convergence theorem (Theorem VII.4.2 in [Do 53], or, e.g., Theorem 19.9 in [Ba 91], Corollary 2.4 in [RY 91]), which is a convergence theorem for martingales indexed by $-\mathbb{N}$. It is very easy to use alternatively Doob's second martingale theorem directly for the second step in the proof.

If one wants to avoid to use Doob's second martingale theorem or its above mentioned consequence, one can simply set

$$\tilde{Z} := \begin{cases} Z & \text{on } \mathbb{C}N \\ 0 & \text{on } N \end{cases}$$

after the first step in the proof. Then \tilde{Z} is measurable with respect to the σ -algebra

$$\bigcap_{n \in -\mathbb{N}} \overline{\mathcal{A}_n},$$

where $\overline{\mathcal{A}_n}$ denotes the augmentation of \mathcal{A}_n by the P -null sets in \mathcal{A} . For the application of this result to Theorems 1 and 2, one has then to augment the filtration considered there appropriately. For example, the conclusion of Theorem 1 would read that X has a measurable modification which is progressive with respect to $\overline{\mathcal{F}}$, where $\overline{\mathcal{F}}$ is the filtration obtained from \mathcal{F} by augmenting it with the P -null sets in \mathcal{A} . The statement of Theorem 2 has to be adjusted accordingly. Of course, this gives somewhat weaker results than Theorem 1 and Theorem 2, but for all practical purposes the differences are immaterial.

Appendix C: Theorem of Choice

In this appendix we give a proof of the classical theorem of choice as one can find it, e.g., in [RN 55]. We begin with a (well-known) lemma.

Lemma C.1 Every Hilbert space is weakly complete.

Proof Let \mathcal{H} be a Hilbert space with inner product (\cdot, \cdot) , norm $\|\cdot\|$ and dual \mathcal{H}^* . Assume that $(u_n, n \in \mathbb{N})$ is a sequence in \mathcal{H} , such that for every $v \in \mathcal{H}$, $((u_n, v), n \in \mathbb{N})$ is Cauchy. We have to show that there exists $u \in \mathcal{H}$ so that for every $v \in \mathcal{H}$, we have that $(u_n - u, v) \rightarrow 0$, as n tends to infinity.

Let $v \in \mathcal{H}$. Because $((u_n, v), n \in \mathbb{N})$ is a Cauchy sequence it converges, and therefore this sequence is bounded. I.e., $((u_n, \cdot), n \in \mathbb{N})$ is pointwise bounded on \mathcal{H} . The Banach–Steinhaus–theorem implies that $((u_n, \cdot), n \in \mathbb{N})$ is bounded in \mathcal{H}^* . By the Riesz representation theorem we know that the norm of (u_n, \cdot) in \mathcal{H}^* is equal to $\|u_n\|$. Thus $(u_n, n \in \mathbb{N})$ is bounded in \mathcal{H} , say $\|u_n\| \leq M$, $M > 0$, for all $n \in \mathbb{N}$.

For $v \in \mathcal{H}$, set

$$L(v) := \lim_{n \rightarrow \infty} (u_n, v).$$

Clearly, $v \mapsto L(v)$ is linear (as the limit of linear mappings). Moreover it is continuous:

$$\begin{aligned} |L(v_1) - L(v_2)| &= \left| \lim_{n \rightarrow \infty} (u_n, v_1 - v_2) \right| \\ &\leq \limsup_{n \rightarrow \infty} |(u_n, v_1 - v_2)| \\ &\leq M \|v_1 - v_2\|. \end{aligned}$$

Therefore, by the Riesz representation theorem there exists $u \in \mathcal{H}$ with $L_v = (u, v)$. \square

Theorem C.2 Every bounded sequence in a Hilbert space contains a weakly convergent subsequence.

Proof Let $(u_n, n \in \mathbb{N})$ be a sequence in a Hilbert space \mathcal{H} which is bounded by $M > 0$. Let \mathcal{H}_0 be the closure of the span of the elements in $(u_n, n \in \mathbb{N})$. Then \mathcal{H}_0 is a Hilbert subspace of \mathcal{H} .

First we show that there is a subsequence $(u_{n'})$ and an element $u \in \mathcal{H}_0$ so that for all $v \in \mathcal{H}_0$, $(u_{n'}, v) \rightarrow (u, v)$ as $n' \rightarrow \infty$. By construction, \mathcal{H}_0 is separable, and we let $(v_k, k \in \mathbb{N})$ be a dense subset. Then for each k , $((u_n, v_k), n \in \mathbb{N})$ is a sequence of real numbers, bounded by $M \cdot \|v_k\|$. The Bolzano–Weierstraß theorem implies that this sequence contains a convergent subsequence. By the diagonal process we can hence extract a subsequence $(u_{n'})$ so that for every $k \in \mathbb{N}$, $(u_{n'}, v_k)$ converges. I.e., the subsequence $(u_{n'})$ converges weakly on a dense subspace of \mathcal{H}_0 . Since $(u_{n'})$ is bounded, $(u_{n'})$ converges weakly on all of \mathcal{H}_0 . (Let $v \in \mathcal{H}_0$ and $\varepsilon > 0$ be given. Choose $k \in \mathbb{N}$ so that $\|v - v_k\| < \varepsilon/2M$, and choose n'_0 large enough so that for all $n' \geq n'_0$, $|(u_{n'}, v_k)| \leq \varepsilon/2$. Then

$$\begin{aligned} |(u_{n'}, v)| &\leq |(u_{n'}, v_k)| + |(u_{n'}, v - v_k)| \\ &\leq \frac{\varepsilon}{2} + M \|v - v_k\| \\ &\leq \varepsilon. \end{aligned}$$

In particular, for every $v \in \mathcal{H}_0$, $((u_{n'}, v))$ is a Cauchy sequence: $(u_{n'})$ is weakly Cauchy in \mathcal{H}_0 . Since \mathcal{H}_0 is Hilbert we can apply Lemma C.1 so that there exists $u \in \mathcal{H}_0$ with $u_{n'} \rightarrow u$ weakly in \mathcal{H}_0 .

Now let $w \in \mathcal{H}$. By the projection theorem (\mathcal{H}_0 is a closed subspace of \mathcal{H}), w admits an orthogonal decomposition $w = w_0 + w^\perp$ with $w_0 \in \mathcal{H}_0$, w^\perp orthogonal to \mathcal{H}_0 . In particular, w^\perp is orthogonal to the elements of the sequence $(u_{n'})$ and to u : $(u_{n'}, w^\perp) = 0$ for all n' , $(u, w^\perp) = 0$. Therefore $(u_{n'} - u, w) = (u_{n'} - u, w_0) \rightarrow 0$ as $n' \rightarrow \infty$. \square

Remark The reader might think that we can conclude the statement of this theorem directly from the Banach–Alaoglu theorem. This is not so, because the weak topology is in general not metrizable, and therefore sequential compactness does not follow from compactness.

Appendix D: Smoothing in $\mathcal{L}^p(\mathbb{R})$

A nice reference for the results in this Appendix, is the book [Zy 88], chapters I and II. (Zygmund considers there (periodic) functions on an interval, but the methods extend without effort to functions on \mathbb{R} or \mathbb{R}^d .)

Lemma D.1 Let $f \in \mathcal{L}^p(\mathbb{R})$, $p \geq 1$, and for $y \in \mathbb{R}$ set $f_y(x) := f(x - y)$, $x \in \mathbb{R}^d$. Then $\|f_y - f\|_p$ converges to zero with $y \rightarrow 0$.

We give a proof which is a modification of the proof of Lemma 23.1 in Bauer (1991).

Proof It suffices to consider $y \in \mathbb{R}$ with $|y| \leq 1$.

First let $f \in C_c(\mathbb{R})$ and choose a compact set K in \mathbb{R} large enough so that for all $y \in \mathbb{R}$ with $|y| \leq 1$, $\text{supp} f_y \subset K$. f is uniformly continuous on K so that we find for given $\varepsilon > 0$ a $\delta > 0$ so that $|x - x'| < \delta$, $x, x' \in K$, implies $|f(x) - f(x')| < \varepsilon$. Without loss of generality we may choose $\delta \leq 1$. But then we have for all $y \in \mathbb{R}$ with $|y| < \delta$ the inequality $|f_y(x) - f(x)| < \varepsilon$ for all $x \in K$. From this we conclude

$$\begin{aligned} \int_{\mathbb{R}} |f_y(x) - f(x)|^p dx &= \int_K |f_y(x) - f(x)|^p dx \\ &\leq \varepsilon^p |K|. \end{aligned}$$

Now let $f \in \mathcal{L}^p(\mathbb{R})$, and let $\varepsilon > 0$ be given. Choose $f' \in C_c(\mathbb{R})$ so that $\|f - f'\|_p \leq \varepsilon$. Then we also have $\|f_y - f'_y\|_p \leq \varepsilon$, and

$$\begin{aligned} \|f_y - f\|_p &\leq \|f_y - f'_y\|_p + \|f - f'\|_p + \|f'_y - f'\|_p \\ &\leq 2\varepsilon + \|f'_y - f'\|_p, \end{aligned}$$

and the last term vanishes as $y \rightarrow 0$ by the first step. □

As usual, we denote for two functions f, g their convolution by $f * g$, i.e.,

$$(f * g)(x) = \int_{\mathbb{R}} f(x - y) g(y) dy, \quad x \in \mathbb{R},$$

whenever the integral on the right hand side is defined for every (or λ -almost every) $x \in \mathbb{R}$. For example, $f * g$ is everywhere defined, if $f \in \mathcal{L}^p(\mathbb{R})$ and $g \in \mathcal{L}^q(\mathbb{R})$ with $p^{-1} + q^{-q} = 1$. Moreover in this case we have the following:

Lemma D.2 Assume that $f \in \mathcal{L}^p(\mathbb{R})$, $g \in \mathcal{L}^q(\mathbb{R})$ with $p^{-1} + q^{-1} = 1$. Then $f * g$ is uniformly continuous.

Proof Let $x, y \in \mathbb{R}$, then

$$|f * g(x + y) - f * g(x)| \leq \|f_y - f\|_p \|g\|_q$$

by Hölder's inequality. Now we can apply Lemma D.1. □

Lemma D.3 Let $(E_i, \mathcal{E}_i, \mu_i)$, $i = 1, 2$, be two σ -finite measure spaces, f a positive $\mathcal{E}_1 \otimes \mathcal{E}_2$ - $\mathcal{B}(\mathbb{R})$ -measurable mapping from $E_1 \times E_2$ into \mathbb{R} . For $p \geq 1$,

$$\left\| \int f(\cdot, y) d\mu_2(y) \right\|_{\mathcal{L}^p(\mu_1)} \leq \int \|f(\cdot, y)\|_{\mathcal{L}^p(\mu_1)} d\mu_2(y).$$

The proof of Lemma D.3 follows the same pattern as the proof of Minkowski's inequality on the basis of Hölder's, and is left as an exercise to the interested reader.

Lemma D.4 (*Young's Inequality*) Let $f \in \mathcal{L}^p(\mathbb{R})$, $g \in \mathcal{L}^q(\mathbb{R})$, $p, q \geq 1$, with $p^{-1} + q^{-1} > 1$. Let r be given by $r^{-1} = p^{-1} + q^{-1} - 1$. Then $f * g \in \mathcal{L}^r(\mathbb{R})$, and

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

A proof of Lemma D.4 can be found, e.g., in [Zy 88], p. 37 f, or in [RS 75], p. 28 f.

Let $p \geq 1$, and let $\varphi \in \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^q(\mathbb{R})$, $p^{-1} + q^{-1} = 1$ with $\varphi \geq 0$, $\|\varphi\|_1 = 1$. For $n \in \mathbb{N}$, set

$$\varphi_n(x) := n \varphi(nx), \quad x \in \mathbb{R}.$$

Assume that $I \subset \mathbb{R}$ is an interval, where we also allow $\pm\infty$ as endpoints. Furthermore let $f \in \mathcal{L}^p(I)$. We associate with f its trivial extension $\tilde{f} \in \mathcal{L}^p(\mathbb{R})$, given by

$$\tilde{f}(x) = \begin{cases} f(x) & x \in I \\ 0 & \text{otherwise.} \end{cases}$$

Set

$$\tilde{f}_n(x) := \varphi_n * \tilde{f}(x), \quad x \in \mathbb{R}, n \in \mathbb{N},$$

which is well-defined and uniformly continuous by Lemma D.2. We denote by f_n the restriction of \tilde{f}_n , $n \in \mathbb{N}$, to I .

Lemma D.5 For every $n \in \mathbb{N}$, $f_n \in \mathcal{L}^p(I)$. Moreover, $(f_n, n \in \mathbb{N})$ converges in $\mathcal{L}^p(I)$ to f .

Proof The first statement follows from Young's inequality (Lemma D.4) for $r = q = 1$. For the second consider

$$\begin{aligned} \|f - f_n\|_{\mathcal{L}^p(I)}^p &\leq \|\tilde{f} - \tilde{f}_n\|_{\mathcal{L}^p(\mathbb{R})}^p \\ &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} (\tilde{f}(x) - \tilde{f}(x-y)) \varphi_n(y) dy \right|^p dx \\ &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} \varphi(y) (\tilde{f}(x) - \tilde{f}(x - \frac{y}{n})) dy \right|^p dx \\ &\leq \left(\int_{\mathbb{R}} \varphi(y) \|\tilde{f} - \tilde{f}_{-y/n}\|_{\mathcal{L}^p(\mathbb{R})} dy \right)^p, \end{aligned}$$

where we used Lemma D.3 in the last step. By Lemma D.1, the integrand of the last integral converges pointwise to zero as $n \rightarrow \infty$. On the other hand,

$$\begin{aligned} \|\tilde{f}_{-y/n}\|_p &= \|\tilde{f}\|_p \\ &= \|f\|_{\mathcal{L}^p(I)}, \end{aligned}$$

so that $2\varphi \|f\|_{\mathcal{L}^p(I)}$ is a uniform majorant. Thus the result follows from an application of the dominated convergence theorem. \square

References

- H. Bauer, *Maß- und Integrationstheorie*, de Gruyter, Berlin, New York (1990).
- H. Bauer, *Wahrscheinlichkeitstheorie*, de Gruyter, Berlin, New York (1991).
- K.L. Chung, J.L. Doob, Fields, optionality and measurability, *Amer. J. Math.* **87** (1964) 397–424.
- K.L. Chung, R.J. Williams, *Introduction to Stochastic Integration*, Birkhäuser, Boston (1983).
- C. Dellacherie, P.-A. Meyer, *Probabilities and Potential*, North-Holland, Amsterdam, New York, Oxford (1978).
- J.L. Doob, *Stochastic Processes*, Wiley, New York, London, Sydney (1953).
- E.B. Dynkin, *Markov Processes*, Springer, Berlin, Heidelberg, New York (1965).
- I. Karatzas, S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, Berlin, Heidelberg, New York (1991).
- A. Friedman, *Stochastic differential equations and applications*, Academic Press, New York, San Francisco, London (1975).
- S. Lang, *Real Analysis*, Addison-Wesley, Reading (1969).
- P.-A. Meyer, *Probabilités et Potentiel*, Hermann, Paris (1966).
- Ph. Protter, *Stochastic Integration and Differential Equations — A New Approach*, Springer, Berlin, Heidelberg, New York (1990).
- D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, Springer, Berlin, Heidelberg, New York (1991).
- M. Reed, B. Simon, *Methods of Modern Mathematical Physics II — Fourier Analysis, Self-Adjointness*, Academic Press, New York, San Francisco, London (1975).
- F. Riesz, B. Sz.-Nagy, *Functional Analysis*, Frederick Ungar Publ. Co., New York (1955).
- H. v. Weizsäcker, G. Winkler, *Stochastic Integrals*, Vieweg, Braunschweig (1990).
- A. Zygmund, *Trigonometric Series*, Cambridge University Press, Cambridge (1988).