On Measurability Properties of Stochastic Processes

Svenja Kaden and Jürgen Potthoff

On Measurability Properties of Stochastic Processes

Svenja Kaden and Jürgen Potthoff

Universität Mannheim
D–68131 Mannheim
Germany

Dedicated to Professor Ernst Binz at the occasion of his 60th birthday

1. Introduction

This paper does not contain anything new, and, in fact, it deals with a rather old–fashioned subject: the theorem of Chung–Doob [CD 64] and Meyer [Me 66] (cf. also [DM 78]) on the existence of a progressively measurable modification of a measurable, adapted process.

The *raison d'être* of this paper is that in our seminar at Mannheim periodically with new students the need comes up to understand the construction of the Itô–integral, and that in most of the common literature the details of a certain argument (see below) are not given or only hinted at. This paper attempts to provide that argument in rather much — maybe even unnecessary — detail, so that the student can understand the probably most direct approach to the stochastic integral with respect to a Brownian motion as the limit of approximations of Riemann type.

Consider a Brownian motion $B = (B_t, t \in \mathbb{R}_+)$ on some probability space $(\Omega, \mathcal{A}, P)$, and suppose that we are given a stochastic process $X = (X_t, t \in \mathbb{R}_+)$. One is interested to construct the Itô integral $\int_a^b X_t \, dB_t$, $0 \le a < b \le +\infty$, 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 that ii) X belongs to $\mathcal{L}^2(\Omega \times [a, b])$. Then one tries to construct $\int_a^b X_t \, dB_t$ as the $\mathcal{L}^2(\Omega)$–limit of approximations of Riemann–type

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

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., [RY 91], [Pr 90], [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 by first making X uniformly bounded by cutting it off, and then making the resulting process — say $\widehat{X}$ — pathwise continuous by setting

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

(Here we have not been very careful about what happens for t and a near zero, cf. section 5 for a more careful discussion.) The problem is, that even though the cutting procedure preserves the adaptedness of the process, the last operation does in general *not*, because it involves uncountable many values of the time parameter. But then we cannot use the Itô isometry for the above sum, and our construction 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 $\mathcal{F}$. Roughly, this means that for given $t \in \mathbb{R}_+$, X as a function of *both* variables is $\mathcal{F}_t \otimes \mathcal{B}([0,t])$–measurable when they range in $\Omega \times [0,t]$. Then it follows from the theorem of Fubini–Tonelli that X^n above is actually adapted. But again this assumption entails some incompatibility with the assumptions usually made in the literature. On the other hand, Chung and Doob proved in [CD 64] that every measurable, separable and adapted process has a progressive modification, and in the book by Meyer [Me 66] (which is most of the time quoted for this result) it is stated that this holds also without the assumption of separability. The argument in [Me 66] has been judged by Karatzas and Shreve in their book [KS 88], p. 5, to be "lengthy and rather demanding". Actually, the proof in [Me 66] has a gap, which has been fixed in [DM 78]. In fact, the proof in [DM 78] makes use of rather heavy machinery: the Dunford–Pettis compactness theorem and a version of the Eberlein–Šmulian theorem. In [KS 88] then the progressive version of X is used to construct its Itô–integral (under the additional assumption that the underlying probability space is complete).

The proof in [DM 78] of the theorem of Chung, Doob and Meyer (without the separability assumption; — it is a part of Theorem IV.30 in [DM 78]) has essentially two steps. The first is to establish that a measurable process X has an approximation by processes of the form

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

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 step, one replaces the $H_{n,k}$ (which are close to X on $A_{n,k}$) by values of X, so that the expressions $1_{A_{n,k}} H_{n,k}$ 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 $1_{[s_{n,k}, t_{n,k}]} X_{s_{n,k}}$ 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 decrease to the infimum of $A_{n,k}$. Once one has the new progressive processes, say $(\widetilde{X}_n, n \in \mathbb{N})$, one may choose

$$\widetilde{X}(t) = \limsup_n \widetilde{X}_n(t)$$

to obtain a progressive modification of X.

2

Essentially, in the present paper this proof is carried 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 rather simple (though still somewhat lengthy if carried out with all the details).

Finally, we combine this result with the theory of smoothing in $\mathcal{L}^p(\mathbb{R})$, as it can, e.g., be found in [Zy 88], to obtain a result which can be used as a basis for the usual construction of the Itô–integral (under the usual assumptions on X): Every process X which belongs to $\mathcal{L}^2(\Omega \times \mathbb{R}_+)$ and is adapted to $\mathcal{F}$ has an approximation in $\mathcal{L}^2(\Omega \times \mathbb{R}_+)$ by a sequence of continuous $\mathcal{F}$–adapted processes.

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 (where often the underlying σ–algebra and the filtration generated by the integrator process are used "as is").

The organization of the paper is as follows. In section 2 we prepare a standard lemma in measure theory. Section 3 is devoted to the (pointwise) approximation of a measurable process X by processes of the form (1). Section 4 deals with the proof of the above described version of the theorem of Chung, Doob and Meyer. In section 5 we combine this theorem with $\mathcal{L}^p$–smoothing. There are four appendices which are supposed to make this paper rather self-contained in view of its above mentioned purpose. 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. Measurability and Pointwise Limits

Let $(E, \mathcal{E})$ be a measurable space, (M, d) a metric space, which we consider as equipped with its Borel σ–algebra $\mathcal{B}_d$.

1 Lemma 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 case that $M = \mathbb{R}$ one finds this result in every textbook on measure and integration theory. However, most of the proofs make use of the lattice structure of $\mathbb{R}$ (e.g., [Ba 90]), which is not available in the general case.

Proof Because the family of open sets is a generator of $\mathcal{B}_d$ which is stable with respect to intersections, it suffices to show that for every open set O in M we have $f^{-1}(O) \in \mathcal{E}$.

First we remark that if U is open, and $x \in f^{-1}(U)$, then U contains a neighborhood of $f(x)$. Since $f(x) = \lim_{n \to \infty} f_n(x)$, we obtain

$$x \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} f_k^{-1}(U).$$

Hence

$$f^{-1}(U) \subset \liminf_n f_n^{-1}(U) \subset \limsup_n f_n^{-1}(U).$$

Next we consider a closed subset A of M. It is straightforward to see that

$$\limsup_n f_n^{-1}(A) \subset f^{-1}(A).$$

(Let us check this: If $x \in \limsup_n f_n^{-1}(A)$, then x is contained in infinitely many sets $f_n^{-1}(A)$. I.e., for a subsequence $(f_{n'})$ we have $f_{n'}(x) \in A$ for all n'. Since this subsequence converges to $f(x)$ and A is closed, $x \in f^{-1}(A)$ follows.)

Now let O be open in M, and define for $m \in \mathbb{N}$

$$U_m := \left\{ y \in O;\, d(y, \partial O)) > \frac{1}{m} \right\},$$

$$A_m := \left\{ y \in O;\, d(y, \partial O) \geq \frac{1}{m} \right\}.$$

Then we have for all $m \in \mathbb{N}$, $U_m \subset A_m \subset O$, and $O = \bigcup_n A_m = \bigcup_m U_m$. Therefore

$$f^{-1}(O) = \bigcup_{m=1}^{\infty} f^{-1}(U_m)$$

$$\subset \bigcup_{m=1}^{\infty} \limsup_n f_n^{-1}(U_m)$$

and

$$f^{-1}(O) = \bigcup_{m=1}^{\infty} f^{-1}(A_m)$$

$$\supset \bigcup_{m=1}^{\infty} \limsup_n f_n^{-1}(A_m)$$

$$\supset \bigcup_{m=1}^{\infty} \limsup_n f_n^{-1}(U_m).$$

Consequently

$$f^{-1}(O) = \bigcup_{m=1}^{\infty} \limsup_{n} f_n^{-1}(U_m)$$

$$= \bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} f_k^{-1}(U_m),$$

and therefore $f^{-1}(O) \in \mathcal{E}$. □

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

Lemma 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 We construct for given $n \in \mathbb{N}$ a countable 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

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

f_n is clearly of the desired type. Now 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 $d(f_n(x), f(x)) < 1/n$. □

3 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 both lemmas into the following

Corollary 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.

3. Measurability of Stochastic Processes

Let $(\Omega, \mathcal{A}, P)$ be a probability space and let T be an interval in $\mathbb{R}_+$ or let $T = \mathbb{R}_+$. A real valued *stochastic process* $X = (X_t, t \in T)$ *with time parameter in* T is a family of real valued random variables indexed by T. (In this paper we will not distinguish between real valued and extended real valued functions; — there will be no danger of confusion.) We say that $\widetilde{X}$ is a *modification* of X, if for every $t \in T$, $P(\widetilde{X}_t \neq X_t) = 0$ holds. We equip T with its Borel σ–algebra $\mathcal{B}(T)$ and $\Omega \times T$ with the product σ–algebra $\mathcal{A} \otimes \mathcal{B}(T)$. According to convenience we shall write $X_t(\omega)$ as well as $X(\omega, t)$ for the evaluation of a stochastic process X at $(\omega, t) \in \Omega \times T$.

1 Definition A stochastic process X is called *measurable*, if the mapping

$$X : \Omega \times T \longrightarrow \mathbb{R}$$

is $\mathcal{A} \otimes \mathcal{B}(T)$–$\mathcal{B}(\mathbb{R})$–measurable.

2 Let a process X be given. Then we can associate with X the mapping

$$\widehat{X} : T \longrightarrow \mathcal{R}_P$$
$$t \longmapsto \widehat{X}_t = [X_t]_P,$$

where $\mathcal{R}_P$ is the space of P–equivalence classes $[Y]_P$ of real valued random variables Y on $(\Omega, \mathcal{A}, P)$.

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 $L^2(P)$. 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.) *uniformly bounded*, if there is $M > 0$ so that for all $t \in T$ and (P–a.e.) $\omega \in \Omega$, $|X(\omega, t)| \leq M$.

3 Lemma Assume that $X = (X_t, t \in T)$ is a measurable stochastic process, so that for all $t \in T$, 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 $E = T$, $\mathcal{E} = \mathcal{B}(T)$, $M = \mathcal{R}_{P,b}$, 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 uniformly bounded. Assume that we have shown that for every measurable, uniformly bounded process Y the properties (i) and (ii) hold for $\widehat{Y}$. Let X be a process as in the hypotheses of the lemma, and for $n \in \mathbb{N}$ set

$$X_t^{(n)}(\omega) := (X_t(\omega) \wedge n) \vee (-n), \qquad \omega \in \Omega, \, t \in T.$$

6

Then $X^{(n)}$ is uniformly 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 $t \in T$, $X_t^{(n)}$ 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 $\widehat{X}^{(n)}$ converges to $\widehat{X}$ in $(\mathcal{R}_{P,b}, \|\cdot\|_2)$. $\widehat{X}^{(n)}$ being measurable, we obtain from lemma 1.1, that so is $\widehat{X}$. 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 uniformly bounded.

We denote by $\mathcal{H}$ the set of all uniformly 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 hypotheses of the monotone class theorem (theorem 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 $\Omega \times T$ to X. Then in particular for every $t \in T$, $(X_n(\cdot, t), n \in \mathbb{N})$ converges uniformly on Ω to $X(\cdot, t)$, and therefore for every $t \in T$, $(\widehat{X}_n(\cdot, t), n \in \mathbb{N})$ converges in $L^2(P)$ to $\widehat{X}(\cdot, t)$. By assumption, for every $n \in \mathbb{N}$ the mapping $t \longmapsto \widehat{X}_n(\cdot, t)$ from T into $(\mathcal{R}_{P,b}, \|\cdot\|_2)$ is measurable, and by lemma 1.1, the same holds for $t \longmapsto \widehat{X}(\cdot, t)$. Moreover, since for all $n \in \mathbb{N}$, $\widehat{X}_n$ is separably valued, so is $\widehat{X}$. (Set $M_{00} = \bigcup_n \widehat{X}_n(T)$, 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 $t \in T$ we have $\widehat{X}(t) = \lim_n \widehat{X}_n(t)$, $\widehat{X}_n(t) \in M_0$. 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.

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 uniformly bounded, too. By the dominated convergence theorem, for every $t \in T$, $\widehat{X}_n(\cdot, t)$ converges to $\widehat{X}(\cdot, t)$ 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

$$X_t(\omega) = 1_A(\omega) \, 1_{[a,b]}(t), \qquad \omega \in \Omega, t \in T,$$

with $a, b \in T$, $a < t$, $A \in \mathcal{A}$. Clearly, $\mathcal{C}$ is closed under multiplication, and $\sigma(\mathcal{C}) = \mathcal{A} \otimes \mathcal{B}(T)$. By the monotone class theorem (theorem 1 in appendix A) $\mathcal{H}$ contains all uniformly bounded, measurable processes. $\qquad \square$

4 Let X be as in the hypothesis of lemma 3.3, and let the sequence in statement (ii) of lemma 3.3 be denoted by $(\widehat{X}_n, n \in \mathbb{N})$. Then for $n \in \mathbb{N}$, $\widehat{X}_n$ is of the following form

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

where $(A_{n,k}, k \in \mathbb{N})$ is a sequence of pairwise disjoint Borel subsets which form a partition of T and $(H_{n,k}, k \in \mathbb{N})$ is a sequence of random variables. Hence statement (ii) of lemma 3.3 reads explicitly as follows:

Corollary There exists a sequence $(X_n, n \in \mathbb{N})$ of processes of the form

$$X_n(\omega, t) = \sum_{k=1}^{\infty} 1_{A_{n,k}}(t) \, H_{n,k}(\omega), \qquad \omega \in \Omega, \, t \in T,$$

so that

$$\sup_{t \in T} \left(\mathbb{E}\big(X(\cdot, t) - X_n(\cdot, t)\big)^2 \right)^{1/2} \xrightarrow[n \to \infty]{} 0.$$

4. Progressive Stochastic Processes

Let $(\Omega, \mathcal{A}, P)$ be a probability space. For convenience we choose for T either an interval of the form $[0, a]$, $a > 0$, or $\mathbb{R}_+$. A *filtration* $\mathcal{F}$ on $(\Omega, \mathcal{A})$ is a family $(\mathcal{F}_t, t \in T)$ of σ–algebras over Ω indexed by T, so that $t > s$ implies $\mathcal{F}_t \supset \mathcal{F}_s$. We assume throughout that for every $t \in T$, $\mathcal{F}_t \subset \mathcal{A}$.

1   Definition Suppose that $X = (X_t, t \in T)$ is a stochastic process and $\mathcal{F} = (\mathcal{F}_t, t \in T)$ is a filtration on $(\Omega, \mathcal{A}, P)$.

(a) X is called $\mathcal{F}$–*adapted*, if for every $t \in T$, X_t is $\mathcal{F}_t$–measurable.

(b) X is called $\mathcal{F}$–*progressive*, if for every $t \in T$, the restriction of X to $\Omega \times [0, t]$ is $\mathcal{F}_t \otimes \mathcal{B}([0, t])$–measurable.

Whenever it is clear from the context which filtration is meant, we shall also simply say "adapted", "progressive" respectively.

If X is $\mathcal{F}$–progressive, then the theorem of Fubini–Tonelli implies that X is $\mathcal{F}$–adapted. The converse is in general false.

In order to prove the main result of this section — a version of the well–known theorem of Chung, Doob and Meyer — we prepare a couple of simple lemmas.

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

$$f_i \otimes f_2(x, y) := f_1(x) f_2(y), \qquad x \in E_1, \, y \in E_2.$$

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

The proof 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 their positive and negative parts. Details are left to the interested reader.

8

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

Lemma Let $A \in \mathcal{B}(T)$, $\tau = \inf A$, 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 $\mathcal{F}_\tau$–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 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 4.2 to conclude that the restriction of Y to $\Omega \times [0, t]$ is $\mathcal{F}_t \otimes \mathcal{B}([0, t])$–measurable. $\qquad\square$

4 **Lemma** Suppose that X is an $\mathcal{F}$–adapted process, $t \in T$, 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 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

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

5   Theorem   Assume that X is a measurable, $\mathcal{F}$–adapted process. Then it has a measurable, $\mathcal{F}$–progressive modification.

Proof   We do the proof in several steps.

Step 1:   With the following argument we reduce to the case that X is in addition uniformly bounded. Suppose that we can show the statement of the theorem for every uniformly bounded, measurable and $\mathcal{F}$–adapted process, and let X be as in the hypothesis. Consider $Y = \arctan X$. Then Y is uniformly bounded, and by our assumption Y has an $\mathcal{F}$–progressive modification $\widetilde{Y}$. Set $\widetilde{X} = \tan \widetilde{Y}$. Then $\widetilde{X}$ is progressive, too, and for $t \in T$ we have $P(X_t = \widetilde{X}_t) = P\big(\arctan X_t = \widetilde{Y}_t\big) = P(Y_t = \widetilde{Y}_t) = 1$. Hence $\widetilde{X}$ is also a modification of X. Therefore, we can from now on assume that X is uniformly bounded.

Step 2:   By corollary 3.4 we know that there exists a sequence of processes $(X_n, n \in \mathbb{N})$ of the form

$$X_n(t) = \sum_{k=1}^{\infty} H_{n,k}\, 1_{A_{n,k}}(t), \qquad t \in T,$$

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

$$\lim_{n \to \infty} \sup_{t \in T} \|X(t) - X_n(t)\|_2 = 0.$$

By choosing a subsequence if necessary we may assume that for every $n \in \mathbb{N}$ we have

$$\sup_{t \in T} \|X(t) - X_n(t)\|_2 \leq 2^{-n}.$$

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

Then we have

$$\sup_{t \in A_{n,k}} \left\| X(t) - \widetilde{X}_{n,k} \right\|_2$$
$$\leq \sup_{t \in A_{n,k}} \left(\|X(t) - H_{n,k}\|_2 + \left\| H_{n,k} - X_{\tau_{n,k}} \right\|_2 \right)$$
$$\leq 2^{-(n-1)}.$$

By lemma 4.3a, $\widetilde{X}_{n,k} \otimes 1_{A_{n,k}}$ is $\mathcal{F}$–progressive.

10

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 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 $\widetilde{X}_{n,k}$. By lemma 4.4a, we may choose $\widetilde{X}_{n,k}$ as $\mathcal{F}_{\tau_{n,k}+}$–measurable. Therefore the process $\widetilde{X}_{n,k} \otimes 1_{A_{n,k}}$ is $\mathcal{F}$–progressive by lemma 4.3b. Moreover, by lemma 4.4b we have

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

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

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

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

In view of lemma 4.2 it is clear that $\widetilde{X}_n$ is measurable and by the previous step, for every $n \in \mathbb{N}$, $\widetilde{X}_n$ is $\mathcal{F}$–progressive. Moreover we have

$$\sup_{t \in T} \|X(t) - \widetilde{X}_n(t)\|_2 \le 2^{-(n-1)}.$$

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

Step 5: Define
$$\widetilde{X} := \limsup_n \widetilde{X}_n.$$

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

5. Approximation of Adapted Stochastic Processes

Throughout this section we suppose that $X = (X_t, t \in T)$ is a real valued stochastic process, which belongs to $\mathcal{L}^2(\Omega \times T, \mathcal{A} \otimes \mathcal{B}(T), P \otimes \lambda)$ and which is adapted to a filtration $\mathcal{F} = (\mathcal{F}_t, t \in T)$. As before, we assume that $T = [0, a]$, $a > 0$, or $T = \mathbb{R}_+$.

1 Definition A real valued stochastic process $Y = (Y_t, t \in T)$ is called P–a.s. *(uniformly) sample continuous* if there exists a P–null set $N \in \mathcal{A}$ so that for all $\omega \in \complement N$, the mapping $t \longmapsto Y(t, \omega)$ from T into $\mathbb{R}$ is (uniformly) continuous.

11

2 Theorem There exists a sequence $(X_n, n \in \mathbb{N})$ of uniformly bounded, measurable, $\mathcal{F}$–adapted processes which are P–a.s. uniformly sample continuous, so that

$$\lim_{n \to \infty} \|X - X_n\|_{\mathcal{L}^2(P \otimes \lambda)} = 0.$$

Proof Consider a measurable, $\mathcal{F}$–progressive modification $\widetilde{X}$ of X, which exists according to theorem 4.5. Because $\widetilde{X}$ is a modification of X, we have that for every $t \in T$,

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

The theorem of Fubini–Tonelli entails that

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

Therefore it suffices to show the statement for $\widetilde{X}$.

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

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

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

Let $\omega \in \Omega$ and extend $\widetilde{X}(\omega, \cdot)$ to all of $\mathbb{R}$ by setting it identically zero outside T. Let $\varphi = 1_{[0,1]}$, and for $n \in \mathbb{N}$, $\varphi_n(u) = n\,\varphi(n\,u)$, $u \in \mathbb{R}$. Put

$$X_n(\omega, t) := \big(\varphi_n * \widetilde{X}(\omega, \cdot)\big)(t), \qquad t \in \mathbb{R},$$

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

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

as $n \longrightarrow \infty$.

It remains to show that X_n is $\mathcal{F}$–adapted. Let $t \in T$. Writing the convolution $\varphi_n * \widetilde{X}$ as an integral, we have

$$X_n(\omega, t) = n \int_{t-1/n}^{t} \widetilde{X}(\omega, s) \, ds$$

$$= n \int_{0}^{t} 1_{[(t-1/n) \vee 0, t]}(s) \, \widetilde{X}(\omega, s) \, ds.$$

12

$\widetilde{X}$ 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 $\omega \longmapsto X_n(\omega, t)$ is $\mathcal{F}_t$–measurable. $\qquad\square$

The last theorem is a convenient starting point for the construction of the Itô integral $\int_0^t X_s \, dB_s$ of the process X against a Brownian motion via approximation by Riemann type sums.

Appendix A: Monotone Class Theorem

We prepare a powerful version of the monotone class theorem which can be found, e.g., in [DM 78].

1 **Theorem**   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 $n \in \mathbb{N}$, $a_0, a_1, \ldots, a_n \in \mathbb{R}$, $f_1, \ldots, f_n \in \mathcal{C}$. 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 $\mathcal{T}_1 \subset \mathcal{T}_2 \subset \ldots \subset \mathcal{T}_n \subset \ldots$ 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 $f g \in \mathcal{T}_k$, and $f g \in \bigcup_k \mathcal{T}_k$. Similarly, we can show that $f + a g$ 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 $\mathcal{T}_1 \subset \mathcal{T}_2 \subset \ldots \subset \mathcal{T}_k \subset \ldots$ 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

13

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 $x \longmapsto |x|$ 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 $g^\pm f_n^p$ increase to $g^\pm f^p$. Therefore we have $gf^p \in \mathcal{H}$. 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
$$g = (f \wedge 1) \vee 0 \in \mathcal{T}_0.$$

14

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

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

Let $(\Omega, \mathcal{A}, P)$ 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

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

1   Lemma 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 $k \geq n$,

$$\mathbb{E}\big(Y\,(\mathbb{E}(Z \,|\, \mathcal{A}_{-n}) - Z_k)\big) = \mathbb{E}\big(Y \,\mathbb{E}(Z - Z_k \,|\, \mathcal{A}_{-n})\big)$$
$$= \mathbb{E}\big(\mathbb{E}(Y \,|\, \mathcal{A}_{-n})\,(Z - Z_k)\big),$$

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

Consequently, for every $n \in \mathbb{N}$ there exists $N_n \in \mathcal{N}$ with $Z = \mathbb{E}(X \,|\, \mathcal{A}_{-n})$ 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 $(\widetilde{Z}_n, n \in -\mathbb{N})$ given by $\widetilde{Z}_n := \mathbb{E}(Z \,|\, \mathcal{A}_n)$, $n \in -\mathbb{N}$, 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 $N \cup M$. $\qquad\square$

Remarks   The theorem of Doob we used in the last step 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

$$\widetilde{Z} := \begin{cases} Z & \text{on } \complement N \\ 0 & \text{on } N \end{cases}$$

after the first step in the proof. Then $\widetilde{Z}$ 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 in sections 4 and 5, one has then to augment the filtration considered there, too. This gives somewhat weaker results than theorem 4.5 and theorem 5.2, 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.

1   Lemma   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 $v\in\mathcal{H}$, $(u_n - u, v)\longrightarrow 0$, 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

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

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

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

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

16

2   **Theorem**   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 $\mathcal{H}$.

We show first that there is a subsequence $(u_{n'})$ and an element $u \in \mathcal{H}_0$ so that for all $v \in \mathcal{H}_0$, $(u_{n'}, v) \longrightarrow (u, v)$ as $n' \to \infty$. 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 $M \|v_k\|$. 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 1 so that there exists $u \in \mathcal{H}_0$ with $u_{n'} \longrightarrow u$ 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 $(u_{n'} - u, w) = (u_{n'} - u, w_0) \longrightarrow 0$ as $n' \to \infty$. $\qquad\square$

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, which are all well–known, 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 $\mathbb{R}^d$.)

1 Lemma 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 $y \to 0$.

The following proof is taken from the book [Ba 91] of H. Bauer (Lemma 23.1).

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 K in $\mathbb{R}$ large enough so that for all $y \in \mathbb{R}$ with $|y| \leq 1$, $\mathrm{supp} f_y \subset K$. 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

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

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

$$\|f_y - f\|_p \leq \|f_y - f'_y\|_p + \|f - f'\|_p + \|f'_y - f'\|_p$$
$$\leq 2\varepsilon + \|f'_y - f'\|_p \, ,$$

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

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

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

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 $f \in \mathcal{L}^p(\mathbb{R})$, $g \in \mathcal{L}^q(\mathbb{R})$ with $p^{-1} + q^{-q} = 1$. Moreover in this case we have the following:

Lemma 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 1. $\qquad \square$

3 Lemma 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 $p \geq 1$,

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

Proof The proof follows the same pattern as the proof of Minkowski's inequality on the basis of Hölder's.

For $p = 1$ we have nothing to prove, so let $p > 1$. Then, using the Fubini–Tonelli–theorem and Hölder's inequality, we get

$$\left\| \int f(\cdot, y)\, d\mu_2(y) \right\|_p^p = \int \left(\int f(x, y)\, d\mu_2(y) \right)^p d\mu_1(x)$$

$$= \int \left(\int f(x, y)\, d\mu_2(y) \left(\int f(x, y)\, d\mu_2(y) \right)^{p-1} \right) d\mu_1(x)$$

$$= \int \int f(x, y) \left(\int f(x, y)\, d\mu_2(x) \right)^{p-1} d\mu_2(y)\, d\mu_1(x)$$

$$= \int \int f(x, y) \left(\int f(x, y)\, d\mu_2(y) \right)^{p-1} d\mu_1(x)\, d\mu_2(y)$$

$$\leq \int \left(\int f(x, y)^p\, d\mu_1(x) \right)^{1/p}$$

$$\cdot \left(\left(\int f(x, y)\, d\mu_2(y) \right)^{q(p-1)} d\mu_1(x) \right)^{1/q} d\mu_2(y)$$

$$= \int \|f(\cdot, y)\|_p\, d\mu_2(y) \left\| \int f(\cdot, y)\, d\mu_2(y) \right\|_p^{p/q}.$$

And the result follows upon division by $\left\| \int f(\cdot, y)\, d\mu_2(y) \right\|_p^{p/q}$. $\square$

4 Lemma (*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 can be found, e.g., in [Zy 88], p. 37 f, or in [RS 75], p. 28 f.

5 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

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

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

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

Set

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

which is well–defined and uniformly continuous by no 2. We denote by f_n the restriction of $\widetilde{f}_n$, $n \in \mathbb{N}$, to I.

Lemma 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 4) for $r = q = 1$. For the second consider

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

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

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

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. $\qquad \square$

References

[Ba 90] H. Bauer, *Maß- und Integrationstheorie*, de Gruyter, Berlin, New York (1990)

[Ba 91] H. Bauer, *Wahrscheinlichkeitstheorie*, de Gruyter, Berlin, New York (1991)

[CD 64] K.L. Chung, J.L. Doob, Fields, optionality and measurability, *Amer. J. Math.* **87** (1964) 397–424

[DM 78] C. Dellacherie, P.–A. Meyer, *Probabilities and Potential*, North–Holland, Amsterdam, New York, Oxford (1978)

[Do 53] J.L. Doob, *Stochastic Processes*, Wiley, New York, London, Sydney (1953)

[Dy 65] E.B. Dynkin, *Markov Processes*, Springer, Berlin, Heidelberg, New York (1965)

[KS 88] I. Karatzas, S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, Berlin, Heidelberg, New York (1991)

[Fr 75] A. Friedman, *Stochastic differential equations and applications*, Academic Press, New York, San Francisco, London (1975)

[Me 66] P.–A. Meyer, *Probabilités et Potentiel*, Hermann, Paris (1966)

[Pr 90]   Ph. Protter, *Stochastic Integration and Differential Equations — A New Approach*, Springer, Berlin, Heidelberg, New York (1990)

[RY 91]   D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, Springer, Berlin, Heidelberg, New York (1991)

[RS 75]   M. Reed, B. Simon, *Methods of Modern Mathematical Physics II — Fourier Analysis, Self-Adjointness*, Academic Press, New York, San Francisco, London (1975)

[RN 55]   F. Riesz, B. Sz.–Nagy, *Functional Analysis*, Frederick Ungar Publ. Co., New York (1955)

[WW 90]   H. v. Weizsäcker, G. Winkler, *Stochastic Integrals*, Vieweg, Braunschweig, Wiesbaden (1990)

[Zy 88]   A. Zygmund, *Trigonometric Series*, Cambridge University Press, Cambridge (1988)