FOUNDATIONS OF MARTINGALE THEORY AND STOCHASTIC CALCULUS FROM A FINANCE PERSPECTIVE

JOSEF TEICHMANN

1. Introduction

The language of mathematical Finance allows to express many results of martingale theory via trading arguments, which makes it somehow easier to appreciate their contents. Just to provide one illustrative example: let $X = (X_n)_{n\geq 0}$ be a martingale and let $N_a^b(t)$ count the upcrossings over the interval $[a, b]$, then we can find a predictable process $V = (V_n)_{n\geq 0}$ such that

$$(b-a)N_a^b(t) \leq (a - X_t)^+ + (V \bullet X)_t\,,$$

where the right hand side's stochastic integral corresponds precisely to the cumulative gains and losses of a buy below a and sell above b strategy plus an option payoff adding for the risk of buying low and loosing even more. Immediately Doob's upcrossing inequality follows by taking expectations. We shall somehow focus on such types of arguments in the sequel, however, pushing them as far as possible. Still most of the following proofs are adapted from Olav Kallenberg's excellent book [3] or from the excellent lecture notes [6], maybe with a bit different wording in each case. For stochastic integration the reading of [6] is recommended.

Another important aspect is the two-sided character of many arguments, which leads, e.g., to reverse martingale results.

2. Filtrations, Stopping times, and all that

Given a probability space $(\Omega, \mathcal{F}, P)$, a *filtration* is an increasing family of sub-σ-algebras $(\mathcal{F}_t)_{t\in T}$ for a given index set $T \subset \mathbb{R} \cup \{\pm\infty\}$.

We shall often assume the "usual conditions" on a filtered probability space, i.e. that a filtration is right continuous and complete, but we first name the properties separately:

(1) A filtration is called *complete* if each $\mathcal{F}_t$ contains all P-null sets from $\mathcal{F}$ for $t \in T$. In particular every σ-algebra $\mathcal{F}_t$ is complete with respect to P.
(2) A filtration is called *right continuous* if $\mathcal{F}_t = \mathcal{F}_{t+} := \cap_{\epsilon>0}\mathcal{F}_{t+\epsilon}$ for all $t \in T$, where $\mathcal{F}_{t+\epsilon} := \cap_{u\in T, u\geq t+\epsilon}\mathcal{F}_u$ for $\epsilon > 0$.

Apparently every filtration $(\mathcal{F}_t)_{t\in T}$ has a smallest right continuous and complete filtration $(\mathcal{G}_t)_{t\in T}$ extending in the sense that $\mathcal{F}_t \subset \mathcal{G}_t$ for $t \in T$, namely

$$\mathcal{G}_t = \overline{\mathcal{F}_{t+}} = \overline{\mathcal{F}}_{t+}\,.$$

It is called the *augmented filtration*.

Notice that usual conditions are necessary if one expects P-nullsets to be part of $\mathcal{F}_0$ to guarantee the process to be adapted, which appears in constructions like regularization procedures. For most purposes right continuous filtrations are enough to obtain regularized processes (e.g. with cadlag trajectories) outside a P-nullset.

Definition 2.1. *A $T \cup \{\infty\}$-valued random variable τ is called an $(\mathcal{F}_t)_{t\in T}$-stopping time or $(\mathcal{F}_t)_{t\in T}$-optional time if $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \in T$. A $T \cup \{\infty\}$-valued random variable τ is called a weakly $(\mathcal{F}_t)_{t\in T}$-optional time if $\{T < t\} \in \mathcal{F}_t$ for all $t \in T$.*

We can collect several results:

(1) Let $(\mathcal{F}_t)_{t\in T}$ be a right continuous filtration, then every weakly optional time is optional.
(2) Suprema of sequences of optional times are optional, infima of sequences of weakly optional times are weakly optional.
(3) Any weakly optional time τ taking values in $\mathbb{R}\cup\{\pm\infty\}$ can be approximated by some countably valued optional time $\tau_n \searrow \tau$, take for instance $\tau_n := 2^{-n}[2^n\tau + 1]$.

A stochastic process is a family of random variables $(X_t)_{t\in T}$ on a filtered probability space $(\Omega, \mathcal{F}, P)$. The process is said to be $(\mathcal{F}_t)_{t\in T}$-adapted if X_t is $\mathcal{F}_t$-measurable for all $t \in T$. Two stochastic processes $(X_t)_{t\in T}$ and $(Y_t)_{t\in T}$ are called *versions* of each other if $X_t = Y_t$ almost surely for all $t \in T$, i.e. for every $t \in T$ there is a null set N_t such that $X_t(\omega) = Y_t(\omega)$ for $\omega \notin N_t$. Two stochastic processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are called *indistinguishable* if almost everywhere $X_t = Y_t$ for all $t \in T$, i.e. there is a set N such that for all $t \in T$ equality $X_t(\omega) = Y_t(\omega)$ holds true for $\omega \notin N$. A stochastic process is called cadlag or RCLL (caglad or LCRL) if the sample paths $t \mapsto X_t(\omega)$ are right continuous and have limits at the left side (left continuous and have limits at the right hand side). Given a stochastic process $(X_t)_{t\in T}$ we denote by $\hat{X}$ the associated mapping from $T \times \Omega$ to $\mathbb{R}$, which maps (t, ω) to $X_t(\omega)$.

Given a general filtration: a stochastic process $(X_t)_{t\in T}$ is called progressively measureable or progressive if $\hat{X} : ((T \cap [-\infty, t]) \times \Omega, \mathcal{B}(T \cap [-\infty, t]) \otimes \mathcal{F}_t) \to \mathbb{R}$ is measureable for all t. An adapted stochastic process with right continuous paths (or left continuous paths) almost surely is progressively measureable. A stochastic process $(X_t)_{t\geq 0}$ is called measureable if $\hat{X}$ is measureable with respect to the product σ algebra. A measureable adapted process has a progressively measureable modification, which is a complicated result.

Given a general filtration we can consider an associated filtration on the convex hull $\overline{T}$ of T, where right from right-discrete points (i.e. $t \in T$ such that there is $\delta > 0$ with $]t, t + \delta[\cap T = \emptyset$), a right continuous extension is performed. We can introduce two σ-algebras on $\overline{T} \times \Omega$, namely the one generated by left continuous processes (the predictable σ-algebra) and the one generated by the right continuous processes (the optional σ-algebra). Given a stochastic processes $(X_t)_{t\in T}$ then we call it predictable if there is a predictable extension on $\overline{T}$.

Most important stopping times are hitting times of stochastic processes $(X_t)_{t\in T}$: given a Borel set B we can define the hitting time τ of a stochastic process $(X_t)_{t\in T}$ via

$$\tau = \inf\{t \in T \mid X_t \in B\}\,.$$

There are several profound results around hitting times useful in potential theory, we provide some simple ones:

(1) Let $(X_t)_{t\in T}$ be an $(\mathcal{F}_t)_{t\in T}$-adapted right continuous process with respect to a general filtration and let B be an open set, then the hitting time is a weak $(\mathcal{F}_t)_{t\in T}$-stopping time.
(2) Let $(X_t)_{t\in T}$ be an $(\mathcal{F}_t)_{t\in T}$-adapted continuous process with respect to a general filtration and let B be a closed set, then the hitting time is a $(\mathcal{F}_t)_{t\in T}$-stopping time.
(3) It is a very deep result in potential theory that all hitting times of Borel subsets are $(\mathcal{F}_t)_{t\in T}$-stopping times, if usual conditions are satisfied and the process is progressively measurable (Debut theorem).

The σ-algebra $\mathcal{F}_t$ represents the set of all (theoretically) observable events up to time t including t, the stopping time σ-algebra $\mathcal{F}_\tau$ represents all (theoretically) observable events up to τ:

$$\mathcal{F}_\tau = \{A \in \mathcal{F}, A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \in T\} .$$

Definition 2.2. *A stochastic process $(X_t)_{t\in T}$ is called martingale (submartingale, supermartingale) with respect to a filtration $(\mathcal{F}_t)_{t\in T}$ if*

$$E[X_t \mid \mathcal{F}_s] = X_s$$

for $t \geq s$ in T ($E[X_t \mid \mathcal{F}_s] \geq X_s$ for submartingales, $E[X_t \mid \mathcal{F}_s] \leq X_s$ for supermartingales).

Proposition 2.3. *Let M be a martingale on an arbitrary index set T with respect to a filtration $(\mathcal{G}_t)_{t\in T}$. Assume a second filtration $(\mathcal{F}_t)_{t\in T}$ such that $\mathcal{F}_t \subset \mathcal{G}_t$ for $t \in T$ and assume that M is actually also $\mathcal{F}$ adapted, then M is also a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in T}$.*

Proof. The proof is a simple consequence of the tower property

$$\mathbb{E}\left[M_t \middle| \mathcal{F}_s\right] = \mathbb{E}\left[\mathbb{E}\left[M_t \middle| \mathcal{G}_s\right] \middle| \mathcal{F}_s\right] = \mathbb{E}\left[M_s \middle| \mathcal{F}_s\right] = M_s$$

for $s \leq t$ in T. □

Remember the Doob-Meyer decomposition for any integrable adapted stochastic process $(X_t)_{t\in\mathbb{N}}$ on the *countable index set* $T = \mathbb{N}$: there is a unique martingale M and a unique predictable process A with $A_0 = 0$ such that $X = M - A$. It can be defined directly via

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

for $t \in \mathbb{N}$, and $X = M - A$. The decomposition is unique since a predictable martingale starting at 0 vanishes. If X is a super-martingale, then the Doob-Meyer decomposition $X = M - A$ yields a martingale and an increasing process A.

We shall need in the sequel a curious inequality for bounded, non-negative super-martingales, see [4] (from where we borrow the proof). We state first a lemma for super-martingales of the type $X_s = E[A_\infty | \mathcal{F}_s] - A_s$, for $s \geq 0$, where A is a non-negative increasing process with $A_0 = 0$ and limit at infinity A_∞ (such a super-martingale is called a potential). Assume $0 \leq X \leq c$, then

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

for natural numbers $p \geq 1$. Indeed

$$\begin{aligned} A_\infty^p &= \sum_{i_1,\dots,i_p} (A_{i_1+1} - A_{i_1}) \dots (A_{i_p+1} - A_{i_p}) \\ &= p \sum_j \sum_{j_1,\dots,j_{p-1} \leq j} (A_{j_1+1} - A_{j_1}) \dots (A_{j_{p-1}+1} - A_{j_{p-1}})(A_{j+1} - A_j) \\ &= p \sum_{j_1,\dots,j_{p-1}} (A_{j_1+1} - A_{j_1}) \dots (A_{j_{p-1}+1} - A_{j_{p-1}})(A_\infty - A_{j_1 \vee \dots \vee j_{p-1}}) . \end{aligned}$$

Taking expectations, observing that all terms except the last one are measurable with respect to $\mathcal{F}_{j_1 \vee \dots \vee j_{p-1}}$ and inserting

$$X_{j_1 \vee \dots \vee j_{p-1}} = E[A_\infty - A_{j_1 \vee \dots \vee j_{p-1}} \mid \mathcal{F}_{j_1 \vee \dots \vee j_{p-1}}] \leq c$$

we obtain the recursive inequality

$$E[A_\infty^p] \leq p E[A_\infty^{p-1}] c ,$$

which leads by induction to the desired result, since $E[A_\infty] = E[X_0]$ by the definition of a potential.

Let us now consider a non-negative super-martingale X bounded by a constant $c \geq 0$, then we can apply the previous inequality for the potential $X_0, \dots, X_n, X_\infty, 0, 0, \dots$ (notice here that also the filtration changes)with corresponding $\tilde{A}_\infty^{(n)} = A_n + E[X_n - X_\infty \mid \mathcal{F}_n] + X_\infty \geq A_n + X_\infty$. If we let n tend to ∞ (notice the almost sure convergence of the bounded super-martingale, see below), we obtain the result, namely that

$$E[M_\infty^p] \leq p! c^{p-1} E[X_0] , \tag{2.1}$$

since $M_\infty = A_\infty + X_\infty = \lim_{n \to \infty} \tilde{A}_\infty^{(n)}$.

3. Optional Sampling: the discrete case

The most important theorem of this section is Doob's optional sampling theorem: it states that the stochastic integral, often also called martingale transform, with respect to a martingale is again a martingale. Most of the proofs stem from Olav Kallenberg's book [3].

Let us consider a finite index set T, a stochastic process $X = (X_t)_{t \in T}$ and an increasing sequence of stopping times $(\tau_k)_{k \geq 0}$ taking values in T together with bounded random variables V_k, which are $\mathcal{F}_{\tau_k}$, then

$$V_t := \sum_{k \geq 0} V_k 1_{\{\tau_k < t \leq \tau_{k+1}\}}$$

is a predictable process. We call such processes *simple predictable* and we can define the stochastic integral (as a finite sum)

$$(V \bullet X)_{s,t} := \sum_{k \geq 0} V_k (X_{s \vee t \wedge \tau_{k+1}} - X_{s \vee t \wedge \tau_k}) .$$

By simple conditioning arguments we can prove the following proposition:

Proposition 3.1. *Assume that the stopping times are deterministic and let X be a martingale, then $(V \bullet X)$ is a martingale. If X is a sub-martingale and $V \geq 0$, then $(V \bullet X)$ is a sub-martingale, too. Furthermore it always holds that $(V \bullet (W \bullet X)) = (VW \bullet X)$ (Associativity).*

This basic proposition can be immediately generalized by considering stopping times instead of deterministic times, since for any two stopping times $\tau_1 \leq \tau_2$ taking values in T we have

$$1_{\{\tau_1 < t \leq \tau_2\}} = \sum_{k \geq 0} 1_{\{\tau_1 < t \leq \tau_2\}} 1_{\{t_k < t \leq t_{k+1}\}}$$

for the sequence $t_0 < t_1 < t_2 < \ldots$ exhausting the finite set T. Whence we can argue by Associativity $(V \bullet (W \bullet X)) = (VW \bullet X)$ that we do always preserve the martingale property, or the sub-martingale property, respectively, in case of non-negative integrands.

We obtain several simple conclusions from these basic facts (optional sampling theorem):

(1) A stochastic process X is a martingale if and only if for each pair σ, τ of stopping times taking values in T we have that $E[M_\tau] = E[M_\sigma]$.

(2) For any stopping time τ taking values in $T \cup \{\infty\}$ and any process M the stopped process M^τ

$$M_t^\tau := M_{\tau \wedge t}$$

for $t \in T$ is well defined. If M is a (sub)martingale, M^τ is a (sub)martingale, too, with respect to $(\mathcal{F}_t)_{t \in T}$ and with respect to the stopped filtration $(\mathcal{F}_{t \wedge \tau})_{t \in T}$.

(3) For two stopping times σ, τ taking values in T, and for any martingale M we obtain

$$\mathbb{E}\,[\,M_\tau |\, \mathcal{F}_\sigma] = M_{\sigma \wedge \tau}\,.$$

For this property we just use that $(M_t - M_{t \wedge \tau})_{t \in T}$ is a martingale with respect to $(\mathcal{F}_t)_{t \in T}$ as difference of two martingales. Stopping this martingale with σ leads to a martingale with respect to $(\mathcal{F}_{t \wedge \sigma})_{t \in T}$, whose evaluation at $\sup T$ and conditional expectation on $\mathcal{F}_\sigma$ leads to the desired property.

These basic considerations are already sufficient to derive fundamental inequalities, from which the important set of maximal inequalities follows.

Theorem 3.2. *Let X be a submartingale on a finite index set T with maximum $supT \in T$, then for any $r \geq 0$ we have*

$$rP[\sup_{t \in T} X_t \geq r] \leq E[X_{\sup T} 1_{\{\sup_{t \in T} X_t \geq r\}}] \leq E[X^+_{\sup T}]$$

and

$$rP[\sup_{t \in T} |X_t| \geq r] \leq 3 \sup_{t \in T} E[|X_t|]\,.$$

Proof. Consider the stopping time $\tau = \inf\{s \in T \mid X_s \geq r\}$ (which can take the value ∞) and the predictable strategy

$$V_t := 1_{\{\tau \leq \sup T\}} 1_{\{\tau < t \leq \sup T\}}\,,$$

then $E[(V \bullet X)] \geq 0$ by Proposition 3.1, hence the first assertion follows. For the second one take the sub-martingale $|X|$. □

Theorem 3.3. *Let M be a martingale on a finite index set T with maximum* $\sup T$

$$E[\sup_{t \in T} |M_t|^p] \leq \big(\frac{p}{p-1}\big)^p E[|M_{\sup T}|^p]$$

for all $p > 1$.

Proof. We apply the Levy-Bernstein inequalities from Theorem 3.2 to the submartingale $|M|$. This yields by Fubini's theorem and the Hölder inequality

$$\begin{aligned} E[\sup_{t\in T} |M_t|^p] &= p\int_0^\infty P[\sup_{t\in T} |M_t| > r] r^{p-1} dr \\ &\leq p\int_0^\infty E[|M_{\sup T}| 1_{\{\sup_{t\in T}|M_t|\geq r\}}] r^{p-2} dr \\ &= pE\big[|M_{\sup T}| \int_0^{\sup_{t\in T}|M_t|} r^{p-2} dr\big] \\ &= \big(\frac{p}{p-1}\big) E\big[|M_{\sup T}| \sup_{t\in T} |M_t|^{p-1}\big] \leq \big(\frac{p}{p-1}\big) \|M_{\sup T}\|_p \|\sup_{t\in T} |M_t|^{p-1}\|_q , \end{aligned}$$

which yields the result. □

Finally we prove Doob's upcrossing inequality, which is the heart of all further convergence theorems, upwards or downwards. Consider an interval $[a,b]$ and consider a submartingale X, then we denote by $N([a,b],X)$ the number of upcrossings of a trajectory of X from below a to above b. We have the following fundamental lemma:

Lemma 3.4. *Let X be s submartingale, then*

$$E[N([a,b],X)] \leq \frac{E[(X_{\sup T} - a)]^+}{b-a}$$

Proof. Denote $\tau_0 = \min T$, then define recursively the stopping times

$$\sigma_k := \inf\{t \geq \tau_{k-1} \mid X_t \leq a\}$$

and

$$\tau_k := \inf\{t \geq \sigma_k \mid X_t \geq b\}$$

for $k \geq 1$. The process

$$V_t := \sum_{k\geq 1} 1_{\{\sigma_k < t \leq \tau_k\}}$$

is predictable and $V \geq 0$ such as $1 - V$. We conclude by

$$\begin{aligned} (b-a)E[N([a,b],X)] &\leq E[(V \bullet (X-a)^+)_{\sup T}] \\ &\leq E[(1 \bullet (X-a)^+)_{\sup T}] \leq E[(X_{\sup T} - a)^+] . \end{aligned}$$

Mind the slight difference of these inequalities to the one in the introduction. □

This remarkable lemma allows to prove the following deep convergence results by passing to countable index sets:

Theorem 3.5. *Let X be an L^1 bounded submartingale on a countable index set T, then there is a set A with probability one such that X_t converges along* any *increasing or decreasing sequence in T.*

Proof. By $\sup_{t\in T} E[|X_t|] < \infty$ we conclude by the Lévy-Bernstein inequality that the measurable random variable $\sup_{t\in T} |X_t|$ is finitely valued, hence along every subsequence there is a finite inferior or superior limit. By monotone convergence we know that for any interval the number of upcrossings is finite almost surely. Consider now A, the intersection of sets with probability one, where the number of upcrossings is finite over intervals with rational endpoints. A has again probability one and on A the process X converges along any increasing or decreasing subsequence, since along monotone sequences a finite number of upcrossings leads to equal inferior and superior limits. Notice that we work here with monotone convergence, since the number of upcrossings for increasing index sets is increasing, however, its expectation is bounded. □

Theorem 3.6. *For any martingale M on any index set we have the following equivalence:*

(1) M is uniformly integrable.
(2) M is closeable at $\sup T$.
(3) M is L^1 convergent at $\sup T$.

Proof. If M is closeable on an arbitrary index set T, then by definition there is $\xi \in L^1(\Omega)$ such that $M_t = E[\xi \mid \mathcal{F}_t]$ for $t \in T$, hence

$$E[M_t 1_A] \leq E[E[|\xi| \mid \mathcal{F}_t] 1_A] = E[|\xi| E[1_A \mid \mathcal{F}_t]]$$

for any $A \in \mathcal{F}$, which tends to zero if $P(A) \to 0$, uniformly in t, hence uniform integrability. On the other hand a uniformly integrable martingale is bounded in L^1 and therefore we have one and the same almost sure limit along any subsequence increasing to $\sup T$. If M is uniformly integrable, an almost sure limit is in fact L^1.

Finally assume $M_t \to \xi$ for $t \to \sup T$ in L^1, hence $M_s \to E[\xi \mid \mathcal{F}_s]$, for $s \in T$ and the martingale property, hence $M_s = E[\xi \mid \mathcal{F}_s]$ for any $s \in T$, which concludes the proof. □

We obtain the following beautiful corollary:

Corollary 3.7. *Let M be a martingale on an arbitrary index set and assume $p > 1$, then M_t converges in L^p for $t \to \sup T$ if and only if it is L^p bounded.*

Proof. If M is L^p bounded, then it is uniformly integrable (by Doob's maximal inequalities from Theorem 3.3) and convergence takes place in L^1 by the previous theorem, which in turn by L^p-boundedness is also a convergence in L^p. On the other hand, if M converges in L^p, then it is by the Jensen's inequality also L^p bounded. □

Finally we may conclude the following two sided version of closedness:

Theorem 3.8. *Let T be a countable index set unbounded above and below, then for any $\xi \in L^1$ we have that*

$$E[\xi \mid \mathcal{F}_t] \to E[\xi \mid \mathcal{F}_{\pm\infty}]$$

for $t \to \pm\infty$.

Proof. By L^1 boundedness we obtain convergence along any increasing or decreasing subsequence towards limits $M_{\pm\infty}$. The upwards version follows from the previous Theorem 3.6, the downwards version follows immediately. □

This last theorem is often called reserve martingale convergence and can be used to prove the strong law of large numbers and almost sure convergence of quadratic variations:

(1) Let $(\xi_i)_{i \geq 1}$ be an i.i.d. sequence of random variables in L^1. Then we can consider the filtration

$$\mathcal{F}_{-n} := \sigma(S_n, S_{n+1}, \ldots)$$

with $S_n := \sum_{i=1}^n \xi_i$, for $n \geq 1$. Since $(\xi_1, S_n, S_{n+1}, \dots) = (\xi_k, S_n, S_{n+1}, \dots)$ for $k \leq n$ in distribution, we obtain

$$E[\xi_1 \mid \mathcal{F}_{-n}] = E[\xi_k \mid \mathcal{F}_{-n}]$$

for $k \leq n$. Whence

$$\begin{aligned} \frac{S_n}{n} &= \frac{1}{n} E[S_n \mid \mathcal{F}_{-n}] = \frac{1}{n} \sum_{i=1}^n E[\xi_i \mid \mathcal{F}_{-n}] \\ &= \frac{1}{n} \sum_{i=1}^n E[\xi_1 \mid \mathcal{F}_{-n}] = E[\xi_1 \mid \mathcal{F}_{-n}] \to E[\xi_1] \,, \end{aligned}$$

since the intersection of all tail σ algebras is trivial, i.e. all elements of the intersection have probability either 0 or 1 (Hewitt-Savage $0-1$ law).

(2) Let $(W_t)_{t \in [0,1]}$ be a Wiener process and denote – for a fixed time $0 \leq t \leq 1$ – by

$$V_t^n := \sum_{t_i \in \Pi^n} (W_{t_{i+1}} - W_{t_i})^2$$

the approximations of quadratic variation t along a refining sequence of partitions $\Pi^n \subset \Pi^{n+1}$ of $[0,t]$, whose meshes tend to zero. We know from the lecture notes that the approximation takes place in L^2, but we do not know whether it actually holds almost surely. Consider the filtration

$$\mathcal{F}_{-n} := \sigma(V_t^n, V_t^{n+1}, \dots)$$

for $n \geq 1$, whose intersection is actually trivial. Without loss of generality we assume that each partition Π^n contains n partition points by possibly adding to the sequence intermediate partitions. Fix now $n \geq 2$, then we consider the difference between Π^{n-1} and Π^n, which is just one point v lying only in Π^n and being surrounded by two nearest neighboring points from Π^{n-1}, i.e. $u < v < w$. Consider now a second Wiener process $\tilde{W}_s = W_{s \wedge v} - (W_s - W_{s \wedge v})$. Apparently

$$(\tilde{V}_t^{n-1}, V_t^n, V_t^{n+1}, \dots) = (V_t^{n-1}, V_t^n, V_t^{n+1}, \dots)$$

in distribution, hence it holds that

$$E[\tilde{V}_t^{n-1} - V_t^n \mid \mathcal{F}_{-n}] = E[V_t^{n-1} - V_t^n \mid \mathcal{F}_{-n}] \,,$$

which in turn means that

$$E[(\tilde{W}_u - \tilde{W}_v)(\tilde{W}_v - \tilde{W}_w) \mid \mathcal{F}_{-n}] = E[(W_u - W_v)(W_v - W_w) \mid \mathcal{F}_{-n}] \,.$$

Inserting the definition of $\tilde{W}$ yields the result that

$$E[V_t^{n-1} - V_t^n \mid \mathcal{F}_{-n}] = 0$$

for $n \geq 1$. Hence we have a martingale on the index set $\mathbb{Z}_{\leq 1}$, which by martingale convergence tends almost surely to its L^2 limit t.

4. Martingales on continuous index sets

Martingale inequalities on uncountable index sets can often be derived from inequalities for the case of countable index sets if certain path properties are guaranteed. From martingale convergence results on countable index sets we can conclude the existence of RCLL versions for processes like martingales, which is the main result of this section. Most of the proofs stem from Olav Kallenberg's book [3].

We need an auxiliary lemma on reverse submartingales first. Of course similar statements hold for supermartingales.

Lemma 4.1. *Let X be a submartingale on $\mathbb{Z}_{\leq 0}$. Then X is uniformly integrable if and only if $E[X]$ is a bounded (from below) sequence.*

Proof. Let $E[X]$ be bounded from below. We can then introduce a Doob-Meyer type decomposition, i.e.

$$A_n := \sum_{k<n} E[X_{k+1} - X_k \mid \mathcal{F}_k]\,,$$

which is well defined since all summands are positive due to submartingality and

$$E[A_0] \leq E[X_0] - \inf_{n \geq 0} E[X_n] < \infty\,.$$

Whence $X = M + A$, where M is a martingale. Since A is uniformly integrable and M is a martingale being closed at 0 by martingale convergence, hence uniformly integrable, also the sum is uniformly integrable. The other direction follows immediately since $E[X_n]$ is decreasing for $n \to \infty$. If it were unbounded from below, it cannot be uniformly integrable. □

From this statement we can conclude by martingale convergence the following fundamental regularization result:

Theorem 4.2. *For any submartingale X on $\mathbb{R}_{\geq 0}$ with restriction Y to $\mathbb{Q}_{\geq 0}$ we have:*

1. *The process of right hand limits Y^+ exists on $\mathbb{R}_{\geq 0}$ outside some nullset A and $Z := 1_{A^c} Y^+$ is an RCLL submartingale with respect to the augmented filtration $\overline{\mathcal{F}_+}$.*
2. *If the filtration is right continuous, then X has a RCLL version, if and only if $t \mapsto E[X_t]$ is right continuous.*

Proof. The process Y is L^1 bounded on bounded intervals since the positive part is an L^1 bounded submartingale by Jensen's inequality, hence by martingale convergence Theorem 3.5 we obtain the existence of right and left hand limits and therefore Y^+ is RCLL. Clearly the process Z is adapted to the augmented filtration $\overline{\mathcal{F}_+}$.

The submartingale property follows readily, too: fix times $s < t$ and choose $s_n \searrow s$ and $t_m \searrow t$, with $s_n < t$ for all $n \geq 1$. Then – by assumption – $E[Y_{t_m} \mid \mathcal{F}_{s_n}] \geq Y_{s_n}$. By martingale convergence to the left we obtain

$$E[Y_{t_m} \mid \mathcal{F}_{s+}] \geq Z_s$$

almost surely. Since the submartingale $(Y_{t_m}))_{m \geq 1}$ has bounded expectations, we conclude L^1-convergence (due to uniform integrability by the previous lemma) and therefore arrive at

$$E[Z_t \mid \overline{\mathcal{F}_{s+}}] \geq Z_s\,.$$

For the second assertion observe that if X is RCLL, then the curve $E[X]$ is right continuous by uniform integrability along decreasing subsequences and the previous lemma. On the other hand if $E[X]$ is right continuous $Z_t = E[Z_t \mid \mathcal{F}_t] \geq X_t$ by limits from the right, but $E[Z_t - X_t] = 0$ by right continuity of $E[X]$, hence Z and X are indistinguishable. $\square$

5. Stochastic Integration for caglad integrands

We follow here mainly Philip Protter's book [5] on stochastic integration, which is inspired by works of Klaus Bichteler, Claude Dellacherie and Paul-Andre Meyer, see [2] and [4]: the idea is to crystallize the essential property (the "good integrator" property), which guarantees the existence of stochastic integrals, and to derive all properties of stochastic integrals from the good integrator property. Finally it can be shown that every good integrator is the sum of a local martingale and a finite variation process. This approach leads to an integration theory for caglad integrands.

Let us introduce some notation: we denote by $\mathbb{S}$ the set of simple predictable processes, i.e.

$$H_0 1_{\{0\}} + \sum_{i=1}^{n} \sum_{i} H_i 1_{]T_i, T_{i+1}]}$$

for an increasing, finite sequence of stopping times $0 = T_0 \leq T_1 \leq \dots T_{n+1} < \infty$ and H_i being $\mathcal{F}_{T_i}$ measurable, by $\mathbb{L}$ the set of caglad processes and by $\mathbb{D}$ the set of cadlag processes on $\mathbb{R}_{\geq 0}$. These vector spaces are endowed with the metric

$$d(X,Y) := \sum_{n \geq 0} \frac{1}{2^n} E\big[|(X-Y)|_n^* \wedge 1\big],$$

which makes $\mathbb{L}$ and $\mathbb{D}$ complete topological vector spaces. We call this topology the ucp-topology ("uniform convergence on compacts in probability"). Notice that predictable strategies as well as integrators are considered $\mathbb{R}$ valued here, which, however, *contains* the $\mathbb{R}^n$ case.

Notice that we are dealing here with topological vector spaces, which are in general not even locally convex. This leads also to the phenomenon that the metric does not detect boundedness of sets, which is defined in the following way: A subset B of a topological vector space $\mathbb{D}$ is called bounded, if it can be absorbed by any open neighborhood U of zero, i.e. there is $R > 0$ such that $B \subset RU$. For instance for the space of random variables $L^0(\Omega)$ this translates to the following equivalent statement: a set B of random variables is bounded in probability if for every $\epsilon > 0$ there is $c > 0$ such that

$$P[|Y| \geq c] < \epsilon$$

for $Y \in B$, which is of course not detectable by the metric.

In order to facilitate the reasoning we shall use the following definition, which seems slightly less general than the original ones which is treated in the subsequent remark.

Definition 5.1. *A cadlag process X is called good integrator if the map*

$$J_X : \mathbb{S} \to \mathbb{D}$$

with $J_X(H) := H_0 X_0 + \sum_{i=1}^n H_i (X_{T_{i+1} \wedge t} - X_{T_i \wedge t})$, for $H \in \mathbb{S}$, is continuous with respect to the ucp-topologies.

Remark 5.2. It would already be sufficient to require a good integrator to satisfy the following property: for every stopped process X^t the map $I_{X^t} : \mathbb{S}_u \to L^0(\Omega)$, where $I_{X^t}(H) := J_X(H)_\infty$, from uniformly bounded, simple predictable processes *with the uniform topology* $\mathbb{S}_u$ to random variables with convergence in probability, is continuous, i.e.

$$I_{X^t}(H^k) := H_0^k X_0 + \sum_{i=1}^{n} H_i^k (X_{T_{i+1}\wedge t} - X_{T_i \wedge t}) \to 0 \tag{5.1}$$

if $H^k \to 0$ uniformly on $\Omega \times \mathbb{R}_{\geq 0}$.

From continuity with respect to the uniform topology continuity with respect to the ucp topologies, as claimed in the definition of a good integrator, immediately follows. Indeed assume that I_X is continuous with respect to the uniform topology, fix $n \geq 0$ and a sequence $H^k \to 0$, which tends to 0 uniformly, then choose $c \geq 0$ and define a sequence of stopping times

$$\tau^k := \inf\{t \mid |(H^k \bullet X)_t| \geq c\}$$

for $k \geq 0$, then

$$P[|(H^k \bullet X)|_n^* \geq c] = P[|(H^k 1_{[0,\tau^k]} \bullet X)_n| \geq c] \to 0$$

as $k \to \infty$ by assumption, hence $J_X : \mathbb{S}_u \to \mathbb{D}$ is continuous, i.e. we can map to processes without loosing continuity.

Take now a sequence $H^k \to 0$ in ucp, and choose $c \geq 0$, $\epsilon > 0$ and $n \geq 0$. Then there is some $\eta > 0$ such that

$$P[|(H \bullet X)_n|^* \geq c] \leq \epsilon$$

for $\|H\|_\infty \leq \eta$ by continuity of $J_X : \mathbb{S}_u \to \mathbb{D}$. Define furthermore stopping times

$$\rho^k := \inf\{s \mid |H_s^k| > \eta\}$$

then we obtain

$$P[|(H^k \bullet X)_n|^* \geq c] \leq P[|(H^k 1_{[0,\rho^k]} \bullet X)_n| \geq c] + P[\rho^k < n] < 2\epsilon$$

if k is large enough since $P[\rho^k < n] \to 0$ as $k \to \infty$.

Clearly Property (5.1) holds if it only holds locally: indeed let τ^n be a localizing sequence, i.e. $\tau^n \nearrow \infty$ with X^{τ^n} being a good integrator. Fix $t \geq 0$ and a sequence $H^k \to 0$, which tends to 0 uniformly, then

$$P[|I_{X^t}(H^k)| \geq c] \leq P[|I_{X^{\tau^n \wedge t}}(H^k)| \geq c] + P[\tau^n \leq t]$$

for every $n \geq 0$. Hence we can choose n large enough, such that the second term is small by the localizing property, and obtain for k large enough that the first term is small by X^{τ^n} being a good integrator.

Remark 5.3. The set $\mathbb{S}$ is dense in the ucp-topology in $\mathbb{L}$, even the bounded simple predictable processes are dense. Consider just the sequence of partitions introduced at the beginning of the next subsection.

Remark 5.4. We provide examples of good integrators:

- Any process of finite variation A with is a good integrator, since for every simple (not even predictable) process it holds that

$$|\int_0^t H_s dA_s| \leq \|H\|_\infty \big(\int_0^t d|A|_s + |A_0| \big)$$

almost surely, for $t \geq 0$.

- By Ito's fundamental insight square integrable martingales M are good integrators, since
$$E\big[(H \bullet M)_t{}^2\big] \leq \|H\|_\infty^2 \big(E[|M_t|^2]\big)$$
holds true for simple, bounded and predictable processes $H \in \mathbb{S}_u$.
- By the following elementary inequality due to Burkholder we can conclude that martingales are good integrators: for every martingale M and every simple, bounded and predictable process $H \in \mathbb{S}_u$ it holds that
$$cP(|(H \bullet M)|_1^* \geq c) \leq 18 \|H\|_\infty \|M_1\|_1$$
for all $c \geq 0$. For an easy proof of this inequality see, e.g., works of Klaus Bichteler [2] or Paul-Andre Meyer [4, Theorem 47, p. 50]. Since the inequality is crucial for our treatment, we shall prove it here, too. Notice that we are just dealing with integrals with respect to simple integrands, hence we can prove it for discrete martingales on a finite set of time points. Let M be a non-negative martingale first and H bounded predictable with $\|H\|_\infty \leq 1$, then $Z := M \wedge c$ is a supermartingale and we have
$$cP(|(H \bullet M)|_1^* \geq c) \leq cP(|M|_1^* \geq c) + cP(|(H \bullet Z)|_1^* \geq c)\,.$$
Since Z is a super-martingale we obtain by the Doob-Meyer decomposition for discrete super-martingales $Z = \tilde{M} - A$ that
$$|(H \bullet Z)| \leq |(H \bullet \tilde{M})| + A\,,$$
i.e. we have an upper bound being a sub-martingale. With $|(H \bullet \tilde{M})| + A$ also its square is a sub-martingale. Hence we can conclude by Lemma 3.2 that
$$cP(|(H \bullet M)|_1^* \geq c) \leq E[M_1] + 2\frac{1}{c}E[(H \bullet \tilde{M})_1^2 + A_1^2]\,,$$
since
$$cP(|(H \bullet Z)|_1^* \geq c) \leq cP\Big(\big|\,|(H \bullet \tilde{M})| + A\big|_1^* \geq c\Big) \leq$$
$$\leq \frac{1}{c}E\Big[\big(|(H \bullet \tilde{M})|_1 + A_1\big)^2\Big] \leq 2\frac{1}{c}E[(H \bullet \tilde{M})_1^2 + A_1^2]\,.$$
Ito's insight allows to estimate the variance of the stochastic integral at time 1 by $E[\tilde{M}_1^2]$. Both quantities $\tilde{M}$ and A of the Doob-Meyer decomposition may, however, be estimated through $E[A_1^2] \leq E[\tilde{M}_1^2] \leq 2cE[Z_0] \leq 2cE[M_0]$, see (2.1), since Z is non-negative (so $A \leq \tilde{M}$ holds true) and $Z \leq c$. This leads to an upper bound
$$cP(|(H \bullet M)|_1^* \geq c) \leq 9E[M_0]\,.$$
Writing a martingale as difference of two non-negative martingales leads to the desired result. Apparently the result translates directly to the fact that M is a good integrator. We actually immediately obtain that $J_X : \mathbb{S}_u \to \mathbb{D}$ is continuous, wherefrom – as we have seen before – the continuity even with respect to the ucp topology on $\mathbb{S}$ follows.

By density and continuity we can extend the map J_X to all caglad processes $Y \in \mathbb{L}$, which defines the stochastic integral $(Y \bullet X)$. As a simple corollary we can prove the following proposition:

Proposition 5.5. *Let $H, G \in \mathbb{L}$ be given and let X be a good integrator, then $(G \bullet X)$ is a good integrator and $(H \bullet (G \bullet X)) = (HG \bullet X)$.*

Proof. Let X be a good integrator, then $J_X : \mathbb{L} \to \mathbb{D}$ is continuous with respect to the ucp topologies. Let (H_k) be a sequence in $\mathbb{S}_u$ converging uniformly to 0, then $H_k G$ converges ucp for every $G \in \mathbb{L}$, whence $J_{(G \bullet X)}$ which satisfies

$$J_{(G \bullet X)}(H) = (HG \bullet X)$$

is obviously continuous, and the desired formula holds by continuous extension. □

5.1. **Approximation results.** We know that $H \mapsto (H \bullet X)$ is continuous with respect to the ucp topologies on the left and right hand side. However, actually a bit more is true.

Most important for the calculation and understanding of stochastic integrals is the following approximation result: a *sequence of partition tending to identity* Π^k consists of stopping times $0 = T_0^k \leq \ldots \leq T_{i_k}^k < \infty$ with mesh $\sup_i (T_{i+1}^k - T_i^k)$ tending to 0 and $\sup_i T_i^k \to \infty$. We call the sequence of cadlag processes

$$Y^{\Pi^k} := \sum_i Y_{T_i^k} 1_{[T_i^k, T_{i+1}^k[}$$

a sampling sequence for a cadlag process Y along Π^k, for $k \geq 0$. Notice that we do not necessarily have that $Y^{\Pi^k} \to Y$ in ucp, nor $Y_-^{\Pi^k} \to Y_-$ due to the presence of large jumps.

Example 5.6. One important sequence of partition is constructed by a truncation of the following one: let $Y \in \mathbb{D}$ be a cadlag process. For $n \geq 0$ we can define a double sequence of stopping times τ_i^n

$$\tau_0^n := 0 \text{ and } \tau_{i+1}^n := \inf\{s \geq \tau_i^n \mid |Y_s - Y_{\tau_i^n}| \geq \frac{1}{2^n}\}$$

for $i \geq 0$. This defines a sequence of partitions

$$\Pi^n = \{\tau_0^n \leq \ldots \tau_{n2^n}^n\}$$

tending to identity. We have that

$$|Y_- - Y_-^{\Pi^n}| \leq \frac{1}{2^n} \,,$$

hence $Y_-^{\Pi^n} \to Y_-$ in the ucp topology, and also $Y^{\Pi^n} \to Y$ in ucp.

Theorem 5.7. *For any good integrator X we obtain that*

$$(Y_-^{\Pi^k} \bullet X) \to (Y_- \bullet X)$$

in the ucp topology in general (even though Y^{Π^k} does not necessarily converge to Y in ucp), as well as the less usually stated but equally true ucp convergence result

$$(Y_-^{\Pi^k} \bullet X^{\Pi^k}) \to (Y_- \bullet X)_- \,.$$

Remark 5.8. Notice that so far we have no understanding on the continuity of $X \mapsto (H \bullet X)$ on the set of good integrators. Of course it is not ucp continuous. The second assertion gives an answer in a very specific case.

Proof. We know by previous remarks that there are sequences Y^l of simple cadlag processes converging ucp to Y, where $Y^l_- \to Y_-$ holds true for the associated left continuous processes. Hence we can write

$$((Y_- - Y_-^{\Pi^k}) \bullet X) = ((Y_- - Y_-^l) \bullet X) + ((Y_-^l - (Y^l)_-^{\Pi^k}) \bullet X) + (((Y^l)_-^{\Pi^k} - Y_-^{\Pi^k}) \bullet X)\,,$$

where the first and third term converge of course in ucp as $l \to \infty$, the third even uniformly in k. The middle term is seen to converge by direct inspection.

The proof of the second assertion follows from the fact that

$$(Y_-^{\Pi^k} \bullet X^{\Pi^k}) = (Y_- \bullet X^{\Pi^k}) \to (Y_- \bullet X)_-\,,$$

where the limit assertion follows from the fact that $(Y_- \bullet X^{\Pi^k})$ only differs from $(Y_-^{\Pi^k} \bullet X)$ on the 'last' interval before t is reached by the stopping times in the partition Π^k, which is a quantity converging to 0 plus the last jump, i.e.

$$Y_{T_i}(X_{T^k_{i+1}\wedge t} - X_{T^k_i \wedge t}) 1_{[T_{i_t}, T_{i_t+1}[}(t) - Y_t \Delta X_t \to 0$$

in ucp. i_t is chosen such that $t \in [T_{i_t}, T_{i_t+1}[$. □

Definition 5.9. *Let X, Y be good integrators, then we define the quadratic (co-) variation process by*

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

Quadratic variation $[X,X]$ is an non-decreasing (hence finite variation) process for any good integrator. As a consequence of the previous approximation theorem we obtain of course for two good integrators X, Y that

$$[X^{\Pi^k}, Y^{\Pi^k}]_t = \sum_{T^k_{i+1} < t} (X_{T^k_{i+1}} - X_{T^k_i})(Y_{T^k_{i+1}} - Y_{T^k_i}) \to [X,Y]_{t-}\,,$$

in ucp, since quadratic co-variation can be expressed by stochastic integrals as given in the definition. Whence

$$\sum_i (X_{T^k_{i+1}\wedge t} - X_{T^k_i \wedge t})(Y_{T^k_{i+1}\wedge t} - Y_{T^k_i \wedge t}) \to [X,Y]_t$$

again in ucp.

Let us fix an important notations here: we shall always assume that $X_{0-} = 0$ (a left limit coming from negative times), whereas X_0 can be different from zero, whence ΔX_0 is not necessarily vanishing.

Proposition 5.10. *Let $H \in \mathbb{L}$ be fixed, as well as two good integrators X, Y. Then*

$$[(H \bullet X), Y] = (H \bullet [X,Y])$$

Proof. Let $H \in \mathbb{S}$ be fixed. Then apparently

$$[(H \bullet X), Y] = \big(H \bullet [X,Y]\big)\,,$$

whence by continuity with respect to ucp topologies also the left hand side is continuous with respect to ucp topologies, which proves the result by continuity of $H \mapsto (H \bullet X)$ with respect to ucp. □

5.2. **Ito's theorem.** The set of semi-martingales is a vector space, in fact even an algebra. More precisely: given finitely many semi-martingales $X^1,\dots,X^n$ then $f(X^1,\dots,X^n)$ is also a semi-martingale for any C^2 function f.

It is remarkable that Ito's theorem can be concluded from its version for piecewise constant processes due to the following continuity lemma, which complements results which have already been established for the approximation of stochastic integrals. We state an additional continuity lemma:

Lemma 5.11. *Let $X^1,\dots,X^n$ be good integrators, Π^k a sequence of partitions tending to the identity and $f:\mathbb{R}^n\to\mathbb{R}$ a C^2 function, then for $t\geq 0$*

$$\sum_{s\leq t}\big\{f(X_s^{\Pi^k})-f(X_{s-}^{\Pi^k})-\sum_{i=1}^n\partial_i f(X_{s-}^{\Pi^k})\Delta X_s^{i,\Pi^k}-\frac{1}{2}\sum_{i,j=1}^n\partial_{ij}^2 f(X_{s-}^{\Pi^k})\Delta X_s^{i,\Pi^k}\Delta X_s^{j,\Pi^k}\big\}$$

$$\to_{k\to\infty}\sum_{s\leq t}\big\{f(X_s)-f(X_{s-})-\sum_{i=1}^n\partial_i f(X_{s-})\Delta X_s^i-\frac{1}{2}\sum_{i,j=1}^n\partial_{ij}^2 f(X_{s-})\Delta X_s^i\Delta X_s^j\big\},$$

where the limit can be even understood in ucp topology.

Remark 5.12. Here we mean with X^{Π^k} the cadlag version of the sampled process as introduced before.

Proof. The proof of this elementary lemma relies on Taylor expansion of f: apparently the finitely many summands of the approximating series are small at s if $(\Delta X_s^{\Pi^k})^{(2+)}$ is small, hence only those jumps remain after the limit, which are at time points where X actually jumps. Let us make this precise: first we know – by the very existence of quadratic variation – that

$$\sum_{s\leq t}(\Delta X^i)^2\leq[X^i,X^i]_t<\infty$$

almost surely. Fix $t\geq 0$ and $\epsilon>0$, then we find for every $\omega\in\Omega$ a finite set A_ω of times up to t, where X jumps in a large way (defined by the condition on B), and a possibly countable set of times B_ω up to t, where X jumps *and* $\sum_{s\in B}\|\Delta X\|^2\leq\epsilon^2$, since every cadlag path has at most countably many jumps up to time t. Furthermore we know that

$$f(y)-f(x)-\sum_{i=1}^n\partial_i f(x)(y-x)^i-\frac{1}{2}\sum_{i,j=1}^n\partial_{ij}^2 f(x)(y-x)^i(y-x)^j=o\big(\|y-x\|\big)\|y-x\|^2$$

as $y\to x$. This means that for $\omega\in\Omega$ we can split the approximating sum into two sums denoted by $\sum_A$, where $]T_i^k(\omega),T_{i+1}^k(\omega)]\cap A_\omega\neq\emptyset$, and $\sum_B$ corresponding to jumps which appear at B and $\sum_C$ over intervals, where no jumps appear in the limit. We then obtain an estimate for the limiting sum $\sum_B$ of the type

$$\sum_B\leq 2\epsilon^2 o\big(\epsilon\big)$$

for k large enough by uniform continuity of continuous functions on compact intervals and jump size at most ϵ. Furthermore we obtain

$$\sum_C\leq\|[X^{T^k},X^{T^k}]\|o\big(\max_i 1_{]T_i^k,T_{i+1}^k]\cap(A\cup B)=\emptyset}\|X_{T_{i+1}^k}-X_{T_i^k}\|\big).$$

The other part $\sum_A$ behaves differently, but is a finite sum, hence it converges to the respective limit (written with A in respective sense). Letting now tend $\epsilon\to 0$

the result follows immediately. The argument is true uniformly along paths in probability. □

We are now able to prove Ito's formula in all generality:

Theorem 5.13. *Let $X^1,\dots,X^n$ be good integrators and $f:\mathbb{R}^n\to\mathbb{R}$ a C^2 function, then for $t\geq 0$*

$$f(X_t)=\sum_{i=1}^n(\partial_i f(X_-)\bullet X^i)_t+\frac{1}{2}\sum_{i,j=1}^n\left(\partial^2_{ij}f(X_-)\bullet[X^i,X^j]\right)_t+$$
$$+\sum_{0\leq s\leq t}\left\{f(X_s)-f(X_{s-})-\sum_{i=1}^n\partial_i f(X_{s-})\Delta X^i_s-\frac{1}{2}\sum_{i,j=1}^n\partial^2_{ij}f(X_{s-})\Delta X^i_s\Delta X^j_s\right\}.$$

Remark 5.14. Notice that $f(X_0)$ is in the second sum since we agreed that $f(X_{0-})=0$.

Proof. Let Π^k be a sequence of partitions tending to the identity, then Ito's formula reads by careful inspection

$$f(X^{\Pi^k}_t)=\sum_{i=1}^n(\partial_i f(X^{\Pi^k}_-)\bullet X^{i,\Pi^k})_t+\frac{1}{2}\sum_{i,j=1}^n\left(\partial^2_{ij}f(X^{\Pi^k}_-)\bullet[X^{i,\Pi^k},X^{j,\Pi^k}]\right)_t+$$
$$+\sum_{0<s\leq t}\left\{f(X^{\Pi^k}_s)-f(X^{\Pi^k}_{s-})-\sum_{i=1}^n\partial_i f(X^{\Pi^k}_{s-})\Delta X^{i,\Pi^k}_s-\frac{1}{2}\sum_{i,j=1}^n\partial^2_{ij}f(X^{\Pi^k}_{s-})\Delta X^{i,\Pi^k}_s\Delta X^{j,\Pi^k}_s\right\},$$

since the process is piece-wise constant and the sum is just telescoping. By the previously stated convergence result, however, this translates directly – even in ucp convergence – to the limit for $k\to\infty$, which is Ito's formula. □

5.3. **Quadratic Pure jump good integrators.** We call a good integrator *quadratic pure jump* if $[X,X]_t=\sum_{s\leq t}(\Delta X_s)^2$ for $t\geq 0$. It follows from Ito's formula that every cadlag, adapted and finite variation process X is quadratic pure jump.

Indeed the finite variation property yields a well-known Ito formula for $f(x)=x^2$ (notice that the second order terms is missing in the sum) of the type

$$X_t^2=2(X_-\bullet X)+\sum_{s\leq t}\left\{X_s^2-X_{s-}^2-2X_{s-}\Delta X_s\right\}=2(X_-\bullet X)+\sum_{s\leq t}(\Delta X_s)^2,$$

which yields the result on the quadratic variation. Hence for every good integrator M we obtain

$$[X,M]_t=\sum_{s\leq t}\Delta X\Delta M$$

for finite variation processes X with complete analogous arguments.

5.4. **Stochastic exponentials.** An instructive example how to calculate with jump processes is given by the following process: let X be a good integrator with $X_0=0$, then the process

$$Z_t=\exp\left(X_t-\frac{1}{2}[X,X]_t\right)\prod_{0\leq s\leq t}(1+\Delta X_s)\exp\left(-\Delta X_s+\frac{1}{2}(\Delta X_s)^2\right)$$

satisfies $Z_t=1+(Z_-\bullet X)_t$ and is called stochastic exponential.

For the proof we have to check that the infinite product is actually converging and defining a good integrator. We show this by proving that it defines an adapted, cadlag process of finite variation. We only have to check this for jumps smaller than $\frac{1}{2}$, i.e. we have to check whether

$$\sum_{s\leq t}\{\log(1+U_s)-U_s+\frac{1}{2}U_s^2\}$$

converges absolutely, where $U_s := \Delta X_s 1_{\{|\Delta X_s|\leq\frac{1}{2}\}}$, for $s\geq 0$. This, however, is true since $|\log(1+x)-x+\frac{1}{2}x^2|\leq Cx^3$ for $|x|\leq\frac{1}{2}$ and $\sum_{s\leq t}\Delta X_s{}^2\leq[X,X]<\infty$ almost surely.

Hence we can apply Ito's formula for the function $\exp(x_1)x_2$ with good integrators

$$X_t^1 = X_t - \frac{1}{2}[X,X]_t$$

and

$$X_t^2 = \prod_{0\leq s\leq t}(1+\Delta X_s)\exp\big(-\Delta X_s+\frac{1}{2}(\Delta X_s)^2\big)\,.$$

This leads to

$$\begin{aligned}
Z_t &= 1+(Z_-\bullet X)_t-\frac{1}{2}(Z_-\bullet[X,X])_t+(\exp(X^1_-)\bullet X^2)_t+\frac{1}{2}(Z_-\bullet[X^1,X^1])_t+\\
&+\sum_{s\leq t}\big\{Z_s-Z_{s-}-Z_{s-}\Delta X_s^1-\exp(X^1_{s-})\Delta X_s^2-\frac{1}{2}Z_{s-}(\Delta X_s^1)^2\big\}=\\
&=1+(Z_-\bullet X)_t-\frac{1}{2}(Z_-\bullet[X,X])_t+\frac{1}{2}(Z_-\bullet[X^1,X^1])_t+\\
&+\sum_{s\leq t}\big\{Z_{s-}\Delta X_s-Z_{s-}\Delta X_s^1-\frac{1}{2}Z_{s-}(\Delta X_s^1)^2\big\}=\\
&=1+(Z_-\bullet X)_t-\frac{1}{2}(Z_-\bullet[X,X])_t+\frac{1}{2}(Z_-\bullet[X^1,X^1])_t+\\
&+\sum_{s\leq t}\big\{\frac{1}{2}Z_{s-}(\Delta X_s)^2-\frac{1}{2}Z_{s-}(\Delta X_s^1)^2\big\}=\\
&=1+(Z_-\bullet X)_t\,,
\end{aligned}$$

since

$$Z_s = Z_{s-}\exp\big(\Delta X_s-\frac{1}{2}(\Delta X_s)^2\big)(1+\Delta X_s)\exp\big(-\Delta X_s+\frac{1}{2}(\Delta X_s)^2\big)$$

holds true, for $s\geq 0$.

5.5. **Lévy's Theorem.** Another remarkable application is Lévy's theorem: consider local martingales $B^1,\ldots,B^n$ starting at 0 with continuous trajectories such that $[B^i,B^j]_t=\delta^{ij}t$ for $t\geq 0$. Then $B^1,\ldots,B^n$ are standard Brownian motions.

For this theorem we have to check that stochastic integrals along locally square integrable martingales are locally square integrable martingales: indeed let X be a locally square-integrable martingale and $H\in\mathbb{L}$, then by localizing and the formula

$$(H\bullet X)^\tau=(H\bullet X^\tau)=(H1_{[0,\tau]}\bullet X^\tau)=(H1_{[0,\tau]}\bullet X)$$

we can assume that H is bounded and X is in fact a square integrable martingale with $E[X_\infty^2]<\infty$. Then, however, we have by Ito's insight that for sequence of partitions tending to identity the process

$$(H^{\Pi^k}\bullet X)$$

is a square integrable martingale, which satisfies additionally

$$E[(H^{\Pi^k} \bullet X)_\infty^2] \leq \|H\|_\infty E[X_\infty^2] \,.$$

By martingale convergence this means that the limit in probability of the stochastic integrals is also a square integrable martingale, whence the result.

This can be readily applied to the stochastic process

$$M_t := \exp\big(\mathrm{i}\langle \lambda, B_t\rangle + \frac{t}{2}\|\lambda\|^2\big)$$

for $t \geq 0$ and $\lambda \in \mathbb{R}^n$, which is a bounded local martingale by the previous consideration and Ito's formula. A bounded local martingale is a martingale, whence

$$E[\exp(\mathrm{i}\langle \lambda, B_t - B_s\rangle) \,|\, \mathcal{F}_s] = \exp(-\frac{t-s}{2}\|\lambda\|^2) \,.$$

6. Bichteler-Dellacherie-Mokobodzki Theorem

A semi-martingale X has a decomposition $X = M + A$, where M is a cadlag local martingale and A is a cadlag process of finite variation. Of course all processes are considered adapted with respect to the given filtration (with usual conditions). The Bichteler-Dellacherie-Mokobodzki Theorem tells every good integrator is a semi-martingale: we present here a proof of Christophe Stricker, which is particularly simple and very instructive, and does not use the Doob-Meyer decomposition in continuous time: instead we are just working with sets bounded in probability.

We can apply in the sequel an L^2 version of Komlos theorem (even though we can also just work with weak convergence): let $(g_n)_{n\geq 1}$ be a bounded sequence in $L^2(P)$, then we can find elements $h_n \in C_n := \operatorname{conv}(g_n, g_{n+1}, \ldots)$ which converge almost surely and in $L^2(P)$ to some element h. For the proof we take

$$A = \sup_{n\geq 1} \inf_{g\in C_n} \|g\|^2 \,,$$

then there are elements $h_n \in \operatorname{conv}(g_n, g_{n+1}, \ldots)$ such that $\|h_n\|^2 \leq A + \frac{1}{n}$. Fix $\epsilon > 0$, then there is n large enough such that for all $k, m \geq n$ the inequality $\|h_k + h_m\|^2 > 4(A - \epsilon)$ holds true, since the sup is along an non-decreasing sequence!. By the parallelogram-identity we then obtain

$$\|h_k - h_m\|^2 = 2\|h_k\|^2 + 2\|h_m\|^2 - \|h_k + h_m\|^2 < 4(A + \frac{1}{n}) - 4(A - \epsilon) = 4\epsilon + \frac{1}{n} \,,$$

which yields the assertion of $L^2(P)$ convergence. By passing to a subsequence the almost sure convergence follows, too.

Remark 6.1. It is of utmost importance to understand that the convex hull of a set of random variables, which is bounded in probability, is not necessarily bounded in probability. Take for instance a sequence of non-negative, independent, identically distributed random variables (X_n) with infinite first moment. Then LLN tells that $\frac{1}{n}\sum_{i=1}^n X_i \to \infty$ almost surely by monotone convergence. Whence the convex hull of (X_n) is not bounded in probability even though apparently the sequence itself is.

We shall actually prove something more general here: we even consider the good integrator property without path properties.

Remark 6.2. Notice that due to the absence of path properties the running maximum process is actually defined as essential supremum over $[0,t]$, i.e. the smallest random variable X_t^* which dominates X_s for all $0 \leq s \leq t$.

Theorem 6.3. *Let X be a stochastic processes (without any assumption on path properties), such that the convex set*

$$\big\{ (H \bullet X)_t \mid \|H\|_\infty \leq 1 \text{ with } H \in \mathbb{S}_u \text{ jumping only at deterministic times } \big\}$$

is bounded in probability for any $t \geq 0$. Then there exists a càdlàg local martingale M and a finite variation process A such that $X = M + A$.

Proof. It is of course enough to fix $t = 1$ and prove the statement on $[0,1]$, where from it follows by concatenation for $[0,\infty[$. We shall also assume $X_0 = 0$. The proof is now split is several steps:

(1) From the assumption it follows immediately that the not necessarily convex set
$$\big\{ |(H \bullet X)_t|^* \mid \|H\|_\infty \leq 1 \big\}$$
is bounded in probability. Indeed, if it were untrue there exists $\epsilon > 0$ and a sequence H^n of strategies uniformly bounded by 1 such that
$$P\big(|(H^n \bullet X)_t|^* \geq n \big) > \epsilon \,,$$
which can also be seen via stopping when n is exceeded, i.e. we find stopping times τ_n (with values in an appropriately chosen finite set depending on n, due to the nature of the essential supremum, and in order to stay on finite deterministic grids)
$$P\big(|(H^n \bullet X)_t|^{\tau_n} = |(H^n 1_{[0,\tau_n]} \bullet X)_t| \geq n \big) > \epsilon \,,$$
which in turn contradicts the assumption. Additionally we can write
$$\sum_i (X_{t_{i+1}} - X_{t_i})^2 = X_{t_n}^2 - 2(H \bullet X)_{t_n} \,,$$
where $H = \sum_i X_{t_i} 1_{]t_i,t_{i+1}]}$ along each deterministic grid. Now $|H|_1^* \leq |X|_1^*$ by the previous consideration, whence we can conclude that the convex hull
$$\big\{ [X,X]^\Pi \mid \text{ for any partition with deterministic times } \Pi \big\}$$
is bounded in probability. Therefore we can find a probability measure $Q \sim P$ such that
$$\sup_\Pi E_Q\big[[X,X]_1^\Pi \big] + E[(|X|_1^*)^2] + \sup_{\|H\|_\infty \leq 1} E_Q[(H \bullet X)] \leq U < \infty$$
by the Nikisin-Yan theorem.

(2) By choosing appropriate strategies H, e.g. taking on $[t_i, t_{i+1}[$ the value
$$\operatorname{sign} E[X_{t_{i+1}} - X_{t_i} | \mathcal{F}_{t_i}]$$
we obtain that the good integrator (without path properties!) is actually of finite mean variation
$$\text{m-var}(X) = \sup_\Pi E_Q\Big[\sum_i \big| E[X_{t_{i+1}} - X_{t_i} | \mathcal{F}_{t_i}] \big| \Big] < U$$
with respect to Q.

Take finally a strategy $H=\sum_i a_i 1_{]t_i,t_{i+1}]}$ uniformly bounded by 1 with grid $\Pi=\{0=t_0<\ldots<t_n\}$, then we can define a discrete martingale

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

where $A_{t_i}=a_i E[X_{t_{i+1}}-X_{t_i}|\mathcal{F}_{t_i}]$. By Doob's maximal inequality we obtain that $E[(|M|^*_{t_n})^2]\leq 4E[M^2_{t_n}]\leq 16U$, whence

$$E_Q[\sup_{t\in\Pi}[|(H\bullet X)|_t]\leq 4\sqrt{U}+\text{m-var}(X)$$

with the right hand side not depending on n and H. So we obtain

$$\sup_{\|H\|_\infty\leq 1} E_Q[|(H\bullet X)|^*_1]<\infty\,.$$

(3) Finally this yields for a strategy $H=\sum_i a_i 1_{]t_i,t_{i+1}]}$ uniformly bounded by 1 with grid $\Pi=\{0=t_0<\ldots<t_n\}$

$$(H\bullet X)^2=\sum_i a_i^2(X_{t_{i+1}}-X_{t_i})^2+2(K\bullet X)$$

with $K=\sum_j a_j(\sum_i^{j-1} a_i(X_{t_{i+1}}-X_{t_i}))1_{]t_j,t_{j+1}}$, where $K\leq|(H\bullet X)|^*$. By the previous result the convex hull of $|(H\bullet X)|^*_1$ is bounded in probability for $\|H\|_\infty\leq 1$, whence the convex hull of $(H\bullet X)^2$ is bounded in probability since the convex set

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

is bounded in probability and the sum of two convex sets bounded in probability is convex and bounded in probability; the first set being the convex hull of

$$\sum_i a_i^2(X_{t_{i+1}}-X_{t_i})^2\,.$$

With these preliminary steps there exists a measure $Q\sim P$ such that

$$\sup_{\|H\|_\infty\leq 1,\Pi} E_Q[(H\bullet X)_1^2+[X,X]^\Pi+(H\bullet X)_1]<\infty\,.$$

The discrete time martingale

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

for $t\in\Pi$ is well defined square integrable martingale with respect to the corresponding discrete filtration . Furthermore

$$\sup_\Pi E_Q\Big[\big(M_1^\Pi\big)^2\Big]<\infty$$

yielding a weakly converging sequence M^{Π^k} (which can be chosen refining with mesh going to 0) converging to a limit M_1. We denote the *continuous time* martingale with càdlàg trajectories generated by M_1 by M, i.e. almost surely $M_t=E[M_1\mid\mathcal{F}_t]$ for $0\leq t\leq 1$ and define $A=X-M$. Notice that A as well as X does a priori not have path properties, in contrast to the càdlàg martingale M.

We consider now the set of all time points belonging to the partitions Π^k and denote it by Π, i.e. $\Pi=\cup\Pi^k$.

Choosing a partition $\sigma = \{0 = s_0 < \ldots < s_n\}$ consisting of points in Π we can readily prove that for any element $Y \in L^2(Q)$ with unit length

$$E[Y \sum_i |A_{s_{i+1}} - A_{s_i}|] \leq \sup_k (E\Big[\big|\sum_i E[X_{t_{i+1}} - X_{t_i}|\mathcal{F}_{t_i}]\big|^2\Big])^{1/2},$$

which in turn can be estimated by weak convergence

$$\sup_{\|H\|_\infty \leq 1, \Pi} E_Q[(H \bullet X)_1^2 + 2[X,X]^\Pi] < \infty .$$

Whence the total variation of A calculated via time points chosen from Π is finite. This is not enough to conclude that the total variation is finite, in contrast to claims [7]. We have to argue here a bit further.

Even though we do not have a finite total variation of A yet we can still follow Rao's classical calculus with so called natural processes, namely for every square integrable càdàg martingale N and every sequence of partitions chosen in Π and tending towards the identity we have

$$\lim_{n\to\infty} E[(N_-^{\sigma_n} \bullet A^{\sigma_n})] = \lim_{n\to\infty} E[\sum_i N_{t_i}(X_{t_{i+1}} - X_{t_i})]$$

$$= \lim_{n\to\infty} E[\sum_i N_1 E[X_{t_{i+1}} - X_{t_i}|\mathcal{F}_{t_i}]] = E[N_1 A_1]$$

since $A = X - M$ and M is an L^2 martingale. Assume now that we have constructed two martingales M^1 and M^2 with associated partition points $\Pi_1 \subset \Pi_2$, then we can ask whether they are equal. Indeed it holds that

$$M^1 - M^2 = A^2 - A^1$$

by construction, i.e. on Π_1 we can consider the square integrable càdlàg martingale $A^2 - A^1$ and perform the above calculus with $N = A^2 - A^1$ and $A = A^2 - A^1$. Hence we obtain $E[(A_t^2 - A_t^2)^2] = 0$ for every $0 \leq t \leq 1$, so A^2 is a version of A^1. Additionally we have that A^1 also has finite total variation on Π_2. Therefore we can conclude that A^1 has finite total variation since Π_2 was arbitrary.

For the rest of the proof we just apply Girsanov's theorem in its general form. □

7. Doob-Meyer decomposition

The decomposition of good integrators X into a finite variation process A and a local martingale M can be refined for super-martingales even beyond càdlàg trajectories. Actually the finite variation process can be chosen predictable and increasing.

We shall first prove the result for bounded, non-negative supermartingales X using estimates for discrete super-martingales. Subsequently The proof will follow lines from [7]. Then we shall use a stopping argument which has to be done with some care due to absence of path properties.

Theorem 7.1. *Let S be a bounded, non-negative super-martingale (without any path properties) on $[0,1]$, then there exists a càdlàg local martingale M and a predictable, increasing finite variation process A such that $S = M - A$. The decomposition is unique up to versions of A and indistinguishibility of M.*

Let S be a non-negative super-martingal such that a sequence τ_n of stopping times exist which stops S at level n, i.e.

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

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

Proof. First we prove the theorem for bounded, non-negative super-martingales $0 \leq S \leq c$. Indeed in this case we obtain for any partiation Π that the sampled super-martingale S^Π admits a discrete Doob-Meyer decomposition with respect to the approriately discretized filtration. For this Doob-Meyer decomposition $S^\Pi = M^\Pi - A^\Pi$

$$E[M_1^\Pi] \leq 2cE[S_0]$$

holds true, whence we can find a sequence of partitions Π^k, refining and with mesh converging to 0, such $M_1^{\Pi^k} \to M_1$ weakly in $L^2(P)$. With this random variable we can define a càdlàg square-integrable martingale M via conditioning, and additionally $A = M - S$, which is a conditionally increasing process. The process is indeed increasing when sampled on time points $\cup \Pi^k$ and therefore of finite variation thereon. Also naturality as in the previous proof holds true, i.e.

$$\lim_{n \to \infty} E[(N_-^{\sigma_n} \bullet A^{\sigma_n})] = E[N_1 A_1] \, ,$$

hence we can conclude that on increasing sets of time points $\Pi_1 \subset \Pi_2$ with associated Doob-Meyer decompositions $S = M^1 - A^1 = M^2 - A^2$ we do actually have uniqueness in version for A and uniqueness up to indistinguishability for M^1 and M^2. Therefore the decomposition into a finite variation process A and a càdlàg martingale is proved, where A is additionally increasing. It remains to prove that A is actually predictable. This, however, follows from general facts on so called natural processes A, i.e. finite variation processes which satisfy

$$E[(N_- \bullet A)_1] = E[N_1 A_1]$$

Second we prove the theorem for a general super-martingale S. Choose a stopping time τ^n, which stops the super-martingale S at level n. Then

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

holds true. Clearly $S \wedge n$ is a bounded super-martingale and therefore we have, by the previous consideration, a Doob-Meyer decomposition. The second term can be written as

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

where the first summand is a càdlàg martingale and the second part is a decreasing process, if n is larger than $E[S_0]$. Whence we have a decomposition into a martingale M^n and an non-decreasing finite variation process A^n. Furthermore this decomposition is unique. Whence for $n \geq m$ apparently

$$(M^n)^{\tau_m} = M^m$$

as well as $(A^n)^{\tau_m} = A^m$. Therefore we can define a càdlàg local martingale M and an increasing finite variation process A such that

$$S = M - A \, .$$

□

Remark 7.2. Assume that S is a non-negative càgàg super-martingale of class (D), i.e. the set of all random variables S_τ for all possible stopping times with values in $[0, 1]$ is uniformly integrable, then the Doob-Meyer decomposition actually can be done with a martingale M. Indeed, since the Doob-Meyer decomposition for stopping times τ_n stopping the process when it reaches level n works with a martingale, we obtain a sequence of martingales M^{τ_n} for the Doob-Meyer decomposition of S^{τ_n}. By uniform integrability M^{τ_n} is uniformly integrable and hence M is a martingale.

8. Stochastic Integration for predictable integrands

For many purposes (i.e. martingale representation) it is not enough to consider only caglad integrands, but predictable integrands are needed. This cannot be achieved universally for all good integrators, but has to be done case by case. The main tool for this purpose are $\mathcal{H}^p$ spaces, for $1 \leq p < \infty$, which are spaces of martingales with certain integrability properties, the most important being $\mathcal{H}^1$. We present first the $\mathcal{H}^p$ and specialize then to $p = 1$ and $p = 2$. This is inspired by [6] and does explicitly not make use of the fundamental theorem of local martingales.

Main tool for the analysis are the Burkholder-Davis-Gundy inequalities:

Theorem 8.1. *For every $p \geq 1$ there are constants $0 < c_p < C_p$ such that for every martingale*

$$c_p E\big[[M,M]_\infty^{\frac{p}{2}}\big] \leq E\big[(|M|_\infty^*)^p\big] \leq C_p E\big[[M,M]_\infty^{\frac{p}{2}}\big]$$

holds true.

Remark 8.2. The inequalities follow from the same inequalities for discrete martingales, which can be proved by deterministic methods. In fact equations of the type

$$(h \bullet M)_T + [M,M]_T^{\frac{p}{2}} \leq (|M|_T^*)^p \leq C_p [M,M]_T^{\frac{p}{2}} + (g \bullet M)_T$$

hold true, with predictable integrands h, g and martingales M on a finite index set with upper bound T hold, see [1].

Let M be a martingale and let us take a sequence of refining partitions Π^n tending to identity, for which $M^{\Pi^n} \to M$ in ucp and $[M^{\Pi^n}, M^{\Pi^n}] \to [M,M]$ in ucp. Fix some time horizon $T > 0$, then by monotone convergence

$$E\big[(|M^{\Pi^n}|_T^*)^p\big] \to E\big[(|M|_T^*)^p\big]$$

as $n \to \infty$, since the sequence of partitions is refining. If $E\big[(|M|_T^*)^p\big] = \infty$, we obtain that all three quantities are infinity. If $E\big[(|M|_T^*)^p\big] < \infty$ we obtain by dominated convergence that

$$E\big[[M^{\Pi^m} - M^{\Pi^n}, M^{\Pi^m} - M^{\Pi^n}]_T^{\frac{p}{2}}\big] \leq E\big[(|M^{\Pi^m} - M^{\Pi^n}|_T^*)^p\big] \to 0$$

which means by

$$E\Big[\big|[M^{\Pi^m}, M^{\Pi^m}]^{\frac{1}{2}} - [M^{\Pi^n}, M^{\Pi^n}]_T^{\frac{1}{2}}\big|^p\Big] \leq E\big[[M^{\Pi^m} - M^{\Pi^n}, M^{\Pi^m} - M^{\Pi^n}]_T^{\frac{p}{2}}\big]$$

the L^p convergence of the quadratic variations to $[M,M]$. This yields the result for any $T > 0$ and hence for $T \to \infty$.

Remark 8.3. The case $p = 2$ can be readily derived from Doob's maximal inequality, see Theorem 3.3, and we obtain

$$E\big[[M,M]_\infty\big] = E\big[M_\infty^2\big] \leq E\big[(|M|_\infty^*)^2\big] \leq 4E\big[M_\infty^2\big] = E\big[[M,M]_\infty\big]\,.$$

Definition 8.4. *Let $p \geq 1$ be given. Define the vector space $\mathcal{H}^p$ as set of martingales M where*

$$\|M\|_{\mathcal{H}^p}^p := E\big[(|M|_\infty^*)^p\big] < \infty$$

holds true.

By the Burkholder-Davis-Gundy inequalities the following theorem easily follows:

Theorem 8.5. *For $p \geq 1$ the space $\mathcal{H}^p$ is a Banach space with equivalent norm*

$$M \mapsto E\big[[M,M]_\infty^{\frac{p}{2}}\big]^{\frac{1}{p}}.$$

For $p = 2$ the equivalent norm is in fact coming from a scalar product

$$(M,N) \mapsto E\big[[M,N]_T\big].$$

Additionally we have the following continuity result: $M^n \to M$ in $\mathcal{H}^p$, then $(Y \bullet M_n) \to (Y \bullet M)$ in ucp for any left-continuous process $Y \in \mathbb{L}$. In particular $[M^n, N] \to [M,N]$ in ucp.

In the next step we consider a weaker topology of L^p type on the set of simple predictable integrands. The following lemma tells about the closure with respect to this topology.

Lemma 8.6. *Let A be an increasing finite variation process and V a predictable process with*

$$(|V|^p \bullet A)_t < \infty\,,$$

then there exists a sequence of bounded, simply predictable processes V^n in $b\mathbb{S}$ such that

$$(|V - V^n|^p \bullet A)_t \to 0\,,$$

as $n \to \infty$.

Proof. By monotone class arguments it is sufficient to prove the lemma for LCRL processes, for which it is, however, clear, since they can be approximated in the ucp topology by simply predictable processes. □

The main line of argument is to construct for predictable processes V satisfying certain integrability conditions with respect to $[M,M]$ a stochastic integral $(V \bullet M)$ for $M \in \mathcal{H}^p$. We take the largest space $\mathcal{H}^1$ in order to stay as general as possible.

Proposition 8.7. *Let $M \in \mathcal{H}^1$ be fixed and let V^n be a sequence of bounded, simple predictable processes such that*

$$E\big[(|V - V^n|^2 \bullet [M,M])_\infty^{\frac{1}{2}}\big] \to 0$$

then the sequence $(V^n \bullet M)$ is a Cauchy sequence in $\mathcal{H}^1$ defining an element $(V \bullet M)$, which does only depend on V and not on the approximating sequence V^n and which is uniquely determined by

$$[(V \bullet M), N] = (V \bullet [M,N])$$

for martingales N.

Proof. This is a direct consequence of the Burkholder-Davis-Gundy inequalities, since

$$E\big[|(V^n \bullet M) - (V^m \bullet M)|^*_\infty)\big] \leq C_1 E\big[((V^n - V^m)^2 \bullet [M,M])^{\frac{1}{2}}_\infty\big] \to 0$$

as $n, m \to \infty$. Whence $(V \bullet M)$ is a well-defined element of $\mathcal{H}^1$, which only depends on V and not on the approximating sequence. For all martingales N and all simple predictable strategies the formula

$$[(V^n \bullet M), N] = (V^n \bullet [M,N])$$

holds true by basic rules for LCRL integrands. By passing to the limit we obtain the general result. Uniqueness is clear since $[M,M] = 0$ means $M =$ by Burkholder-Davis-Gundy inequalities. □

Definition 8.8. *Let $M \in \mathcal{H}^1$, then we denote by $L^1(M)$ the set of predictable processes V such that*

$$E\big[(|V|^2 \bullet [M,M])^{\frac{1}{2}}_\infty\big] < \infty\,.$$

Apparently we have constructed a bounded linear map $L^1(M) \to \mathcal{H}^1$, $V \mapsto (V \bullet H)$. The set of integrands $L^1(M)$ is not the largest one, we can still generalize it by localization, which defines the set $L(M)$: a predictable process V is said to belong to $L(M)$ if $(|V|^2 \bullet [M,M])^{\frac{1}{2}}_\infty$ is locally integrable, which means for bounded variation processes nothing else than just being finite, see [6]. Notice that this is the largest set of integrands given that we require that the integral is a semi-martingal having a quadratic variation, which coincides with $(V^2 \bullet [M,M])$. Notice also that by the same argument every local martingale is in fact locally $\mathcal{H}^1$, which in turn means that we can define for any semi-martingale a largest set of integrands.

References

[1] Mathias Beiglböck and Pietro Siorpaes. Pathwise versions of the Burkholder-Davis-Gundy inequalities. *preprint*, 2013.

[2] Klaus Bichteler. Stochastic integration and L^p-theory of semimartingales. *Ann. Probab.*, 9(1):49–89, 1981.

[3] Olav Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[4] Paul-André Meyer. *Martingales and stochastic integrals. I. Lecture Notes in Mathematics. Vol. 284.* Springer-Verlag, Berlin-New York, 1972.

[5] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York).* Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.

[6] A. N. Shiryaev and A. S. Cherny. A vector stochastic integral and the fundamental theorem of asset pricing. *Tr. Mat. Inst. Steklova*, 237(Stokhast. Finans. Mat.):12–56, 2002.

[7] C. Stricker. Caractérisation des semimartingales. In *Seminar on probability, XVIII*, volume 1059 of *Lecture Notes in Math.*, pages 148–153. Springer, Berlin, 1984.

ETH Zürich, D-Math, Rämistrasse 101, CH-8092 Zürich, Switzerland
Email address: jteichma@math.ethz.ch