4. Stochastic Integral

4.1. Continuous Time Processes

Throughout, $(\Omega, \mathcal{F}, P)$ denotes a probability space. From now on however we no longer assume that $\mathbf{T} = \mathbb{N}_0$ but[41]

$$\mathbf{T} = [0, \infty[\quad \text{and} \quad T = \sup \mathbf{T} = \infty$$

Remark 4.1. If not otherwise specified, elements of $\mathbf{T}$ are designed by the letter $s, t, u, \ldots$. Countable and dense subset of $\mathbf{T}$ is denoted by $\mathbf{Q}$ and their elements are designed by the letter q, r. Often $\mathbf{Q} = \mathbb{Q}$.♦

As random variable may be identified in the P-almost sure sense, for processes we have several possibilities at hand.

Definition 4.2. Let X and Y be two processes. We say that

- X is a *modification* of Y if $X_t = Y_t$ P-almost surely for every $t \in \mathbf{T}$, that is

$$P[X_t = Y_t] = 1, \quad \text{for every } t \in \mathbf{T}.$$

- X is indistinguishable from Y if $X_t = Y_t$ for all $t \in \mathbf{T}$ P-almost surely, that is

$$P[X_t = Y_t; \text{ for every } t \in \mathbf{T}] = 1.$$

In the case where $\mathbf{T}$ is countable, these two notions are equivalent. However, if $\mathbf{T}$ is no longer countable, modification and indistinguishability are different in that we may have to take into account uncountably many null sets as the following example shows.

Example 4.3. Set $(\Omega, \mathcal{F}, P) := ([0,1], \mathcal{B}([0,1]), dt)$, where dt is the Lebesgue measure and $\mathcal{B}([0,1])$ is the Borel σ-algebra, and $\mathbf{T} = [0,1]$. Define the processes X and Y by

$$X_t = 0 \quad \text{and} \quad Y_t = \begin{cases} Y_t(\omega) = 1 & \text{if } \omega = t \\ Y_t(\omega) = 0 & \text{otherwise} \end{cases}$$

for every $t \in [0,1]$. It follows that $P[X_t = Y_t] = P[\{\omega \in [0,1] : \omega \neq t\}] = 1$, whereas $P[X_t = Y_t :$ for every $t] = P[\{\omega \in [0,1] : \omega \neq t \text{ for every } t\}] = P[\emptyset] = 0$. ♢

We see that uncountably many sets of measure zero can add up to something that may no longer have measure zero. However, if the time set is countable, or if we can infer from the structure of the trajectories that it is sufficient to consider countably many times, then these two conditions will coincide. We say that a process X

- has "*des limites à gauche*" – left limits – or "*des limites à droite*" – right limits – if

$$P\left[\liminf_{s \nearrow t} X_s = \limsup_{s \nearrow t} X_s : \text{ for every } t > 0, t \in \mathbf{T}\right] = 1,$$

  or

$$P\left[\liminf_{s \searrow t} X_s = \limsup_{s \searrow t} X_s : \text{ for every } t < T, t \in \mathbf{T}\right] = 1,$$

  respectively.

[41] Eventually, $\mathbf{T}$ may be equal to $[0, T]$ for some $T < \infty$.

- is "*continue à gauche*" – left-continuous – or "*continue à droite*" – right-continuous – if

$$P\left[\lim_{s\nearrow t} X_s = X_t : \text{ for every } t > 0, t \in \mathbf{T}\right] = 1,$$

or

$$P\left[\lim_{s\searrow t} X_s = X_t : \text{ for every } t < T, t \in \mathbf{T}\right] = 1,$$

respectively.[42]

A process X is said to be càdlàg, càglàd, or làdlàg, if it is "continue à droite avec des limites à gauche", "continue à gauche avec des limites à droite" or "limité à gauche comme à droite", respectively. In general, these concepts are meaningful if $\mathbf{T}$ is a dense subset of an interval in $[0, \infty]$.

Lemma 4.4. *Suppose that X and Y are modifications of each other and both are either right-continuous or left-continuous. Then X and Y are indistinguishable.*

Proof. Let X and Y be right-continuous and modifications of each other. Since X and Y are right-continuous, it follows that

$$\begin{aligned} P\left[X_t \neq Y_t : \text{for some } t \in \mathbf{T}\right] &= P\left[\lim_{q\searrow t} X_q \neq \lim_{q\searrow t} Y_q : \text{ for some } t \in \mathbf{T}\right] \\ &= P\left[X_q \neq Y_q : \text{for some } q \in \mathbb{Q}\right] = P\left[\cup_{q\in\mathbb{Q}}\{X_q \neq Y_q\}\right] \leq \sum P[X_q \neq Y_q] = 0 \end{aligned}$$

Hence,

$$P\left[X_t = Y_t \colon \text{for every } t\right] = 1$$

showing that X and Y are indistinguishable. □

Exercice 4.5. The assumption of left- or right-continuity is central. Find an example of two làdlàg processes X, Y, modifications of each other, which are not indistinguishable. ◊

As such, a process is nothing else than an arbitrary family of random variables indexed by the time. It can also be seen as a mapping $X : \Omega \times \mathbf{T} \to S$.

Definition 4.6. We say that a process X is *measurable* if it is measurable with respect to the product σ-algebra $\mathcal{F} \otimes \mathcal{B}(\mathbf{T})$.

The definition of a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbf{T}}$ does not change as an increasing family of σ-algebra indexed by time. However, we can also define the right and left filtration $\mathbb{F}^+ = (\mathcal{F}_t^+)_{t\in\mathbf{T}}$ and $\mathbb{F}^- = (\mathcal{F}_t^-)_{t\in\mathbf{T}}$ as follows

$$\mathcal{F}_t^+ = \bigcap_{s>t} \mathcal{F}_s \quad \text{and} \quad \mathcal{F}_t^- = \bigvee_{s<t} \mathcal{F}_s := \sigma\left(\mathcal{F}_s : s < t\right)$$

for $t \in \mathbf{T}$. Clearly these definitions make sense if t is neither minimum or maximum of $\mathbf{T}$. In these extreme cases, we set $\mathcal{F}_0^- = \mathcal{F}_0$ and $\mathcal{F}_T^+ = \mathcal{F}_T$ if $T \in \mathbf{T}$. We say that the filtration is left- or right-continuous if $\mathbb{F} = \mathbb{F}^-$ or $\mathbb{F} = \mathbb{F}^+$, and continuous if it is both.

Remark 4.7. From the definition, $\mathbb{F}^\pm$ are themselves filtrations and it holds $\mathcal{F}_t^- \subseteq \mathcal{F}_t \subseteq \mathcal{F}_t^+$ as well as $\mathcal{F}_s^+ \subseteq \mathcal{F}_t^-$ whenever $s < t$. Hence $(\mathbb{F}^-)^+ = (\mathbb{F})^+ = (\mathbb{F}^+)^+$ as well as $(\mathbb{F}^+)^- = (\mathbb{F})^- = (\mathbb{F}^-)^-$. Clearly, if $\mathbf{T}$ is discrete, for instance $\mathbf{T} = \mathbb{N}_0$, left- and right-continuous filtrations make little sense. ♦

[42] Naturally, when we write lim we implicitly mean that it exists.

Definition 4.8. We say that a stochastic process X is

- $\mathbb{F}$-*adapted* – or simply adapted – if X_t is $\mathcal{F}_t$-measurable for every $t \in \mathbf{T}$;
- $\mathbb{F}$-*progressively measurable* – or simply progressively measurable – if $(\omega, s) \mapsto X_s(\omega)$, $s \leq t$ is $\mathcal{F}_t \otimes \mathcal{B}(\mathbf{T} \cap [0,t])$-measurable for every $t \in \mathbf{T}$.

In particular, progresively measurable processes are automatically adapted. The reciprocal is true if the paths of the process are regular enough.

Proposition 4.9. *Let X be right-continuous or left-continuous $\mathbb{F}$-adapted process. Then X is progressively measurable.*

Proof. Suppose that X is right-continuous and fix $t \in \mathbf{T}$. Define

$$X_s^n = X_{\frac{k+1}{2^n}t}, \quad \text{for } \frac{k}{2^n}t < s \leq \frac{k+1}{2^n}t$$

Since X is right-continuous, it follows that $\lim X^n = X$ on $\Omega \times [0,t]$ up to the null set of those ω on which X does not have right-continuous paths. Furthermore, since X is adapted, it follows that the piecewise constant process X^n is $\mathcal{F}_t \otimes \mathcal{B}(\mathbf{T} \cap [0,t])$-measurable. Hence X is progressively measurable.□

The previous result makes use of the regularity of paths to derive progressive measurability from adaptiveness. The following result goes a step further by showing that measurability together with adaptiveness yields progressive measurability, up to a modification though.

Theorem 4.10. *Any measurable and adapted process admits a progressive modification.*

The proof of this theorem is not trivial and somewhat lengthy in the standard literature, see for instance [1]. Since this result will be needed for the construction of the stochastic integral we provide a recent simpler version of the proof from Ondrejat and Seidler [5] in Section 4.3.
The notion of stopping times also has to be slightly modified in the continuous time

Definition 4.11. On a probability space, a *random time* is a measurable mapping $\tau : \Omega \to \mathbf{T} \cup T$. Given a filtration, a random time is a(n)

- *optional time* if $\{\tau < t\} \in \mathcal{F}_t$ for every $t \in \mathbf{T}$.
- *stopping time* if $\{\tau \leq t\} \in \mathcal{F}_t$ for every $t \in \mathbf{T}$.

Proposition 4.12. *Every stopping time is is an optional time, and every optional time is a stopping time for the right-filtration. In particular, the two notions coincide if the filtration is right-continuous.*

Proof. The first assertion is trivial. As for the second, let τ be an optional time, and $t \in \mathbf{T}$. It follows that $\{\tau \leq t\} = \cap_n \{\tau < t + 1/n\} \in \mathcal{F}_t^+$. □

For a process X and a subset V of the state space we define the *hitting time* of X in V as

$$\tau_V(\omega) = \inf\{t \in \mathbf{T} : X_t(\omega) \in V\}.$$

This function is not necessarily random even if X is adapted, however we have the following.

Proposition 4.13. *If X is an adapted right-continuous process and V is open, then τ_V is an optional time. If X is a continuous adapted process and V is closed, then τ_V is a stopping time.*

Proof. It holds $\{\tau_V < t\} = \{\omega \in \Omega : X_s(\omega) \in V, s < t\}$. Since X is right-continuous and V is open, $X_s(\omega) \in V$ implies the existence of a rational $q < s$ such that $X_q(\omega) \in V$. Hence $\{\tau_V < t\} = \{X_q \in V : q < t\} = \cup_{q<t}\{X_q \in V\} \in \mathcal{F}_t$. For the case of X being continuous and V closed, define the open sets $V_n = \{x : d(x, V) < 1/n\} \supseteq V$. Then by continuity of X we obtain

$$\{\tau_V \leq t\} = \{X_t \in V\} \bigcup \left(\bigcap_n \bigcup_{q<t} \{X_q \in V_n\} \right) \in \mathcal{F}_t. \qquad \square$$

Let us collect some standard properties of optional and stopping times.

Proposition 4.14. *The following assertions hold*

(a) *Every constant t is a stopping time.*

(b) *$\tau + \sigma$, $\tau \vee \sigma$ and $\tau \wedge \sigma$ are stopping/optional times as soon as τ, σ are stopping/optional times.*

(c) *$\lim \tau^n$ is a stopping time as soon as (τ^n) is an increasing sequence of stopping times.*

(d) *$\lim \tau^n$ is an optional time as soon as (τ^n) is a decreasing sequence of optional times. It is a stopping time if (τ^n) are stationary stopping times, that is, $\tau^m(\omega) = \tau^n(\omega)$ for all m greater than a given n, for P-almost all $\omega \in \Omega$.*[43]

(e) *If τ is a stopping time, then the collection $\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t\}$ is a σ-algebra and τ is $\mathcal{F}_\tau$-measurable.*

(f) *For any two stopping times, it holds $\mathcal{F}_\sigma \cap \mathcal{F}_\tau = \mathcal{F}_{\sigma\wedge\tau}$. In particular, $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$, if $\sigma \leq \tau$. For every random variable X, it holds $E[E[X \mid \mathcal{F}_\sigma] \mid \mathcal{F}_\tau] = E[X \mid \mathcal{F}_{\sigma\wedge\tau}]$.*

Proof. The proof follows the same argumentation as in the discrete time by noting that $\mathbb{Q}$ is a countable dense subset of $\mathbf{T}$. Only the following two points need a certain care.

(c) let τ and σ be two stopping times, let us show that the summ is still a topping time. Noting that τ is a stopping times if and only if $\{\tau > t\} = \{\tau \leq t\}^c \in \mathcal{F}_t$ for every t, the following decomposition holds

$$\{\tau + \sigma > t\} = \{\tau = 0, \sigma > t\} \cup \{\sigma = 0, \tau > t\} \cup \{\tau \geq t, \sigma > 0\} \cup \{\sigma + \tau > t, 0 < \tau < t\}$$

Noting that $\{\tau = 0\} = \{\tau \leq 0\} \in \mathcal{F}_0 \subseteq \mathcal{F}_t$, the same for $\{\sigma = 0\} \in \mathcal{F}_t$, it follows immediately that the first two sets are in $\mathcal{F}_t$. Further, $\{\tau \geq t\} = \cap_n\{\tau > t - 1/n\} \in \mathcal{F}_{t-} \subseteq \mathcal{F}_t$ and $\{\sigma > 0\} \in \mathcal{F}_0$ showing that the third set in this decomposition is in $\mathcal{F}_t$. As for the last one, note that

$$\{\sigma+\tau > t, 0 < \tau < t\} = \cup_{0<q<t}\{\sigma > t-q\}\cap\{t > \tau > q\} = \cup_{0<q<t}\{\sigma > t-q\}\cap\{\tau > q\}\cap\{\tau < t\}$$

which is for the same reasong as before in $\mathcal{F}_t$ since $0 < q < t$.

(d) Suppose that τ^n is a decreasing sequence of optional times. It follows from $\{\lim \tau^n < t\} = \{\tau^n < t : \text{ for some } n\} = \cup_n\{\tau^n < t\} \in \mathcal{F}_t$ that $\lim \tau^n$ is an optional time. If τ^n are stopping times, it only holds $\{\lim \tau^n \leq t\} = \cap_{q>0}\{\tau^n \leq t + q : \text{for some } n\} \in \mathcal{F}_t^+$ and therefore $\lim \tau^n$ is optional. However, defining $A_n = \{\tau^n = \tau^m : \text{for all } m \geq n\}$, it follows from stationarity that A_n is increasing to Ω. Furthermore, $A_n \in \mathcal{F}_{\tau^n}$ and hence $\{\lim \tau^n \leq t\} = \cup_n\{\tau^n \leq t\} \cap A_n \in \mathcal{F}_t$. $\square$

Proposition 4.15. *Let X be a progressively measurable process and τ a stopping time. Then $X_\tau(\omega) := X_{\tau(\omega)}(\omega)$ is an $\mathcal{F}_\tau$-measurable random variable. Furthermore, $X^\tau := (X_{\cdot\wedge\tau})$ is a progressive process.*

[43] Note that n may depend on ω.

Proof. First, τ being a stopping time implies that $(\omega, s) \mapsto h(\omega, s) := (\omega, \tau(\omega) \wedge s)$ from $\Omega \times \mathbf{T} \cap [0, t]$ onto itself is $\mathcal{F}_t \otimes \mathcal{B}(\mathbf{T} \cap [0, t])$-measurable for every t. Since X is progressive and $X_s^\tau(\omega) = X \circ h(\omega, s)$ for every $s \leq t$, it follows that $(s, \omega) \mapsto X_s^\tau(\omega)$ is also $\mathcal{F}_t \otimes \mathcal{B}(\mathbf{T} \cap [0, t])$-measurable. Thus X^τ is progressive and, in particular, X_τ is $\mathcal{F}_\tau$ measurable. □

The null sets on a probability space plays a central role. They allow to identify random variables in the almost sure sense. With regard to a filtration indexed by an uncountable time set, this may yield some tricky problems – this is mainly due to the problem of right continuous version of processes not further discussed here, see [1, Theorem III-44 p. 64, Theorems IV-32-33 pp. 102–103]. In order to get rid of these problems and the identification between optional and stopping times we will work with the following assumption.

Definition 4.16. A filtration $\mathbb{F}$ is said to

- be *complete* if $\mathcal{F}_0$ contains all the P-negligible sets of $\mathcal{F}$;
- satisfy the *usual conditions* if it is complete and right-continuous, that is $\mathbb{F}^+ = \mathbb{F}$.

From now on:

$$\mathbb{F} = \mathbb{F}^+ \quad \text{and} \quad \mathcal{F}_0 \text{ contains all the } P \text{ null sets of } \mathcal{F}$$

For a stopping/optional time τ, we denote by $[\tau] = \{(\omega, t) \in \Omega \times \mathbf{T} : \tau(\omega) = t\}$ its graph.

Proposition 4.17. *Let X be a càdlàg, adapted process on a filtration satisfying the usual conditions. Then there exists a sequence of stopping times (τ^n) which exhausts the jumps*[44] *$\Delta X = X - X_-$ of X, that is*

$$\{\Delta X \neq 0\} \subseteq \bigcup_n [\tau^n] .$$

The definition of (Super-/Sub-)Martingales do not change in the continuous time. The martingale properties and theorems extend to the continuous time. However some restrictions have to be made in terms of path regularity. Indeed, as mentioned earlier, the infinite amount of null sets that may add up has to be countably controlled.

Definition 4.18. We denote by $\mathcal{S}$ be the set of *simple predictable* processes H of the form

$$H_t = H_0 1_{\{0\}} + \sum_{k=1}^{n} H_k 1_{]\tau_{k-1}, \tau_k]}(t)$$

for a finite sequence $0 = \tau_0 \leq \tau_1 \leq \ldots \leq \tau_n$ of stopping times where τ_n is bounded, and $H_k \in L^0_{\tau_{k-1}}$ for $k = 0, \ldots, n$.

Per definition, any $H \in \mathcal{S}$ is a predictable progressive process. For a progressive process X and $H \in \mathcal{S}$, we denote by $H \bullet X$ the process

$$H \bullet X := H_0 X_0 + \sum_{k=1}^{n} H_k \left(X^{\tau_k} - X^{\tau_{k-1}} \right) .$$

This process may be seen as the simple integral of H with respect to X, also denoted by $\int H dX$.

Theorem 4.19. *Let X be a process either on $\mathbf{T} = \mathbb{Q}$ countable, or right-continuous on $\mathbf{T} = [0, \infty[$, and $H \in \mathcal{S}$. The following assertions hold true.*

[44] For X càdlàg, the jump process ΔX is the difference of X with the càglàd version X_- of X

(a) If X is a martingale and $H \bullet X_t$ is integrable for every t, then $H \bullet X$ is a martingale. If X is a super/submartingale, $H \bullet X_t$ is integrable for every t and H is positive, then $H \bullet X$ is a super/submartingale. In particular if τ is a bounded stopping time then X^τ is a martingale or super/submartingale, respectively.

(b) Let X be a submartingale, $t \in \mathbf{T}$ and $\lambda > 0$. Then it holds

$$\lambda P\left[\bar{X}_t \geq \lambda\right] \leq E\left[1_{\{\bar{X}_t<\lambda\}} X_t\right] \leq E\left[X_t^+\right] \tag{4.1}$$

$$\lambda P\left[\underline{X}_t \leq -\lambda\right] \leq E\left[1_{\{\underline{X}_t>-\lambda\}} X_t\right] - E\left[X_0\right] \leq E\left[X_t^+\right] - E\left[X_0\right] \tag{4.2}$$

(c) Let X be a positive submartingale and $p > 1$, it holds

$$\left\|\sup_{s \leq t} X_s\right\|_p \leq q \left\|X_t\right\|_p$$

where $q = p/(p-1)$ is the conjugate of p.

(d) Let X be a supermartingale, then for every two reals $x < y$, the numbers of up-crossing of $[x, y]$ by X up to time t, $U_{[0,t]}(x, y, X)$ is a random variable and it holds

$$E\left[U_{[0,t]}\left(x, y, X\right)\right] \leq \frac{E\left[(x - X_t)^+\right]}{y - x} \leq \frac{|x| + E\left[|X_t|\right]}{y - x}, \quad t \in \mathbf{T}.$$

In particular, if X is a martingale, and $p > 1$, then $|X|^p$ is a positive submartingale – Jensen inequality – and so

$$\left\|X_t^*\right\|_p \leq \left(\frac{p}{p-1}\right) \left\|X_t\right\|_p$$

for every $p > 1$.

Proof. The inequalities hold true if the process X is sampled on any finite discretization of $[0, t]$ containing 0 and t. Hence, passing to the limit, these inequalities hold for $([0, t] \cap \mathbb{Q}) \cup \{0, t\}$, showing the case $\mathbf{T} = \mathbb{Q}$. In case where $\mathbf{T}$ is continuous and the paths of X are right-continuous, the inequalities also follow as seen before. The single thing to check is whether $U_{[0,t]}(x, y, X)$ is a well defined random variable. However, for any finite $F \subseteq [0, t]$, since X is right-continuous, the τ^k and σ^k in the construction of the $U_F(x, y, X)$ are stopping times according to Proposition 4.13, therefore $U_F(x, y, X)$ is a random variable. It follows that $U_{([0,t]\cap\mathbb{Q})\cup\{0,t\}}\left(x, y, X\right)$ is a random variable. Since X is right-continuous, this set takes into account all the up-crossing on $[0, t]$. $\square$

Theorem 4.20. *Any right-continuous submartingale is càdlàg and every sample path is almost surely bounded on any compact interval. Furthermore, X is a submartingale with respect to $\mathbb{F}^+$ as well as with respect to the augmentation of $\mathbb{F}$.*

Proof. Let X be a right-continuous submartingale. The boundedness of the sample paths on any compact interval almost surely follows from (4.1) and (4.2). As for the làg property, for $x < y$ two reals, define

$$A = \bigcup_{n \in \mathbb{N}} \bigcup_{p,q \in \mathbb{Q}, p<q} \left\{\omega \in \Omega : U_{[0,n]}(p, q, X(\omega)) = \infty\right\}.$$

By means of the up-crossing inequality, it follows that this countable union is of measure 0. However, A contains the set

$$\left\{\omega \in \Omega : \liminf_{s \nearrow t} X_s(\omega) < \limsup_{s \nearrow t} X_s(\omega), t \in \mathbf{T}\right\}.$$

Hence X is càdlàg. The fact that X is a supermartingale with respect to $\mathbb{F}^+$ is immediate. As for the augmentation, observe that null sets do not modify the supermartingale inequalities. $\square$

As noticed, the up-crossing inequality shows that supermartingales have some nice regularity of paths. However, we assumed from the beginning that these supermartingale were right-continuous, central to derive Doob's maximal inequalities. Let us show that up to modification, any supermartingale has nice properties, however in the right-continuous filtration or in a filtration satisfying the usual conditions. From now on, $\mathbf{T}$ is a continuous time interval and $\mathbb{Q}$ is a countable order dense subset of it.

Theorem 4.21. *Let X be a submartingale. Then the following holds true.*

(a) Almost surely, the limits

$$X_{t+} = \lim_{q \searrow t} X_q \quad \text{and} \quad X_{t-} = \lim_{q \nearrow t} X_q$$

exist for every $t \in \mathbf{T}$ and thereby define two processes X_+ and X_-,respectively.

(b) The process X_+ is a $\mathbb{F}^+$ submartingale and is a martingale if X is. Analogously, the process X_- is a $\mathbb{F}^-$ submartingale and is a martingale if X is. Furthermore

$$X_t \geq E\left[X_{t+} \,|\, \mathcal{F}_t\right] \tag{4.3}$$

$$X_{t-} \geq E\left[X_t \,|\, \mathcal{F}_{t-}\right] \tag{4.4}$$

with equality in (4.3) if $t \mapsto E[X_t]$ is right-continuous and equality in (4.4) if $t \mapsto E[X_t]$ is left-continuous. In particular, equality holds in (4.3) and (4.4) if X is a martingale.

Proof. (a) Unlike in the previous proof we can only estimate the up-crossing of X over a countable bounded interval. Define

$$A = \bigcup_{n \in \mathbb{N}} \bigcup_{p<q, p,q \in \mathbb{Q}} \left\{\omega \in \Omega : U_{[0,n] \cap \mathbb{Q}}(p, q, X(\omega)) = \infty\right\}.$$

This set is of measure 0. Hence with the same argumentation as in the previous proof, it follows that

$$P\left[\liminf_{q \nearrow t, q \in \mathbb{Q}} X_q < \limsup_{q \nearrow t, q \in \mathbb{Q}} X_q : \text{for some } t \in \mathbf{T}\right] = 0,$$

$$P\left[\liminf_{q \searrow t, q \in \mathbb{Q}} X_q < \limsup_{q \searrow t, q \in \mathbb{Q}} X_q : \text{for some } t \in \mathbf{T}\right] = 0.$$

We can then define the processes X_- and X_+ by

$$X_{t+} = \lim_{q \searrow t} X_q, \text{ for } t < T \quad \text{ and } \quad X_{t-} = \lim_{q \nearrow t} X_q, \text{ for } t > 0,$$

with the conventions that $X_{0-} = X_0$ and $X_{T+} = X_T$ if $T \in \mathbf{T}$.

(b) Clearly X_+ and X_- are $\mathbb{F}^+$- and $\mathbb{F}^-$-adapted processes, respectively. Let $(q_n) \subseteq \mathbb{Q}$ be a sequence decreasing to $t \in \mathbf{T}$. From the previous step, X_{q_n} converges P-almost surely to X_{t+}. Further, $E[X_t] \leq E[X_{q_n}] \leq E[X_{q_0}]$ for every n, so (X_{q_n}) is uniformly bounded in L^1, and $E[X_{q_n}]$ is a decreasing sequence converging to $\lim E[X_{q_n}] > E[X_t] > -\infty$. Hence, for $\lambda > 0$, and $\varepsilon > 0$, let n_0 be such that $E[X_{q_n}] \geq E[X_{q_{n_0}}] - \varepsilon$ for every $n \geq n_0$. As X is a submartingale, it follows that

$$\begin{aligned}
E\left[|X_{q_n}| : |X_{q_n}| > \lambda\right] &= E\left[X_{q_n} : X_{q_n} > \lambda\right] - E\left[X_{q_n} : X_{q_n} < -\lambda\right] \\
&= E\left[X_{q_n} : X_{q_n} > \lambda\right] - E\left[X_{q_n}\right] + E\left[X_{q_n} : X_{q_n} \geq -\lambda\right] \\
&\leq E\left[X_{q_{n_0}} : X_{q_n} > \lambda\right] + \varepsilon - E\left[X_{q_{n_0}}\right] + E\left[X_{q_{n_0}} : X_{q_n} \geq -\lambda\right] \\
&= E\left[X_{q_{n_0}} : X_{q_n} > \lambda\right] - E\left[X_{q_{n_0}} : X_{q_n} < -\lambda\right] + \varepsilon \\
&\leq E\left[\left|X_{q_{n_0}}\right| : |X_{q_n}| > \lambda\right] + \varepsilon.
\end{aligned}$$

By Markov's inequality, $P[|X_{q_n}| > \lambda] \leq \sup_n E[|X_{q_n}|]/\lambda = C/\lambda$ for $0 < C < \infty$, showing therefore that (X_{q_n}) is uniformly integrable. Together with the P-almost sure convergence, it follows that X_{q_n} converges in L^1 to X_{t+}. Thus X_{t+} is integrable and it holds

$$X_t \leq \lim E\left[X_{q_n} \,|\, \mathcal{F}_t\right] = E\left[X_{t+} \,\middle|\, \mathcal{F}_t\right].$$

Further, for $s < t$, and $q_n \searrow s$ with $q_n < t$, it holds

$$X_{q_n} \leq E\left[X_t \,|\, \mathcal{F}_{q_n}\right] \leq E\left[E\left[X_{t+} \,|\, \mathcal{F}_t\right] \,|\, \mathcal{F}_{q_n}\right] = E\left[X_{t+} \,|\, \mathcal{F}_{q_n}\right]$$

for every n. The same arguments as above show that $E\left[X_{t+} \,|\, \mathcal{F}_{q_n}\right]$ is uniformly integrable and converges P-almost surely and in L^1 and that the limit is $E[X_{t+} \,|\, \mathcal{F}_{s+}]$. Thus X_+ is a $\mathbb{F}^+$-submartingale. Finally, if $t \mapsto E[X_t]$ is right-continuous, it follows that $E[X_{t+}] = \lim E[X_{q_n}] = E[X_t]$. Hence, the positive random variable $X_t - E\left[X_{t+} \,|\, \mathcal{F}_t\right]$ has zero expectation and therefore is zero.

As for the case of X_-, the a similar argumentation holds using submartingale convergence theorem for the existence and integrability of X_{t-} and inequality (4.4). Furthermore, by $X_{s-} \leq E[X_s \,|\, \mathcal{F}_{s-}] \leq E[E[X_{t-} \,|\, \mathcal{F}_s] \,|\, \mathcal{F}_{s-}] = E[X_{t-} \,|\, \mathcal{F}_{s-}]$ it follows that X is a $\mathbb{F}^-$ submartingale. The equality in (4.4) if $t \mapsto E[X_t]$ is left-continuous follows by an analogous argumentation. □

Theorem 4.22. *Let X be a supermartingale with respect to a filtration satisfying the usual assumptions. Suppose further that $t \mapsto E[X_t]$ is right-continuous. Then X has a càdlàg modification.*

Proof. According to the previous theorem, set $Y = X_+$ outside the negligible set A up to which X_+ and X_- are defined, and 0 on A. Since $A \in \mathcal{F}_0$, it follows that Y is càdlàg. Furthermore, from $t \mapsto E[X_t]$ right-continuous, by the previous theorem it holds $X_t = E[X_{t+} \,|\, \mathcal{F}_t] = E[Y_t \,|\, \mathcal{F}_t]$. However, since $\mathbb{F}$ is right-continuous, it follows that Y_t is $\mathcal{F}_t$-measurable and so $X_t = Y_t$ almost surely for every t. □

4.2. Integration

Recall that we assume that the filtrations on continuous time interval satisfy the usual conditions.

4.2.1. Lebesgue-Stieljes Integration

Proposition 4.23 (Lebesgue-Stieljes measure). *Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing right continuous function. There exists a unique measure dF on the Borel σ-algebra of the real line such that*

$$dF[]a, b]] = F(b) - F(a), \quad a, b \in \mathbb{R}, a \leq b$$

This measure is called the Lebesgue-Stieljes measure.

Proof. Let $\Omega = \mathbb{R}$ and $\mathcal{B}$ be the Borel σ-algebra which is generated by the semi-ring $\mathcal{R} = \{]a, b] \colon a \leq b\}$ with the convention that $]a, a] = \emptyset$. Define

$$P[]a, b]] = F(a) - F(b), \quad]a, b] \in \mathcal{R}.$$

Straightforward inspection shows that P is additive, such that $P[\emptyset] = 0$ and P is sub-additive. To show that P extends uniquely to a σ-finite measure on $\mathcal{B}$, we just have to check that P is σ-subadditive on $\mathcal{R}$. Let $A =]a, b] \in \mathcal{R}$ and $(A_n) = (]a_n, b_n])$ a countable family in $\mathcal{R}$ such that $A \subseteq \cup A_n$. Taking $\varepsilon > 0$, by right-continuity of F, choose some $a^\varepsilon \in]a, b[$ such that $F(a^\varepsilon) - F(a) < \varepsilon/2$. Also using the right-continuity of F, choose $b_n^\varepsilon > b_n$ for every n such that $F(b_n^\varepsilon) - F(b_n) \leq \varepsilon 2^{-n-1}$. It follows that

$$[a^\varepsilon, b] \subseteq]a^\varepsilon, b] \subseteq]a, b] \subseteq \cup]a_n, b_n] \subseteq \cup]a_n, b_n^\varepsilon[.$$

However, $[a^\varepsilon, b]$ is a compact set, therefore, the open covering $\cup]a_n, b_n^\varepsilon[$ of $[a^\varepsilon, b]$ can be choosen finite, hence, there exists n_0 such that

$$[a^\varepsilon, b] \subseteq]a^\varepsilon, b] \subseteq \cup_{k\leq n_0}]a_k, b_k^\varepsilon[\subseteq \cup_{k\leq n_0}]a_k, b_k^\varepsilon]$$

and therefore[45]

$$\begin{aligned} P[[a,b]] = F(b) - F(a) \leq \varepsilon/2 + F(b) - F(a^\varepsilon) \leq \varepsilon/2 + \sum_{k=1}^{n_0}(F(b_k^\varepsilon) - F(a_k)) \\ \leq \varepsilon + \sum (F(b_n) - F(a_n)) = \varepsilon + \sum P[]a_n, b_n]] \end{aligned}$$

showing that P extends to a measure on the real line. This measure is also σ-finite in the sense that there exists an increasing sequence of sets $(]a_n, b_n])$ such that $\mathbb{R} = \cup]a_n, b_n]$ and $P[]a_n, b_n]) < \infty$ for every n. Hence, this extension is unique and we denote it dF. □

Example 4.24. 1: The classical Lebesgue measure dx on the real line is derived from the continuous increasing function $F(x) = x$, for which it holds

$$dx[]a,b]] = b - a, \quad a \leq b$$

2: if we consider the function $F(x) = 1_{[y,\infty[}(x)$ for $y \in \mathbb{R}$ which is increasing, it gives rise to the Dirac measure

$$dF[A] = \delta_y[A] = \begin{cases} 1 & \text{if } y \in A \\ 0 & \text{otherwise} \end{cases}$$ ◊

We will mainly be interested in integrating functions with respect to their paths which are not necessarily increasing.

Definition 4.25. We say that a function[46] $F : [0, \infty[, t \mapsto F(t) := F_t$ is of bounded variation if

$$S_t = \sup_{\Pi \text{ subdivision of } [0,t]} S_t^\Pi < \infty$$

for every $t \geq 0$ where

$$S_t^\Pi = \sum_{1\leq k\leq n} |F_{t_{k+1}} - F_{t_k}|, \quad \Pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}$$

Functions of bounded variations actually describe any functions that can be defined as a difference between two increasing functions.

Proposition 4.26. *A function is of bounded variations if and only if it can be written as the difference between to increasing functions.*

Proof. Let F be a function of bounded variations. Inspection shows that $F^+ = (S + F)/2$ and $F^- = (S - F)/2$ are two increasing functions which difference is equal to F. This decomposition is actually the minimal one, in the sense that if $F = A - B$ for two increasing functions A and B, then it follows that $F^+ \leq A$ and $F^- \leq B$. The reciprocal is easy. □

[45] As an exercise, check that μ is finitely sub-additive. This is in general true for content on a semi-ring, see Appendix.

[46] This can also be defined on any interval of the real line, however we will mostly deal with integration with respect to time, and therefore consider the interval $[0, \infty[$.

Given a right continuous function F of bounded variations, we can therefore defined a so called signed measure and the absolute value of this measure

$$dF = dF^+ - dF^- \quad \text{and} \quad |dF| = dF^+ + dF^-$$

If $f : [0,\infty[\to \mathbb{R}$ is a locally bounded[47] and $\mathcal{B}([0,\infty[)$-measurable function on the real line, it follows that we can define the integral

$$\int_0^t f_s dF_s = \int_{]0,t]} f_s dF_s := \int_0^t f_s dF_s^+ - \int_0^t f_s dF_s^-$$

which is called the Stieljes integral of f with respect to F. The integral is understood over the interval $]0,t]$ so that $\int_0^t dF_s = F_t - F_0$.

Proposition 4.27 (Integration by part). *Let F and G be two right continuous functions of finite variations, then it holds*

$$F_t G_t = F_0 G_0 + \int_0^t F_s dG_s + \int_0^t G_{s-} dF_s = F_0 G_0 + \int_0^t F_{s-} dG_s + \int_0^t G_{s-} dF_s + \sum_{s \leq t} \Delta F_s \Delta G_s$$

where $F_{s-} = \lim_{u \nearrow s} F_u$ and $\Delta F_s = F_s - F_{s-} = dF[\{s\}]$.

Proof. Considering the product measure $dF \otimes dG$ on the product space $[0,\infty[\times[0,\infty[$, using the triangular equality

$$1_{]0,t]}(s_1)1_{]0,t]}(s_2) = 1_{]0,t]}(s_1)1_{]0,s_1]}(s_2) + 1_{]0,s_2[}(s_1)1_{]0,t]}(s_2)$$

we obtain using Fubini-Tonelli that

$$\begin{aligned} dF \otimes dG\,[]0,t]\times]0,t]] = (F_t - F_0)(G_t - G_0) &= \int\int 1_{]0,t]}(s_1)1_{]0,t]}(s_2) dF_{s_1} dG_{s_2} \\ &= \int_{]0,t]} \left(\int_{]0,s_1]} dG_{s_2} \right) dF_{s_1} + \int_{]0,t]} \left(\int_{]0,s_2[} dF_{s_1} \right) dG_{s_2} \\ &= \int_{]0,t]} G_s dF_s + \int_{]0,t]} F_{s-} dG_s - G_0 \left(F_t - F_0\right) - F_0 \left(G_t - G_0\right) \end{aligned}$$

showing the first side of the equality. As for the second one, noting that $F = F_- + \Delta F$, and since F can only have countably many discontinuity points, it holds

$$\int_{]0,t]} F_s dG_s = \int_{]0,t]} F_{s-} dG_s + \int_{]0,t]} \Delta F_s dG_s = \int_{]0,t]} F_{s-} dG_s + \sum_{s \leq t} \Delta F_s \Delta G_s$$

□

Proposition 4.28 (Chain Rule Formula). *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable function. It follows that*

$$f(F_t) = f(F_0) + \int_0^t f'(F_{s-}) dF_s + \sum_{s \leq t} \left(f(F_s) - f(F_{s-}) - f'(F_{s-}) \Delta F_s \right)$$

In particular, if F is continuous, it holds

$$f(F_t) = f(F_0) + \int_0^t f'(F_s) dF_s$$

[47] That is bounded on any compact interval $[0,t]$.

Proof. The proposition holds for $f(x) = x$, and by the integration by part Proposition 4.27, it also holds for $f(x) = x^2$. Hence, by recursion, it holds for every monomial of the form $f(x) = x^n$ where n is an integer. The chain rule formula being linear, it holds for every polynomial $f(x) = \sum_{k\leq n} \alpha_k x^k$. Since we can approximate uniformly every continuously differentiable function by interpolation on any compact interval, the formula also holds in the limit for every continuously differentiable function. □

We can extend this integration procedure for every ω-dependent paths as follows.

Definition 4.29. A process A is called *increasing* if

- A is adapted;
- $E[A_t] < \infty$ for all t and $A_0 = 0$;
- A is càdlàg and almost all sample paths are increasing[48].

An increasing process A is called integrable if $E[A_T] < \infty$ where $A_T = \lim_t A_t$.
A process A is called of *bounded variations* if A is the difference between two increasing processes.

We denote by dA the ω-wise σ-finite signed measure $dA_t(\omega)$ induced by A. For every locally bounded[49] measurable process[50] X, we can define the ω-wise integral

$$\int_0^t X_s(\omega) dA_s(\omega), \quad \omega \in \Omega, t \in [0, \infty[$$

Proposition 4.30. *If X is a locally bounded measurable process and A is a process of bounded variations, it follows that*

$$\int X_s dA_s = \left(\int_0^t X_s dA_s \right)$$

defines right continuous measurable process. If furthermore X is progressive, then $\int X dA_s$ is progressive and right continuous.

Proof. The proof is quite easy, and only needs to approximate X be sequences of simple step processes with the right measurability. □

Exercice 4.31. Using Radon-Nikodym, show that for A, B two increasing processes such that $A - B$ is still an increasing process, there exists an adapted jointly measurable process H such that $B = H \bullet A$. In particular, if A is of bounded variations, there exists H adapted and measurable such that $A = \int H\, |dA|$.◊

Definition 4.32. An increasing process A is called natural if for every bounded right-continuous martingale M

$$E\left[\int_0^t M_s dA_s\right] = E\left[\int_0^t M_{s-} dA_s\right], \quad \text{for all } t \in \mathbf{T}. \tag{4.5}$$

Remark 4.33. Note that every increasing and continuous process is automatically natural. Indeed

$$\int_0^t (M_s - M_{s-})\, dA_s = 0$$

almost surely since every paths $s \mapsto M_s$ has only countably many discountinuities and therefore is a set of dA null measure. ♦

[48] The natural – and only – way of being increasing means that $s \leq t$ implies $A_s \leq A_t$!
[49] That is $(\omega, s) \mapsto X_s(\omega)$ is uniformly bounded for every $\omega \in \Omega$ and $s \in [0, t]$.
[50] Recall that a measurable process is a process such that $(\omega, t) \mapsto X_t(\omega)$ is $\mathcal{F} \otimes \mathcal{B}([0, \infty[)$, in particular $t \mapsto X_t(\omega)$ is $\mathcal{B}([0, \infty[)$ measurable for every ω.

Lemma 4.34. *Note that (4.5) is equivalent to*

$$E\left[M_tA_t\right]=E\left[\int_0^t M_{s-}dA_s\right],\quad \textit{for all } t\in\mathbf{T}. \tag{4.6}$$

Proof. It suffices to show that $E[M_tA_t]=\int_0^t M_s dA_s$. Let $0=t_0<t_1<\cdots<t_n=t$ and define $M^n=\sum_{k=1}^n M_{t_k}1_{]t_{k-1},t_k]}$. By the Martingale property, it follows that

$$\begin{aligned}E\left[\int_0^t M^n_s dA_s\right]&=\sum_{k=1}^n E\left[M_{t_k}\left(A_{t_k}-A_{t_{k-1}}\right)\right]\\&=E\left[M_tA_t\right]-\sum_{k=1}^{n-1}E\left[A_{t_k}\left(M_{t_{k+1}}-M_{t_k}\right)\right]=E\left[M_tA_t\right].\end{aligned}$$

By letting the mesh of the subdivision tending to 0, it holds $M^n\to M$ P-almost surely, and by dominated convergence, it follows that

$$E\left[M_tA_t\right]=E\left[\int_0^t M_s dA_s\right]$$

which ends the proof. □

Lemma 4.35. *If $\mathbf{T}$ is discrete, an increasing process is natural, if and only if it is predictable and integrable.*

Proof. If an increasing process is predictable and integrable, then clearly it is natural. Reciprocally, define the bounded positive martingale $M_s=E[1_B\,|\,\mathcal{F}_s]$ where $B=\{A_t-E[A_t\,|\,\mathcal{F}_{t-1}]>\varepsilon\}$. It follows that

$$\begin{aligned}0=E\left[(M_t-M_{t-1})A_t\right]&=E[1_BA_t]-E[E[1_B\,|\,\mathcal{F}_{t-1}]A_t]\\&\geq\varepsilon P[B]+E\left[1_BE[A_t\,|\,\mathcal{F}_{t-1}]\right]-E\left[1_BE[A_t\,|\,\mathcal{F}_{t-1}]\right]=\varepsilon P[B].\end{aligned}$$

Hence $P[B]=0$. The same holds for $B=\{A_t-E[A_t\,|\,\mathcal{F}_{t-1}]<-\varepsilon\}$ showing that $A_t=E[A_t\,|\,\mathcal{F}_{t-1}]$□

Proposition 4.36. *Let M be a right continuous martingale which can be written as the difference between two natural increasing processes. Then M is indistinguishable from 0.*

Proof. By (4.6), $E[X_tM_t]=E[\int_0^t X_{s-}dM_s]$ for every càdlàg bounded martingale X. Now take $X^n=\sum_{k=1}^n X_{t_{k-1}}1_{[t_{k-1},t_k[}+X_t1_{\{t\}}$ for the subdivision $0=t_0<t_1<\ldots<t_n=t$ for which holds $X^n_-=X_01_{\{0\}}+\sum_{k=1}^n X_{t_{k-1}}1_{]t_{k-1},t_k]}$. Hence, by the Martingale property of M it holds

$$E\left[\int_0^t X^n_{s-}dM_s\right]=E\left[\sum_{k=1}^n X_{t_{k-1}}\left(M_{t_k}-M_{t_{k-1}}\right)\right]=0.$$

By letting the mesh of the subdivision converging to 0, it follows by dominated convergence that

$$E\left[X_tM_t\right]=E\left[\int_0^t X_{s-}dA_s\right]=0$$

Let $X=E[1_A\,|\,\mathcal{F}_\cdot]$ where $A=\{M_t>\varepsilon\}$. Hence X is a bounded martingale, and therefore, up to a modification it is càdlàg. It follows that $0=E[X_tM_t]=E[1_AM_t]\geq\varepsilon P[A]$ and so $P[A]=0$. The same holds for $A=\{M_t<-\varepsilon\}$ showing that $M_t=0$. From M and 0 càdlàg follows indistinguishability. □

4.2.2. Doob-Meyer Decomposition

Definition 4.37. A right-continuous process X is said to belong to class

- (D) if the collection $\{X_\tau : \tau \text{ stopping time}, \tau < \infty\}$ is uniformly integrable;
- (DL) if the collections $\{X_\tau : \tau \text{ stopping time}, \tau < t\}$ is uniformly integrable for every $t \in \mathbf{T}$.

Theorem 4.38. *Let X be a càdlàg supermartingale of class (DL). Then X can be decomposed uniquely – up to indistinguishability – into*

$$X = M - A \tag{4.7}$$

where M is a càdlàg martingale and A is a natural increasing process. If furthermore, X is of class (D), then M is uniformly integrable and A is integrable.

For the proof of the Theorem which follows Rao [6] we will make use of the following Theorem of Dunford-Pettis, the proof of which can be found in Dunford and Schwartz [2] for instance.

Theorem 4.39. *A subset of L^1 is $\sigma(L^1, L^\infty)$-relatively compact if and only if it is uniformly integrable.*

Proof. Let us first address the uniqueness by setting $X = M - A = M' - A'$. It follows that $A - A' = M - M'$ is a càdlàg martingale which is a difference of two increasing natural processes. It follows from the previous Proposition 4.36 that $A - A' = M - M'$ is indistinguishable from 0, hence the uniqueness. We fix some $a \in \mathbf{T}$ and prove the existence for $t \in [0, a]$, the uniqueness yields then the result on $\mathbf{T}$. Without loss of generality, we may assume that $X_0 = 0$ and define $\Pi^n = \{t_k^n = ka/2^n : k = 0, \ldots, 2^n\}$ subdivision of $[0, a]$ and $\Pi = \cup \Pi^n$. For any process Y we define $\Delta_k^n Y = Y_{t_{k+1}^n} - Y_{t_k^n}$ the difference process along Π^n. Let then

$$A_t^n = \sum_{k=0, k<2^n t/a}^{n} -E\left[\Delta_k^n X \,|\, \mathcal{F}_{t_k^n}\right]$$

Since X is a supermartingale, it follows that A^n is an increasing process and by definition $M^n = X + A^n$ is a – discrete time – martingale on Π^n. Indeed,

$$\begin{aligned}
E\left[\Delta_k^n M^n \,|\, \mathcal{F}_{t_k^n}\right] &= E\left[\Delta_k^n X \,|\, \mathcal{F}_{t_k^n}\right] + \sum_{\tilde{k}=0}^{k-1} E\left[\Delta_{\tilde{k}}^n X \,|\, \mathcal{F}_{t_{\tilde{k}}^n}\right] - \sum_{\tilde{k}=0}^{k} E\left[\Delta_{\tilde{k}}^n X \,|\, \mathcal{F}_{t_{\tilde{k}}^n}\right] \\
&= E\left[\Delta_k^n X - \Delta_k^n X \,|\, \mathcal{F}_{t_k^n}\right] = 0.
\end{aligned}$$

Let now $\tau_\lambda^n = \inf\{t \in \mathbf{T} : A_t^n > \lambda\} \wedge a$ which is a stopping time bounded by a. It follows from the optional sampling theorem that for every $t \in [0, a]$ it holds

$$\begin{aligned}
\frac{1}{2} E\left[A_t^n : A_t^n > 2\lambda\right] &\leq E\left[A_t^n - A_t^n \wedge n\right] = E\left[A_t^n - A_{\tau_\lambda^n \wedge t}^n\right] \\
&= E\left[M_t^n - M_{\tau_\lambda^n \wedge t}^n\right] - E\left[X_t - X_{\tau_\lambda^n \wedge t}\right] = E\left[X_{\tau_\lambda^n \wedge t} - X_t\right] = E\left[X_{\tau_\lambda^n \wedge t} - X_t : A_t^n > \lambda\right].
\end{aligned}$$

However, by Markov's inequality and $\{X_t : t \leq a\}$ uniformly integrable, it holds $P[A_t^n > \lambda] \leq E\left[A_t^n\right]/\lambda = E[X_t]/\lambda \leq \sup_{t \leq a} E[|X_t|]/\lambda$. Hence, $P[A_t^n > \lambda] \to 0$ uniformly in n and $t \in [0, a]$ as λ goes to ∞. From the uniform integrability of $\{X_t - X_{\tau_\lambda^n \wedge t} : n, \lambda > 0, t \in [0, a]\}$, it follows that $\{A_t^n : n \in \mathbb{N}\}$ is uniformly integrable. This follows in particular if $t = a$. If X is of class (D), then we can remove from this argumentation the boundary a and modify with $\Pi^n = \{k/2^n : k \in \mathbb{N}\}$ and find that $\{A_T^n : n \in \mathbb{N}\}$ is uniformly integrable.
By Dunford-Pettis'Theorem 4.39, there exists a subsequence, again denoted A^n, and $A_a \in L^1$ such that $E[\xi A_a^n] \to E[\xi A_a]$ for every $\xi \in L^\infty$. Define $M = E[X_a - A_a \,|\, \mathcal{F}_\cdot]$, which being a martingale can be

selected up to a modification as being càdlàg. Let further $A = M - X^a$ where X^a is X stopped at a. By the very choice of M, A is also càdlàg, and it holds $X = M - A$ on $[0, a]$. Let $t \in \Pi$ and n large enough that $t \in \Pi^n$. For every $\xi \in L^\infty$, it holds

$$E\left[\xi\left(A_t^n - A_t\right)\right] = E\left[E[\xi \,|\, \mathcal{F}_t]\left(M_t^n - M_t\right)\right] = E\left[E\left[\xi \,|\, \mathcal{F}_t\right]\left(M_a^n - M_a\right)\right] = E\left[E\left[\xi \,|\, \mathcal{F}_t\right]\left(A_a^n - A_a\right)\right],$$

which converges to 0 showing that $A_t^n - A_t$ converges in $\sigma(L^1, L^\infty)$ to 0 for every $t \in \Pi$. In particular, for every $t, s \in \Pi$, $s < t$ and n large enough,

$$0 \leq E\left[A_t^n - A_s^n : A_t < A_s\right] \xrightarrow[n\to\infty]{} E\left[(A_t - A_s) \wedge 0\right] \leq 0$$

showing that A is increasing along Π and A being càdlàg it follows that A is an increasing process. We are left to show that A is natural on $[0, a]$. For N bounded càdlàg martingale, using the fact that $A^n = A + M^n - M$, $M^n - M$ is a martingale along Π^n and Lemma 4.34 it holds

$$\begin{aligned} E[N_a A_a^n] &= E\left[\int_0^a N_{s-} dA_s^n\right] = \sum E\left[N_{t_k^n} \Delta_k^n A^n\right] \\ &= \sum E\left[N_{t_k^n} \Delta_k^n A\right] + \sum E\left[N_{t_k^n} \Delta_k^n \left(M^n - M\right)\right] = E\left[\sum N_{t_k^n} \Delta_k^n A\right]. \end{aligned}$$

On the left-hand side use the weak-convergence, and on the right-hand side the dominated convergence to get that $E[N_a A_a] = E[\int_0^a N_{s-} dA_s]$ which by Lemma 4.34 ends the proof for the case (DL).
If X is of class D, it follows from the argumentation about the uniform integrability that $\{A_T^n : n \in \mathbb{N}\}$ is uniformly integrable, hence $A_T \in L^1$. Furthermore, from Proposition (thm:supermartingaleconvergence (the L1 convergence theorems for u.i. cases has to be added!) it follows that $\sup E[|X_t|] < \infty$. Hence,

$$E[M_t : M_t > \lambda] = E\left[X_t + A_t : X_t + A_t > \lambda\right] \leq E\left[X_t + A_T : X_t + A_t > \lambda\right].$$

From Markov's inequality, it follows that $P[X_t + A_t > \lambda] \leq (E\left[|X_t|\right] + E\left[A_T\right])/\lambda \leq \sup E\left[|X_t|\right]/\lambda + E\left[A_T\right]/\lambda$ converging uniformly in t to 0 as $\lambda \to \infty$. From the uniform integrability of $\{X_t + A_T : t \in \mathbf{T}\}$ follows the uniform integrability of M. □

In the following we will define the stochastic integral with respect to continuous martingale using the Doob-Meyer decomposition. So one may think that if X is continuous, then the Doob-Meyer decomposition is also continuous. This is not straightforward a-priori, and is the subject of the following theorem that will be addressed in an Appendix.

Theorem 4.40. *Let X be a càdlàg supermartingale such that $E[X_{\tau^n}] \to E[X_\tau]$ for every increasing sequence of stopping times (τ^n) with $\sup \tau^n = \tau < a$ where $a \in \mathbf{T}$. Then, the natural increasing process in the Doob-Meyer decomposition of X is continuous.*

4.2.3. Stochastic Integral

We will construct the stochastic integral with respect to continuous martingales, since this is the only case of integral we will meet in this lecture. Let us fix some notations. By $\mathcal{M}_c^2$ we denote the space of square integrable continuous martingales M such that $M_0 = 0$. On the space $\mathcal{M}_c^2$, we define the – translation invariant – distance

$$\|M\| = \sum \frac{1}{2^n} \frac{\sup_{t\leq n} \|M_t\|_2}{1 + \sup_{t\leq n} \|M_t\|_2}.$$

The fact that this is a translation invariant distance is classical: It is a Fréchet distance generated by a countable family of half norms.

Lemma 4.41. *The space* $(\mathcal{M}_c^2, \|\cdot\|)$*, where two elements are identified if they are indistinguishable, is a Fréchet space, that is, complete.*

Proof. Let (M^n) be a Cauchy sequence in $\mathcal{M}_c^2$. It follows that (M_t^n) is Cauchy in L_t^2, hence converges in L^2 to M_t. Let further $A \in \mathcal{F}_s$. It follows from the martingale property of M^n that $E[M_t : A] = \lim_n E[M_t^n : A] = \lim_n E[M_s^n : A] = E[M_s : A]$. Choosing the càdlàg version of M shows that M is a càdlàg martingale with $M_t \in L_t^2$ for every t. Let us prove that M is actually continuous. By Doob's Maximal inequality one has

$$P\left[\sup_{s\leq t} |M_s^n - M_s| > \lambda\right] \leq \frac{1}{\lambda}\|M_t^n - M_t\|_2^2 \xrightarrow[n\to\infty]{} 0.$$

Hence $P[\sup_{s\leq t} |M_s^{n_k} - M_s| > \lambda] \leq 1/2^k$ for some n_k and every k. Applying Borel-Cantelli, it follows that $P[\liminf\{\sup_{s\leq t} |M_s^n - M_s| \leq \lambda\}] = 1$ for every λ, showing that for P almost all $\omega \in \Omega$ the continuous path $M^n(\omega)$ converges uniformly on $[0, t]$ to $M(\omega)$, that is M is continuous. $\square$

Definition 4.42. According to Theorems 4.38 and 4.40, we denote by $\langle M\rangle$ the unique natural increasing and continuous process in the Doob-Meyer decomposition of the supermartingale $-M^2$ for $M \in \mathcal{M}_c^2$. We call $\langle M\rangle$ the *quadratic variation* of M.

By the Doob-Meyer theorem, $\langle M\rangle$ is the unique increasing and natural process such that $M^2 - \langle M\rangle$ a continuous martingale. In particular, it follows that $\langle M^\tau\rangle = \langle M\rangle^\tau$ for every bounded stopping time.
Before defining the stochastic integral we will see that we may approximate the quadratic variation by simple step functions. To do so, for $M \in \mathcal{M}_c^2$, let

$$V(M, \Pi) = \sum_{k=1}^{n}\left(M_{t_k} - M_{t_{k-1}}\right)^2, \quad t \in \mathbf{T}$$

where Π is a set of points $0 = t_0 \leq t_1 \leq \cdots \leq t_n = t$ for some $t \in \mathbf{T}$.

Proposition 4.43. *Let* $M \in \mathcal{M}_c^2$*. Then,* $V(M, \Pi^n)$ *converges in probability to* $\langle M\rangle_t$ *where* Π^n *is a sequence of subdivisions of* $[0, t]$ *the mesh of which converges to* 0*.*

Proof. Let $M \in \mathcal{M}_c^2$. Throughout, Π denotes generically a subdivision $0 = t_0 \leq t_1, \cdots, t_n = t$, $|\Pi| = \sup\{|t_k - t_{k-1}| : t_k \in \Pi, k \geq 0\}$ the mesh of the subdivision, and with a slight abuse, for a process Y, we use the notation $\Delta_k Y = \Delta_k^\Pi Y = Y_{t_k} - Y_{t_{k-1}}$. For a function f, we denote by $m(f, \Pi) = \sup\{|f(s) - f(u)| : |s - u| < |\Pi|, s, u \in [0, t]\}$. If f is continuous, $m(f, \Pi) \to 0$ as $|\Pi| \to 0$ since we are on a compact interval. Straightforward inspection shows

$$E\left[(M_t - M_s)^2 \mid \mathcal{F}_s\right] = E\left[M_t^2 - M_s^2 \mid \mathcal{F}_s\right] = E\left[\langle M\rangle_t - \langle M_s\rangle \mid \mathcal{F}_s\right]$$

where $s \leq t$. In particular, if M is bounded by K, it holds $E[\sum(\Delta_k M)^2] = E[M_t^2] \leq K^2$. Suppose first that M and $\langle M\rangle$ are bounded. Using the aforementioned inequalities, $2a^2 + 2b^2 - (a-b)^2 = (a+b)^2 \geq 0$, and the fact that $M^2 - \langle M\rangle$ is a martingale, it holds

$$\begin{aligned}
E\left[\left(V\left(M, \Pi^n\right) - \langle M\rangle_t\right)^2\right] &= E\left[\left(\sum\left(\Delta_k M\right)^2 - \left(\Delta_k\langle M\rangle\right)\right)^2\right] \\
&= E\left[\sum\left(\left(\Delta_k M\right)^2 - \left(\Delta_k\langle M\rangle\right)\right)^2\right] \leq 2\sum E\left[\left(\Delta_k M\right)^4\right] + 2\sum E\left[\left(\Delta_k\langle M\rangle\right)^2\right] \\
&\leq 2\sum E\left[m\left(M, \Pi\right)^2\left(\Delta_k M\right)^2\right] + 2\sum E\left[m\left(\langle M\rangle, \Pi\right)\langle M\rangle_t\right].
\end{aligned}$$

Since $m(\langle M\rangle, \Pi)\langle M\rangle_t \to 0$ almost surely as $|\Pi| \to 0$, the uniform boundedness of $\langle M\rangle$ in combination with dominated convergence yields that the second term on the right hand-side converges to 0 as $|\Pi| \to 0$. As for the first term, applying Cauchy-Schwartz yields

$$E\left[m(M,\Pi)^2 \sum (\Delta_k M)^2\right] \leq E\left[m(M,\Pi)^4\right]^{1/2} E\left[\left(\sum (\Delta_k M)^2\right)^2\right]^{1/2}.$$

By dominated convergence, $E[m(M,\Pi)^4]$ converges to 0 as the mesh of the subdivision goes to 0. On the other hand, denoting by K the bound on M, the aforementioned inequalities yield

$$\begin{aligned} E\left[\left(\sum (\Delta_k M)^2\right)^2\right] &= E\left[\sum (\Delta_k M)^4\right] + 2E\left[\sum_k \sum_{l>k} E\left[(\Delta_l M)^2 (\Delta_k M)^2 \mid \mathcal{F}_{t_k}\right]\right] \\ &\leq 2K^2 E\left[\sum (\Delta_k M)^2\right] + 2K^2 E\left[\sum (\Delta_k M)^2\right] \leq 6K^4, \end{aligned}$$

showing the convergence in L^2 if both M and $\langle M\rangle$ are bounded. Otherwise, let $\tau^n = \inf\{t : |M_t| > n \text{ or } \langle M\rangle_t > n\}$. It follows that

$$\{|V(M,\Pi) - \langle M\rangle_t| > \varepsilon\} \subseteq \{\tau^n \leq t\} \cup \left\{\left|V\left(M^{\tau^n}, \Pi\right) - \langle M^{\tau^n}\rangle\right| > \varepsilon\right\}.$$

As for the second set, since both M^{τ^n} and $\langle M^{\tau^n}\rangle$ are bounded, the probability of this set can be made arbitrarily small. As for the first set, since $\lim P[\tau^n > t] = 1$, its probability can also be made arbitrarily small for n large enough. $\square$

For a fixed $M \in \mathcal{M}_c^2$, we define the space $L^2(M) := L^2(P \otimes d\langle M \rangle)$ of measurable processes with finite Fréchet distance

$$\|H\| = \sum \frac{1}{2^n} \frac{\|H\|_{2,n}}{1 + \|H\|_{2,n}},$$

where

$$\|H\|_{2,t} = E\left[\int_0^t H_s^2 d\langle M \rangle_s\right]^{1/2}.$$

This is a subset of the measurable processes $L^0(P \otimes d\langle M \rangle)$ identified in the $P \otimes d\langle M \rangle$-almost everywhere sense. The fact that $L^2(P \otimes d\langle M \rangle)$ is a Fréchet lattice is classical.

Definition 4.44. Let $M \in \mathcal{M}_c^2$. By $\mathcal{L}^2(M)$ we denote the subspace of those $H \in L^2(M)$ which are progressive.

If there is no risk of confusion, we often drop the reference to M and use the notation $\mathcal{L}^2 := \mathcal{L}^2(M)$. Note that $\mathcal{L}^2$ is a closed subspace of L^2. Indeed, let $(H^n) \subseteq \mathcal{L}^2$ such that $\|H^n - H\| \to 0$. For each m, it follows that $\|H^n - H\|_{2,m} \to 0$ which is the $\|\cdot\|_2$ norm in the Banach space $L^2(\Omega \times [0,m], \mathcal{F}_t \otimes \mathcal{B}([0,m]), P \otimes d\langle M \rangle_{[0,m]})$. Hence, up to a subsequence, $H^n \to H$ $P \otimes d\langle M \rangle$ on $\Omega \times [0,m]$. Taking a diagonal subsequence along m, it follows that $H^n \to H$ $P \otimes d\langle M \rangle$. Since H^n is progressive, it follows that H is itself progressive and therefore element of $\mathcal{L}^2$.
As previously mentioned, continuous local martingales have unbounded variation unless they are constant, see for instance the Brownian Motion W. Consequently a path-wise definition in the sense of a Riemann-Stieltjes integration is not applicable here. However, making use of L^2-isometry arguments going back to Itô allows us to define a proper notion of the stochastic integral. Recall that $\mathcal{S}$ is the set of simple integrands, that is those càglàd stochastic processes of the form

$$H = H_0 + \sum_{k=1}^{n} H_k 1_{]\tau_{k-1}, \tau_k]}$$

where $0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_n < a$ is a sequence of stopping times $a \in \mathbf{T}$ and $H_k \in L^\infty_{\tau_{k-1}}$ for $k = 0, \ldots, n$. Note that $\mathcal{S}$ is a linear subspace of $\mathcal{L}^2(M)$ independently of M. Furthermore

Lemma 4.45. *Let $M \in \mathcal{M}_c^2$, then the space $\mathcal{S}$ is dense in $\mathcal{L}^2(M)$. In particular, for each $H \in \mathcal{L}^2(M)$ there exists a sequence $(H^n) \subseteq \mathcal{S}$ satisfying*

$$\lim_{n \to \infty} \|H - H^n\|_{2,t} = 0, \quad \text{for all } t \in \mathbf{T}.$$

Remark 4.46. This lemma is the key to extend the stochastic integral from $\mathcal{S}$ to $\mathcal{L}^2(M)$. In particular, the size of the closure of $\mathcal{S}$ depends on the "regularity" of $\langle M \rangle$ allowing for more or less integrands for the stochastic integral with respect to M. On the one hand, if $\langle M \rangle$ is absolutely continuous – for instance in the case of the Brownian motion – then it is even possible to define a stochastic integral with respect to integrands in $L^2(P \otimes d\langle M \rangle)$ which are "only" measurable and adapted and not necessarily progressive. On the other hand if M were an element of $\mathcal{M}^2$ – those càdlàg martingales such that $E[M_t^2] < \infty$ for all $t \in \mathbf{T}$ such as some class of Levy Processes – then it is also possible to define a stochastic integral but with respect to a smaller set of integrands, namely those predictable processes in $\mathcal{L}^2(M)$.[51] ♦

Proof. Let $H \in \mathcal{L}^2$.

[51] The processes measurable with respect to the filtration generated by the càglàd processes.

Step 1: Assume first that the paths of $\langle M\rangle$ are absolutely continuous almost surely. In particular, $P \otimes d\langle M\rangle \ll P \otimes dt$.

Fix m. Assume that H is continuous adapted and bounded, therefore, by proposition 4.9 progressive. It follows that

$$H_s^n = H_{(k)2^{-n}m}, \quad \text{for } km2^n < s \leq (k+1)m2^{-n} \tag{4.8}$$

is converging $P \otimes dt$ to X on $\Omega \times [0,m]$, and by Lebesgue's dominated convergence, it follows that $\|H^n - H\|_{2,m} \to 0$.

Assume now that H is progressively measurable and bounded and define $F_t(\omega) = \int_0^t H_t(\omega)dt$ with the assumption that $F_t(\omega) = 0$ for $t < 0$. It follows that F is a bounded adapted and continuous process on $[0,m]$, and so is $G_t^n(\omega) = n(F_t(\omega) - F_{t-1/m}(\omega))$. By the fundamental theorem of calculus, it follows that for almost all $\omega \in \Omega$, $G_t^n(\omega) \to H_t(\omega)$ for dt-almost all $t \in [0,m]$. Hence $\{(\omega,t) : \lim G_t^n(\omega) \neq H_t(\omega)\}$ is a $P \otimes dt$-null measure set and a-fortiori $P \otimes d\langle M\rangle$-null measure set. By Lebesgues'dominated convergence, we then have $\|G^n - H\|_{2,m} \to 0$. From what have been done before, since each G^m can be approximated by a sequence of $\mathcal{S}$, it follows that every bounded progressive process H can be approximated in $\|\|_{2,m}$ on $[0,m]$ by $(H^n) \subseteq \mathcal{S}$.

Let now H be a progressive process and define $G_t^n = H_t 1_{\{|H_t| \leq n\}}$ which is progressive and bounded. From $\|H^n - H\|_{2,m} \to 0$ and H^n itself being approximated by elements of $\mathcal{S}$ in $\|\|_{2,m}$ it follows that H is approximated in $\|\|_{2,m}$ by elements in $\mathcal{S}$. Now to get in the Fréchet-Distance, consider $H^{n,m}$ an approximation in $\|\|_{2,m}$ of H on $[0,m]$. Up to a quick subsequence, we may assume that $\|H^{n,n} - H\|_{2,n} \leq 2^n$. It follows that the diagonal process $H^{n,n}$ converges in the Frechet distance to H.

Step 2: Let us now address the case where $d\langle M\rangle$ is not absolutely continuous with respect to dt. Just as previously, it is enough to show that every H progressive, uniformly bounded by a constant C, and non zero only on $[0,m]$ can be approximated by elements of $\mathcal{S}$ in the $\|\|_{2,m}$-norm. The idea is to "tweak" a bit $d\langle M\rangle$ to a measure absolutely continuous with respect to dt by means of a time change. Since $\langle M\rangle_t + t$ is strictly increasing and continuous, there exists a strictly increasing and continuous inverse $T : \Omega \times \mathbf{T} \to \mathbf{T}$ such that $\langle M\rangle_{T_s(\omega)}(\omega) + T_s(\omega) = s$ for all $s \in \mathbf{T}$. In particular[52] $T(s) \leq s$ since $A_t + t \geq t$ and $\{T_s \leq t\} = \{A_t + t \geq s\} \in \mathcal{F}_t$ for all t. It follows that T_s is a bounded stopping time for every s. We then define the new filtration $\mathcal{G}_s := \mathcal{F}_{T_s}$ for $s \in \mathbf{T}$ and the process $G_s = H_{T_s}$ which is $\mathbb{G}$-adapted and measurable since H is $\mathbb{F}$ progressive. However, we do not know wether G is progressive. Applying Theorem 4.10, we can modify it and assume that it is progressive and adapted. From the previous argumentation, there exists a simple processes G^ε of the form (4.8) such that $E[\int_0^\lambda |G_t - G_t^\varepsilon|^2 ds] \leq \varepsilon/2$ for a given λ. However, since $E[\int G_s^2 ds] \leq E[\int 1_{T_s \leq m} H_{T_s}^2 ds] \leq CE[\langle M\rangle_m + m] < \infty$, we may choose λ large enough and G^ε zero outside $[0,\lambda]$ to get $E[\int |G_t^\varepsilon - Y_t|^2 ds] < \varepsilon$. Reversing the time clock, it follows that

$$H_t^\varepsilon = X_0 1_0 + \sum G_{s_k}^\varepsilon 1_{]T_{s_k}, T_{s_{k+1}}]}$$

which is a simple process since T_{s_k} is a bounded stopping time and $G_{s_k}^\varepsilon$ is $\mathcal{F}_{T_{s_k}}$-measurable. Hence, by definition

$$E\left[\int_0^m |H_t^\varepsilon - H_t|^2 d\langle M\rangle_t\right] \leq E\left[\int_0^m |H_t^\varepsilon - H_t|^2 (d\langle M\rangle_t + dt)\right] \leq E\left[\int |G_s^\varepsilon - G_s|^2 ds\right] \leq \varepsilon,$$

which ends the proof. □

[52] Recall that the general right-continuous inverse of an increasing function f is given by $f^{-1}(s) = \sup\{t : f(t) \leq s\}$. However, if f is strictly increasing and continuous, it follows that $f^{-1}(s) = \inf\{t : s \leq f(t)\}$. Furthermore, if $f(t) \geq t$, then $f^{-1}(s) \leq s$.

So far so good, up to the fact that we are in need of Theorem 4.10 in order to obtain the density when $d\langle A\rangle$ is not absolutely continuous with respect to dt. You can find the proof at the end of this section. We are now in place to define the stochastic integral.

Definition 4.47. For $H \in \mathcal{S}$ we define the (elementary) stochastic integral of H with respect to M as the process

$$I(H) := H_0 M_0 + \sum_{k=1}^{n} H_k \left(M^{\tau_k} - M^{\tau_{k-1}}\right) = \sum_{k=1}^{n} H_k \left(M^{\tau_k} - M^{\tau_{k-1}}\right).$$

Lemma 4.48. *Let $M \in \mathcal{M}_c^2$. For each $H \in \mathcal{S}$, $I(H) \in \mathcal{M}_c^2$ with quadratic variation*

$$\langle I(H)\rangle = \int H_s^2 d\langle M\rangle_s$$

and it holds

$$\|I_t(H)\|_2 = \|H\|_{2,t}, \quad \text{for every } t \in \mathbf{T}.$$

In particular, I_t is an isometrie between $\mathcal{S}$ and L_t^2.

Proof. Observe that for $t > \tau_n$, M begin a martingale and H predictable, it holds

$$\|I_t(H)\|_2^2 = E\left[\left(\sum_{k=1}^{n} H_k \left(M_{\tau_k} - M_{\tau_{k-1}}\right)\right)^2\right] = E\left[\sum_{k=1}^{n} H_k^2 \left(M_{\tau_k}^2 - M_{\tau_{k-1}}^2\right)\right].$$

Since $M^2 - \langle M\rangle$ is a martingale it follows that

$$\|I_t(H)\|_2^2 = E\left[\sum_{k=1}^{n} H_k^2 \left(\langle M\rangle_{\tau_k} - \langle M\rangle_{\tau_{k-1}}\right)\right] = \|H\|_{2,t}^2.$$ □

You should notice that H being predictable is essential for the proof above.

Theorem 4.49. *Let $M \in \mathcal{M}_c^2$. For all $H \in \mathcal{L}^2$, there exists a unique – up to indistinguishability – process $I(H) \in \mathcal{M}_c^2$ such that for every $(H^n) \subseteq \mathcal{S}$ with $\|H^n - H\| \to 0$ for every t if follows that $\|I(H^n) - I(H)\| \to 0$. Furthermore, its quadratic variation are given by*

$$\langle I(H)\rangle = \int H^2 d\langle M\rangle.$$

Proof. Let $H \in \mathcal{L}^2(M)$ and according to 4.45 pick $(H^n) \subseteq \mathcal{S}$ such that $\|H^n - H\|_{2,t} \to 0$ for every $t \in \mathbf{T}$. By Isometry, this shows that $(I(H^n))$ is a Cauchy sequence in the complete space $\mathcal{M}_c^2$, see Lemma 4.45. Denoting this limit by $I(H)$, it is independent of the choice of (H^n). Indeed, if you consider another sequence, considering the alternate sequence between both, it will is then Cauchy and converges to a limit which is common to both. □

We call this operator the *stochastic integral* of H with respect to M and denote it either $H \bullet M$ or

$$\int H dM$$

Let us collect first properties the proof of which are immediate consequences of the approximation method and thus omitted here.

Proposition 4.50. *Let $M \in \mathcal{M}_c^2$, $G, H \in \mathcal{L}^2$ as well as $\alpha \in \mathbb{R}$.*

(i) $(\alpha H + G) \bullet M = \alpha H \bullet M + G \bullet M$.

(ii) $(1_{[0,\tau]} H) \bullet M = H \bullet M^\tau = (H \bullet M)^\tau$.

For $M, N \in \mathcal{M}_c^2$, the covariation of M, N – which is a continuous process of bounded variation – is given by the Polar formula

$$\langle M, N \rangle = \frac{1}{4} \left(\langle M + N \rangle - \langle M - N \rangle \right).$$

The following relation holds

Proposition 4.51. *Let $M, N \in \mathcal{M}_c^2$ and $G \in \mathcal{L}^2(M)$, $H \in \mathcal{L}^2(N)$ then it holds*

$$\langle G \bullet M, H \bullet N \rangle = \int G d\langle M, H \bullet N \rangle = \int GH d\langle M, N \rangle$$

In particular for $G \in \mathcal{L}^2(M)$ and $H \in \mathcal{L}^2(G \bullet M)$, it holds $GH \in \mathcal{L}^(M)$ and the chain rule

$$\int H d\left(G \bullet M \right) = \int GH dM.$$

This relation is straightforward for $G, H \in \mathcal{S}$. The passage to the limit is left as an exercise by using the following proposition known as the Kunita-Watanabe inequality.

Proposition 4.52. *Let $M, N \in \mathcal{M}_c^2$ and $G \in \mathcal{L}^2(M)$, $H \in \mathcal{L}^2(N)$ then it holds*

$$\int_0^t |GH| \, d \left| \langle M, N \rangle \right| \leq \left(\int_0^t |G| \, d\langle M \rangle \right)^{1/2} \left(\int_0^t |H| \, d\langle N \rangle \right)^{1/2}$$

for every $t \in \mathbf{T}$.

Proof. From the fact that $|GH| = sign(GH)GH$ and that the density process $d\langle M, N \rangle / d \left| \langle M, N \rangle \right|$ takes value in $\{0, 1\}$ is is enough to show the inequality for the $\left| \int GH d\langle M, N \rangle \right|$ on the left hand side. Using the Cauchy Schwartz inequality for bilinear forms, it holds for simple increments that $|\Delta \langle M, N \rangle| \leq (\Delta \langle M \rangle)^{1/2} (\Delta \langle N \rangle)^{1/2}$. With simple integrands it holds

$$\begin{aligned} \left| \int_0^t GH d\langle M, N \rangle \right| &\leq \sum |G_k H_k| \left| \Delta_k \langle M, N \rangle \right| \leq \sum |G_k| \, |H_k| \left(\Delta_k \langle M \rangle \right)^{1/2} (\Delta_k \langle N \rangle)^{1/2} \\ &\leq \left(\sum |H|_k^2 \, \Delta_k \langle M \rangle \right)^{1/2} \left(\sum |G|_k^2 \, \Delta_k \langle N \rangle \right)^{1/2} = \left(\int_0^t |G| \, d\langle M \rangle \right)^{1/2} \left(\int_0^t |H| \, d\langle N \rangle \right)^{1/2}. \end{aligned}$$

The rest follows by passing to the limit. □

We conclude this section by stating and proving the famous Burkholder-Davis-Gundy inequalities for a martingale M. We provide two totally different approaches to prove the theorem inspired by Karatzas and Shreve [4] on the one hand and by Rogers and Williams [7] on the other.

Theorem 4.53. *Let M be a continuous martingale which, along with its quadratic variation process $\langle M \rangle$, is bounded. Then, for every $p > 0$ there exists universal constants c_p and C_p such that*

$$c_p E \left[\langle M \rangle_T^{\frac{p}{2}} \right] \leq E \left[(M_T^*)^p \right] \leq C_P \left[\langle M \rangle_T^{\frac{p}{2}} \right].$$

Proof. Version 1: Following Karatzas and Shreve [4]. We set $m := \frac{p}{2}$. Consider the process

$$Y_t = \delta + \varepsilon \langle M \rangle_t + M_t^2 = \delta + (1+\varepsilon)\langle M \rangle_t + 2\int_0^t M_u dM_u$$

where $\delta > 0$ and $\varepsilon \geq 0$ will be chosen appropriately later on. Applying Itô's formula to Y^m we obtain

$$Y_t^m = \delta^m + m(1+\varepsilon)\int_0^t Y_u^{m-1} d\langle M \rangle_u + 2m(m-1)\int_0^t Y_u^{m-2} M_u^2 d\langle M \rangle_u + 2m \int_0^t Y_u^{m-1} M_u dM_u.$$

As M is bounded and Y is bounded away from zero, we have $E[\int_0^T Y_u^{m-1} M_u dM_u] = 0$ and thus

$$E\left[Y_T^m\right] = \delta^m + m(1+\varepsilon)E\left[\int_0^T Y_u^{m-1} d\langle M \rangle_u\right] + 2m(m-1)E\left[\int_0^T Y_u^{m-2} M_u^2 d\langle M \rangle_u\right]. \quad (4.9)$$

Case 1: Upper bound if $0 < m \leq 1$. The last term on the right-hand side of (4.9) is nonpositive and hence, letting $\delta \downarrow 0$, we obtain by recalling that $0 < m \leq 1$

$$\begin{aligned} E\left[\left(\varepsilon\langle M \rangle_T + M_T^2\right)^m\right] &\leq m(1+\varepsilon)E\left[\int_0^T \left(\varepsilon\langle M \rangle_u + M_u^2\right)^{m-1} d\langle M \rangle_u\right] \\ &\leq m(1+\varepsilon)\varepsilon^{m-1}E\left[\int_0^T \langle M \rangle_u^{m-1} d\langle M \rangle_u\right] \\ &= (1+\varepsilon)\varepsilon^{m-1}E\left[\langle M \rangle_T^m\right]. \end{aligned} \quad (4.10)$$

Using concavity of the function $f(x) = x^m, x \geq 0$ for m under consideration, we get

$$2^{m-1}(x^m + y^m) \leq (x+y)^m \qquad \text{for all } x, y \geq 0. \quad (4.11)$$

Thus, (4.10) yields $\varepsilon^m E[\langle M \rangle_T^m] + E[|M_T|^{2m}] \leq (1+\varepsilon)(\varepsilon/2)^{m-1}E[\langle M \rangle_T^m]$ by which we deduce

$$E\left[|M_T|^{2m}\right] \leq \left[(1+\varepsilon)\left(\frac{\varepsilon}{2}\right)^{m-1}\right] E\left[\langle M \rangle_T^m\right]. \quad (4.12)$$

Case 2: Lower bound if $m > 1$. In this case, the last term in (4.9) is nonnegative and thus the direction of all inequalities in (4.10)-(4.12) is reversed. Consequently, we end up with

$$E\left[|M_T|^{2m}\right] \geq \left[(1+\varepsilon)\left(\frac{\varepsilon}{2}\right)^{m-1}\right] E\left[\langle M \rangle_T^m\right]. \quad (4.13)$$

Of course, ε has to be chosen accordingly to ensure positivity of the constant appearing in (4.12) and (4.13).

Case 3: Lower bound if $\frac{1}{2} < m \leq 1$. Let us consider (4.9) with $\varepsilon = 0$ and let $\delta \downarrow 0$. This yields

$$E\left[|M_T|^{2m}\right] = 2m(m-\frac{1}{2})E\left[\int_0^T |M_u|^{2(m-1)} d\langle M \rangle_u\right]. \quad (4.14)$$

Furthermore, (4.9) together with (4.11) and $\frac{1}{2} < m \leq 1$ gives

$$\begin{aligned} 2^{m-1}\left[\varepsilon^m E\left[\langle M \rangle_T^m\right] + \left(E\left[\delta + M_T^2\right]\right)^m\right] &\leq E\left[\left(\varepsilon\langle M \rangle_T + (\delta + M_T^2)\right)^m\right] \\ &\leq \delta^m + m(1+\varepsilon)E\left[\int_0^T (\delta + M_u^2)^{m-1} d\langle M \rangle_u\right]. \end{aligned}$$

Letting $\delta \downarrow 0$ implies

$$2^{m-1}\left[\varepsilon^m E\left[\langle M\rangle_T^m\right]+\left(E\left[\delta+M_T^2\right]\right)^m\right] \leq m(1+\varepsilon) E\left[\int_0^T\left|M_u\right|^{2(m-1)} d\langle M\rangle_u\right]. \tag{4.15}$$

Combining (4.14) and (4.15) provides the following lower bound

$$E\left[\left|M_T\right|^{2m}\right] \geq \varepsilon^m\left(\frac{(1+\varepsilon) 2^{1-m}}{2m-1}-1\right)^{-1} E\left[\langle M\rangle_T^m\right]. \tag{4.16}$$

Case 4: Upper bound if $m>1$. In this case, the inequality (4.15) is reversed, yielding

$$E\left[\left|M_T\right|^{2m}\right] \leq \varepsilon^m\left(\frac{(1+\varepsilon) 2^{1-m}}{2m-1}-1\right)^{-1} E\left[\langle M\rangle_T^m\right]. \tag{4.17}$$

Once again, care has to be taken to assure positivity of the constant appearing above. Thus, for $m>\frac{1}{2}$ we obtain by Doob's maximal inequality

$$\begin{aligned} B_m E\left[\langle M\rangle_T^m\right] &\leq E\left[\left|M_T\right|^{2m}\right] \leq E\left[\left(M_T^*\right)^{2m}\right] \\ &\leq\left(\frac{2m}{2m-1}\right)^{2m} E\left[\left|M_T\right|^{2m}\right] \leq K_m\left(\frac{2m}{2m-1}\right)^{2m} E\left[\langle M\rangle_T^m\right] \end{aligned}$$

where the constants B_m and K_m are determined by the pair of inequalities (4.13) and (4.16) and (4.12) and (4.17), respectively. For the case $0<m \leq \frac{1}{2}$ we refer the reader to [4, Theorem 3.28].

Version 2: Following Rogers and Williams [7]. We begin by proving an auxiliary result. Given a function F such that $F(0)=0$ and for every $\alpha>1$ it holds $K_F:=\sup_{x \geq 0} \frac{F(\alpha x)}{F(x)}>\infty$, consider two nonnegative random variables X and Y and constants $\beta>1$ and $\delta, \varepsilon>0$ such that for all $\lambda>0$

$$P(X>\beta \lambda, Y \leq \delta \lambda) \leq \varepsilon P(X>\lambda). \tag{4.18}$$

Let γ and η be such that $F(\beta \lambda) \leq \gamma F(\lambda)$ and $F(\delta^{-1} \lambda) \leq \eta F(\lambda)$. If $\gamma \varepsilon<1$, then

$$E[F(X)] \leq \frac{\gamma \eta}{1-\gamma \varepsilon} E[F(Y)].$$

As to the proof of this claim, observe that (4.18) implies for $\lambda>0$

$$P(X>\beta \lambda)=P(X>\beta \lambda, Y \leq \delta \lambda)+P(X>\beta \lambda, Y>\delta \lambda) \leq \varepsilon P(X>\lambda)+P(Y>\delta \lambda). \tag{4.19}$$

As $F(0)=0$, it follows $F(x)=\int_0^x dF(\lambda)=\int_0^{\infty} 1_{\{\lambda<x\}} dF(\lambda)$. Thus, by Fubini's theorem it holds

$$E[F(X)]=\int_0^{\infty} P(X>\lambda) dF(\lambda)$$

Using (4.19) we obtain

$$\begin{aligned} E[F(X/\beta)] &= \int_0^{\infty} P(X>\beta \lambda) dF(\lambda) \\ &\leq \int_0^{\infty} \varepsilon P(X>\lambda) dF(\lambda)+\int_0^{\infty} P(Y>\delta \lambda) dF(\lambda) \leq \varepsilon E[F(X)]+E[F(Y/\delta)]. \end{aligned}$$

From the conditions on η and γ it follows that

$$E[F(X/\beta)] \leq \varepsilon\gamma E[F(X/\beta)] + \eta E[F(Y)]$$

As we assumed $\varepsilon\gamma < 1$, by using the condition on γ we deduce

$$E[F(X)] \leq \gamma E[F(X/\beta)] \leq \frac{\gamma\eta}{1-\varepsilon\gamma} E[F(Y)]$$

which is the desired result. We are now in the position to prove the Burkholder-Davis-Gundy inequality. To this end, let us define the stopping time $\tau := \inf\{t \geq 0 : |M_t| > \lambda\}$ and the process $N_t := (M_{t+\tau} - M_\tau)^2 - (\langle M\rangle_{t+\tau} - \langle M\rangle_\tau)$ which is an $\mathcal{F}_{\tau+t}$-adapted local martingale. Choose $\beta > 1$ and $0 < \delta < 1$. On the event $\{\sup_{t\geq 0}|M_t| > \beta\lambda, \langle M\rangle_\infty \leq \delta^2\lambda^2\}$ the martingale N must hit the level $(\beta-1)^2\lambda^2 - \delta^2\lambda^2$ before it hits $-\delta^2\lambda^2$. Using the optional sampling theorem it is an easy exercise to verify that the probability of a martingale hitting a level b before a level a is given by $-a/(b-a)$ and thus

$$P\left(\sup_{t\geq 0}|M_t| > \beta\lambda, \langle M\rangle_\infty \leq \delta^2\lambda^2 \,|\, \mathcal{F}_\tau\right) \leq \frac{\delta^2}{(\beta-1)^2}.$$

Since $\beta > 1$ it holds

$$\begin{aligned}
P\left(\sup_{t\geq 0}|M_t| > \beta\lambda, \langle M\rangle_\infty \leq \delta^2\lambda^2\right) &= P\left(\sup_{t\geq 0}|M_t| > \beta\lambda, \langle M\rangle_\infty \leq \delta^2\lambda^2, \tau < \infty\right) \\
&= E\left[P\left(\sup_{t\geq 0}|M_t| > \beta\lambda, \langle M\rangle_\infty \leq \delta^2\lambda^2 \,|\, \mathcal{F}_\tau\right) 1_{\{\tau<\infty\}}\right] \\
&\leq \frac{\delta^2}{(\beta-1)^2} P(\tau < \infty). \qquad (4.20)
\end{aligned}$$

Having chosen $\beta > 1$, by definition we have $F(\beta\lambda) \leq K_F F(\lambda)$ and analogously, as $\delta < 1$ it holds $F(\lambda/\delta) \leq K_F F(\lambda)$. Thus we may choose $\gamma = \eta = K_F$. Setting $X := \sup_{t\geq 0}|M_t|$ and $Y := \langle M\rangle_\infty$ in combination with $P(\tau < \infty) = P(\sup_{t\geq 0}|M_t| > \lambda)$, (4.20) implies that (4.18) is satisfied. For $F(x) = x^p$ we choose δ sufficiently small to ensure that $\varepsilon := \delta^2/(\beta-1)^2 < 1/K_F$ which in turn implies

$$E\left[\left(\sup_{t\geq 0}|M_t|\right)^p\right] \leq \frac{K_F^2}{1-K_F\varepsilon} E\left[\langle M\rangle_\infty^{\frac{p}{2}}\right]. \qquad (4.21)$$

The reverse inequality follows by interchanging the role of $\sup_{t\geq 0}|M_t|$ and $\langle M\rangle_\infty^{\frac{p}{2}}$ together with a slight modification of the arguments above. Indeed, we now need to consider the continuous local martingale $\bar{N} = -N$, defined by means of the stopping time $\tau := \inf\{t \geq 0 : \langle M\rangle_\infty > \lambda^2\}$, and the event $\{\langle M\rangle_\infty > \beta^2\lambda^2, \sup_{t\geq 0}|M_t| \leq \delta\lambda\}$. As before, we deduce that $\bar{N}$ hits $(\beta^2-1)\lambda^2 - 4\delta^2\lambda^2$ before hitting $-4\delta^2\lambda^2$ and thus

$$P\left(\langle M\rangle_\infty > \beta^2\lambda^2, \sup_{t\geq 0}|M_t| \leq \delta\lambda \,|\, \mathcal{F}_\tau\right) \leq \frac{4\delta^2}{(\beta^2-1)}.$$

This yields the analogues to (4.20) and finally also to (4.21), that is

$$E\left[\langle M\rangle_\infty^{\frac{p}{2}}\right] \leq \frac{K_F^2}{1-K_F\varepsilon} E\left[\left(\sup_{t\geq 0}|M_t|\right)^p\right]$$

which, together with considering the stopped martingale $M_{\cdot\wedge T}$, finishes the proof. □

Remark 4.54. By standard localization methods the assertion of the preceding theorem can be shown to hold for all continuous local martingales. ♦

4.2.4. Local Versions, Semi-Martingales, Itô-Formula

Since this lecture is not about stochastic integration as in the previous part where the construction bypass many properties about the localization and where the proofs are sometimes only sketch, this will also here be the same. This is rather to give you an idea how the stochastic integration can be extended and how you get the Itô integral.
Up to now we defined the stochastic integral with respect to continuous square integrable martingale. However, we saw that we may localise the construction by stopping the processes, so that we can define a stochastic integral by localizing.

Definition 4.55. Let $\mathcal{M}_c^{loc}$ be the set of continuous martingales such that there exists a sequence of stopping times $\tau^n \nearrow T$ with $M^{\tau^n} \in \mathcal{M}_c^2$ for every n. Similarly, we define $\mathcal{L}^{loc}(M)$ for $M \in \mathcal{L}^{loc}$ as the set of progressive measurable processes such that $\int_0^t H d\langle M\rangle < \infty$ P-almost surely for every $t \in \mathbf{T}$.

Proposition 4.56. *Let $M \in \mathcal{M}_c^{loc}$ and $H \in \mathcal{L}^{loc}(M)$, then there exists a unique continuous local martingale $H \bullet M \in \mathcal{M}_c^{loc}$ given by $H \bullet M = H \bullet M^{\tau^n}$ on $[0, \tau^n]$ for every n.*

This proposition is simple by localising M as well as $\langle M\rangle$. All the properties mentioned before also holds by localising.
This allows us now to address the celebrated Itô-Formula.

Definition 4.57. A semi-martingale is a process X with decomposition

$$X = X_0 + M + A$$

where A is the difference of two increasing continuous processes and $M \in \mathcal{M}_c^{loc}$.

We denote by $\langle X\rangle = \langle M\rangle$ the quadratic variation of the semi-martingale $X = X_0 + M + A$ with $M \in \mathcal{M}_c^{loc}$ and A of bounded variation.

Remark 4.58. A simple proof in the line of [3, Theorem 13.9] shows that $V(\Pi, X)$ converges in probability to $\langle M\rangle_t$. Applying the polar formula, it follows that $\langle X, Y\rangle = \langle M, N\rangle$ for $Y = Y_0 + N + B$. In particular the covariation with respect to a process of bounded variation is always equal to 0. ♦

A progressive process H is called locally bounded if there exists a sequence of stopping times $\tau^n \nearrow T$ such that H^{τ^n} is bounded for every n. Given such a locally bounded progressive process and $X = X_0 + M + A$ semi-martingale it is then possible to define

$$\int H dX := \int H dM + \int H dA$$

where $\int H dM$ is the stochastic integral with respect to $M \in \mathcal{M}_c^{loc}$ and $\int H dA$ is the Lebesgues-Stieljes integral of H with respect to A.
Before addressing the Itô integral, let us state the following proposition which is a stochastic version of the dominated convergence theorem of Lebesgues.

Proposition 4.59. *Let X be a local martingale and (H^n) a sequence of locally bounded progressive processes converging point-wise to 0. Suppose that $|H^n| \leq H$, then it follows that $H^n \bullet X$ converges to zero uniformly on compact in probability, that is*

$$P\left[\sup_{s \leq t}\left|\int_0^s H^n dX\right| > \varepsilon\right] \to 0$$

for every $t \in \mathbf{T}$.

Proof. For $X = A + M$, then $\sup_{s \leq t} |\int_0^t HdA|$ converges to 0 point-wise by means of Lebesgue's dominated convergence. So we just have to show the result for $X = M$. Up to a localizing sequence of stopping time, we assume that H is bounded as well as $M \in \mathcal{M}_c^2$. By Lebesgue's dominated convergence, it follows that $\|H^n\|_{2,t} \to 0$ and by Itô-isometrie $\|H^n \bullet M_t\|_2 \to 0$ Convergence in L^2 implies convergence in probability of the supremum over $s \in [0,t]$ by Doob's maximal inequalities, hence the result by noticing that $\{\tau^n \leq t\}$ goes uniformly to 0 in probability for the localising sequence. □

Remark 4.60. In particular, if H is a continuous adapted process, then, $H^n \bullet X$ converges uniformly on in probability on $[0,t]$ to $H \bullet X$ for $H^n = H_0 + \sum H_{t_k^n} 1_{]t_k^n, t_{k+1}^n]}$ and for a subdivision $\Pi^n = \{0 = t_0^n, \cdots, t_{k_n}^n = t\}$ whose mesh is converging to 0. ♦

Theorem 4.61. *Let X and Y be continuous semi-martingales, then it holds*

$$X_t Y_t = X_0 Y_0 + \int_0^t X dY + \int_0^t Y dX + \int_0^t d\langle X, Y\rangle \tag{4.22}$$

Proof. Note that for $X = Y$, Relation (4.22) reads as follows

$$X_t^2 = X_0 + 2\int_0^t X dX + \int_0^t d\langle X\rangle = X_0 + 2\int_0^t XdX + \langle X\rangle_t \tag{4.23}$$

If we show this relation, then (4.22) follows from the polar relation $XY = ((X+Y)^2 - (X-Y)^2)/4$ and the linearity of the relation (4.23). If we consider a subdivision Π, it follows that

$$X_t^2 - X_0^2 = 2\sum X_{t_k}\left(X_{t_k+1} - X_{t_k}\right) + \sum \left(X_{t_{k+1}} - X_{t_k}\right)^2$$

To the first term, we apply Remark 4.60 and for the second one Remark 4.58 to get the result as the mesh of the subdivision converges to 0. □

Note that we get by induction that

$$X_t^n = X_0 + n\int_0^t X^{n-1} dX + \frac{n(n-1)}{2}\int_0^t X^{n-2} d\langle X\rangle \tag{4.24}$$

Indeed for $n = 1$, it is immediate and for $n = 2$ it is what we have in the previous proof. Assume that it holds for every $k \leq n-1$ for $n \geq 3$, setting $Y = X^{n-1}$ we have

$$X_t^n = X_0^n + \int_0^t X dX^{n-1} + \int_0^t X^{n-1} dX + \int_0^t d\langle X, X^{n-1}\rangle \tag{4.25}$$

Applying the hypotheses of recurrence as well as the Proposition 4.51 which also applies for semi-martingales it follows that

$$\begin{aligned}
\int_0^t X dX^{n-1} &= \int_0^t X d\left(X_0 + (n-1)\int_0^t X^{n-2} dX + \frac{(n-1)(n-2)}{2}\int_0^t X^{n-3} d\langle X\rangle\right) \\
&= (n-1)\int_0^t X X^{n-2} dX + \frac{(n-1)(n-2)}{2}\int_0^t X X^{n-3} d\langle X\rangle \\
&= (n-1)\int_0^t X^{n-1} dX + \frac{(n-1)(n-2)}{2}\int_0^t X^{n-2} d\langle X\rangle
\end{aligned}$$

and since the covariation where one of the process is of bounded variation is 0 Proposition 4.51 also yields

$$\begin{aligned}\langle X, X^{n-1}\rangle &= \left\langle X, X_0 + (n-1)\int X^{n-2}dX + \frac{(n-1)(n-2)}{2}\int X^{n-3}d\langle X\rangle \right\rangle \\ &= (n-1)\int X^{n-2}d\langle X\rangle.\end{aligned}$$

Substituting both relation in (4.25) yields (4.24). By linearity it follows that for every polynomial function $f : \mathbb{R} \to \mathbb{R}$ it holds

$$f(X_t) = f(X_0) + \int_0^t f'(X)dX + \frac{1}{2}\int_0^t f''(X)d\langle X\rangle.$$

This is actually the Itô's formula when the function is even C^2.

Theorem 4.62. *Let X be a continuous semi-martingale, and $f : \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function. Then it holds*

$$f(X_t) = f(X_0) + \int_0^t f'(X)dX + \frac{1}{2}\int_0^t f''(X)d\langle X\rangle. \tag{4.26}$$

Proof. We will use the fact that (4.26) holds for the dense subset of polynomes for the supremum norm on compact intervals $[-K, K]$ denoted by $\|\cdot\|_K$. Suppose first that X is bounded by some $K > 0$. Fixing n, pick a polynome p_n such that $\|p_n - f\|_K + \|p' - f'\|_K + \|p'' - f''\|_K < 1/n$. Let

$$\begin{aligned}R = f(X_t) - p_n(X_t) - (f(X_0) - p_n(X_0))& \\ - \int_0^t (f'(X) - p_n'(X))\, dX - \frac{1}{2}\int_0^t (f''(X) - p_n''(X))\, d\langle X\rangle&.\end{aligned}$$

The first two differences can be made arbitrarily small. As for the stochastic integral difference, the uniform convergence on compact in probability follows from Proposition 4.59. Hence, up to a rapid subsequence, we have almost sure convergence. Finally the Lebesgues-Stieltjes integral converges to 0 due to Lebesgues' dominated convergence.

Assuming now that X is not bounded, as usual we localise by setting $\tau^n = \int\{t : |X_t| > n\}$ and apply the result for X^{τ^n}. It follows that

$$\begin{aligned}f(X_{t\wedge\tau^n}) = f(X_t^{\tau^n}) &= f(X_0) + \int_0^t f'(X^{\tau^n})dX^{\tau^n} + \frac{1}{2}\int_0^t f''(X^{\tau^n})d\langle X^{\tau^n}\rangle \\ &= f(X_0) + \int_0^{t\wedge\tau^n} f'(X)dX + \frac{1}{2}\int_0^{t\wedge\tau^n} f''(X)d\langle X\rangle.\end{aligned}$$

Since the probability of $\{t < \tau^n\}$ goes to 1 as n goes to infinite, we obtain the desired result. □