Stochastic integration

P.J.C. Spreij

this version: February 27, 2006

Contents

1 Stochastic processes

In this section we review some fundamental facts from the general theory of stochastic processes.

1.1 General theory

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We will use a set of time instants I. When this set is not specified, it will be $[0, \infty)$, occasionally $[0, \infty]$, or $\mathbb{N}$. Let $(E, \mathcal{E})$ another measurable space. If the set E is endowed with a metric, the Borel σ-algebra that is generated by this metric will be denoted by $\mathcal{B}(E)$ or $\mathcal{B}E$. Mostly E will be $\mathbb{R}^d$ and $\mathcal{E}$ the ordinary Borel σ-algebra on it.

Definition 1.1. A random element of E is a map from Ω into E that is $\mathcal{F}/\mathcal{E}$-measurable. A stochastic process X with time set I is a collection $\{X_t, t \in I\}$ of random elements of E. For each ω the map $t \mapsto X_t(\omega)$ is called a (sample) path, trajectory or realization of X.

Since we will mainly encounter processes where $I = [0, \infty)$, we will discuss processes whose paths are continuous, or right-continuous, or *cadlag.* The latter means that all paths are right-continuous functions with finite left limits at each $t > 0$. We will also encounter processes that satisfy these properties almost surely.

Often we have to specify in which sense two (stochastic) processes are the same. The following concepts are used.

Definition 1.2. Two processes X and Y are called indistinguishable if the set $\{X_t = Y_t, \forall t \in I\}$ contains a set of probability one (hence the paths of indistinguishable processes are a.s. equal). They are called modifications of each other if $\mathbb{P}(X_t = Y_t) = 1$, for all $t \in I$. The processes are said to have the same finite dimensional distributions if for any n-tuple $(t_1, \ldots, t_n)$ with the $t_i \in I$ the laws of the random vectors $(X_{t_1}, \ldots, X_{t_n})$ and $(Y_{t_1}, \ldots, Y_{t_n})$ coincide.

Clearly, the first of these three concepts is the strongest, the last the weakest. Whereas the first two definitions only make sense for processes defined on the same probability space, for the last one this is not necessary.

Example 1.3. Let T be a nonnegative real random variable with a continuous distribution. Let $X = 0$ and Y be defined by $Y_t(\omega) = 1_{t=T(\omega)}$, $t \in [0, \infty)$. Then X is a modification of Y, whereas $\mathbb{P}(X_t = Y_t, \forall t \geq 0) = 0$.

Proposition 1.4 *Let Y be a modification of X and assume that all paths of X and Y are right-continuous. Then X and Y are indistinguishable.*

Proof. Right-continuity allows us to write $\{X_t = Y_t, \forall t \geq 0\} = \{X_t = Y_t, \forall t \in [0, \infty) \cap \mathbb{Q}\}$. Since Y is a modification of X, the last set (is measurable and) has probability one. $\square$

Throughout the course we will need various measurability properties of stochastic processes. Viewing a process X as a map from $[0,\infty) \times \Omega$ into E, we call this process measurable if $X^{-1}(A)$ belongs to $\mathcal{B}[0,\infty) \times \mathcal{F}$ for all $A \in \mathcal{B}(E)$.

Definition 1.5. A filtration $\mathbb{F} = \{\mathcal{F}_t, t \geq 0\}$ is a collection of sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{F}_s \subset \mathcal{F}_t$ for all $s \leq t$. We put $\mathcal{F}_\infty = \sigma(\mathcal{F}_t, t \geq 0)$. Given a stochastic process X we denote by $\mathcal{F}_t^X$ the smallest σ-algebra for which all X_s, with $s \leq t$, are measurable and $\mathbb{F}^X = \{\mathcal{F}_t^X, t \geq 0\}$.
Given a filtration $\mathbb{F}$ for $t \geq 0$ the σ-algebras $\mathcal{F}_{t+}$ and $\mathcal{F}_{t-}$ for $t > 0$ are defined as follows. $\mathcal{F}_{t+} = \bigcap_{h>0} \mathcal{F}_{t+h}$ and $\mathcal{F}_{t-} = \sigma(\mathcal{F}_{t-h}, h > 0)$. We will call a filtration $\mathbb{F}$ right-continuous if $\mathcal{F}_t = \mathcal{F}_{t+}$ for all $t \geq 0$, left-continuous if $\mathcal{F}_t = \mathcal{F}_{t-}$ for all $t > 0$ and continuous if it is both left- and right-continuous. A filtration is said to satisfy the *usual conditions* if it is right-continuous and if $\mathcal{F}_0$ contains all $\mathcal{F}$-null sets. We use the notation $\mathbb{F}^+$ for the filtration $\{\mathcal{F}_{t+}, t \geq 0\}$.

Definition 1.6. Given a filtration $\mathbb{F}$ a process X is called $\mathbb{F}$-adapted, adapted to $\mathbb{F}$, or simply adapted, if for all $t \geq 0$ the random variable X_t is $\mathcal{F}_t$-measurable. Clearly, any process X is adapted to $\mathbb{F}^X$. A process X is called progressive or progressively measurable, if for all $t \geq 0$ the map $(s,\omega) \mapsto X_s(\omega)$ from $[0,t] \times \Omega$ into $\mathbb{R}$ is $\mathcal{B}[0,t] \otimes \mathcal{F}_t$-measurable. A progressive process is always adapted (exercise 1.10) and measurable.

Proposition 1.7 *Let X be a process that is (left-) or right-continuous. Then it is measurable. Such a process is progressive if it is adapted.*

Proof. We will prove the latter assertion only, the first one can be established by a similar argument. Assume that X is right-continuous and fix $t > 0$ (the proof for a left-continuous process is analogous). Then, since X is right-continuous, it is on $[0,t]$ the pointwise limit of

$$X^n = X_0 1_{\{0\}}(\cdot) + \sum_{k=1}^{2^n} 1_{((k-1)2^{-n}t, k2^{-n}t]}(\cdot) X_{k2^{-n}t},$$

All terms in the summation on the right hand side are easily seen to be $\mathcal{B}[0,t] \times \mathcal{F}_t$-measurable, and so are X^n and its limit X. □

1.2 Stopping times

Definition 1.8. A map $T : \Omega \to [0,\infty]$ is a called a random time, if T is a random variable. T is called a stopping time (w.r.t. the filtration $\mathbb{F}$, or an $\mathbb{F}$-stopping time) if for all $t \geq 0$ the set $\{T \leq t\} \in \mathcal{F}_t$.

If $T < \infty$, then T is called finite and if $\mathbb{P}(T < \infty) = 1$ it is called a.s. finite. Likewise we say that T is bounded if there is a $K \in [0,\infty)$ such that $T(\omega) \leq K$ and that T is a.s. bounded if for such a K we have $\mathbb{P}(T \leq K) = 1$. For a random time T and a process X one defines on the (measurable) set $\{T < \infty\}$ the function X_T by

$$X_T(\omega) = X_{T(\omega)}(\omega). \tag{1.1}$$

If next to the process X we also have a random variable X_∞, then X_T is defined on the whole set Ω.

The set of stopping times is closed under many operations.

Proposition 1.9 *Let $S, T, T_1, T_2, \ldots$ be stopping times. Then all random variables $S \vee T$, $S \wedge T$, $S + T$, $\sup_n T_n$ are stopping times. The random variable $\inf_n T_n$ is an $\mathbb{F}^+$-stopping time. If $a > 0$, then $T + a$ is a stopping time as well.*

Proof. Exercise 1.5. □

Suppose that E is endowed with a metric d and that the $\mathcal{E}$ is the Borel σ-algebra on E. Let $E \subset \mathbb{R}$ and X a process. The hitting time H_E is defined as $H_E = \inf\{t \geq 0 : X_t \in E\}$. Hitting times are stopping times under extra conditions.

Proposition 1.10 *If G is an open set and X a right-continuous $\mathbb{F}$-adapted process, then H_G is an $\mathbb{F}^+$-stopping time. Let F be a closed set and X a cadlag process. Define $\tilde{H}_F = \inf\{t \geq 0 : X_t \in F \text{ or } X_{t-} \in F\}$, then $\tilde{H}_F$ is an $\mathbb{F}$-stopping time. If X is a continuous process, then H_F is an $\mathbb{F}$-stopping time.*

Proof. Notice first that $\{H_G < t\} = \cup_{s<t}\{X_s \in G\}$. Since G is open and X is right-continuous, we may replace the latter union with $\cup_{s<t, s\in\mathbb{Q}}\{X_s \in G\}$. Hence $\{H_G < t\} \in \mathcal{F}_t$ and thus H_G is an $\mathbb{F}^+$-stopping time in view of exercise 1.4.
Since F is closed it is the intersection $\cap_{n=1}^\infty F^n$ of the open sets $F^n = \{x \in E : d(x, F) < \frac{1}{n}\}$. The event $\{\tilde{H}_F \leq t\}$ can be written as the union of $\{X_t \in F\}$, $\{X_{t-} \in F\}$ and $\cap_{n\geq 1} \cup_{s<t, s\in\mathbb{Q}} \{X_s \in F^n\}$ by an argument similar to the one we used above. The result follows. □

Definition 1.11. For a stopping time T we define $\mathcal{F}_T$ as the collection $\{F \in \mathcal{F}_\infty : F \cap \{T \leq t\} \in \mathcal{F}_t, \forall t \geq 0\}$. Since a constant random variable $T \equiv t_0$ is a stopping time, one has for this stopping time $\mathcal{F}_T = \mathcal{F}_{t_0}$. In this sense the notation $\mathcal{F}_T$ is unambiguous. Similarly, we have $\mathcal{F}_{T+} = \{F \in \mathcal{F} : F \cap \{T \leq t\} \in \mathcal{F}_{t+}, \forall t \geq 0\}$.

Proposition 1.12 *If S and T are stopping times, then for all $A \in \mathcal{F}_S$ it holds that $A \cap \{S \leq T\} \in \mathcal{F}_T$. Moreover $\mathcal{F}_S \cap \mathcal{F}_T = \mathcal{F}_{S\wedge T}$ and $\{S \leq T\} \in \mathcal{F}_S \cap \mathcal{F}_T$. If $S \leq T$, then $\mathcal{F}_S \subset \mathcal{F}_T$.*

Proof. Write $A \cap \{S \leq T\} \cap \{T \leq t\}$ as $A \cap \{S \leq t\} \cap \{S \wedge t \leq T \wedge t\} \cap \{T \leq t\}$. The intersection of the first two sets belongs to $\mathcal{F}_t$ if $A \in \mathcal{F}_S$, the fourth one obviously too. That the third set belongs to $\mathcal{F}_t$ is left as exercise 1.6.
It now follows from the first assertion (why?) that $\mathcal{F}_S \cap \mathcal{F}_T \supset \mathcal{F}_{S\wedge T}$. Let

$A \in \mathcal{F}_S \cap \mathcal{F}_T$. Then $A \cap \{S \wedge T \le t\} = (A \cap \{S \le t\}) \cup (A \cap \{T \le t\})$, and this obviously belongs to $\mathcal{F}_t$. From the previous assertions it follows that $\{S \le T\} \in \mathcal{F}_T$. But then we also have $\{S > T\} \in \mathcal{F}_T$ and by symmetry $\{S < T\} \in \mathcal{F}_S$. Likewise we have for every $n \in \mathbb{N}$ that $\{S < T + \frac{1}{n}\} \in \mathcal{F}_S$. Taking intersections, we get $\{S \le T\} \in \mathcal{F}_S$. □

For a stopping time T and a stochastic process X we define the stopped process X^T by $X^T_t = X_{T \wedge t}$, for $t \ge 0$.

Proposition 1.13 *Let T be a stopping time, X a progressive process. Then $Y := X_T 1_{T<\infty}$ is $\mathcal{F}_T$-measurable and the stopped process X^T is progressive too.*

Proof. First we show that X^T is progressive. Let $t > 0$. The map $\phi : ([0,t] \times \Omega, \mathcal{B}[0,t] \times \mathcal{F}_t) \to ([0,t] \times \Omega, \mathcal{B}[0,t] \times \mathcal{F}_t)$ defined by $\phi : (s,\omega) \to (T(\omega) \wedge s, \omega)$ is measurable and so is the composition $(s,\omega) \mapsto X(\phi(s,\omega)) = X_{T(\omega)\wedge s}(\omega)$, which shows that X^T is progressive. Fubini's theorem then says that the section map $\omega \mapsto X_{T(\omega)\wedge t}(\omega)$ is $\mathcal{F}_t$-measurable. To show that Y is $\mathcal{F}_T$-measurable we have to show that $\{Y \in B\} \cap \{T \le t\} = \{X_{T\wedge t} \in B\} \cap \{T \le t\} \in \mathcal{F}_t$. This has by now become obvious. □

1.3 Exercises

1.1 Let X be a measurable process on $[0,\infty)$. Then the maps $t \mapsto X_t(\omega)$ are Borel-measurable. If $\mathbb{E}\,|X_t| < \infty$ for all t, then also $t \mapsto \mathbb{E}\,X_t$ is measurable and if $\int_0^T \mathbb{E}\,|X_t|\,dt < \infty$, then $\int_0^T \mathbb{E}\,X_t\,dt = \mathbb{E}\int_0^T X_t\,dt$. Prove these statements. Show also that the process $\int_0^\cdot X_s\,ds$ is progressive if X is progressive.

1.2 Let X be a cadlag adapted process and A the event that X is continuous on an interval $[0,t)$. Then $A \in \mathcal{F}_t$.

1.3 Let X be a measurable process and T a random time. Then $X_T 1_{T<\infty}$ is a random variable.

1.4 A random time T is an $\mathbb{F}^+$-stopping time iff for all $t > 0$ one has $\{T < t\} \in \mathcal{F}_t$.

1.5 Prove proposition 1.9.

1.6 Let S, T be stopping times bounded by a constant t_0. Then $\sigma(S) \subset \mathcal{F}_{t_0}$ and $\{S \le T\} \in \mathcal{F}_{t_0}$.

1.7 Let T be a stopping time and S a random time such that $S \ge T$. If S is $\mathcal{F}_T$-measurable, then S is a stopping time too.

1.8 Let S and T be stopping times and Z an integrable random variable. Then $\mathbb{E}\,[\mathbb{E}\,[Z|\mathcal{F}_T]|\mathcal{F}_S] = \mathbb{E}\,[Z|\mathcal{F}_{S\wedge T}]$.

1.9 Let S and T be stopping times such that $S \le T$ and even $S < T$ on the set $\{S < \infty\}$. Then $\mathcal{F}_{S+} \subset \mathcal{F}_T$.

1.10 Show that a progressive process is adapted and measurable.

1.11 Show that $\mathbb{F}^+$ is a right-continuous filtration.

1.12 Show that $\mathbb{F}^X$ is a left-continuous filtration if X is a left-continuous process.

1.13 Let X be a process that is adapted to a filtration $\mathbb{F}$. Let Y be a modification of X. Show that also Y is adapted to $\mathbb{F}$ if $\mathbb{F}$ satisfies the usual conditions.

1.14 Show that proposition 1.13 can be refined as follows. Under the stated assumptions the process X^T is even progressive w.r.t. the filtration $\mathcal{F}_{t\wedge T}$.

2 Martingales

In this section we review some properties of martingales and supermartingales. Throughout the section we work with a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\mathbb{F}$. Familiarity with the theory of martingales in discrete time is assumed.

2.1 Generalities

We start with the definition of martingales, submartingales and supermartingales in continuous time.

Definition 2.1. A real-valued process X is called a supermartingale, if it is adapted, if all X_t are integrable and $\mathbb{E}\,[X_t|\mathcal{F}_s] \leq X_s$ a.s. for all $t \geq s$. If $-X$ is a supermartingale, we call X a submartingale and X will be called a martingale if it is both a submartingale and a supermartingale.

The number of upcrossings of a stochastic process X over an interval $[a, b]$ when time runs through a finite set F is denoted by $U([a, b]; F)$. Then we define the number of upcrossings over the interval $[a, b]$ when time runs through an interval $[s, t]$ by $\sup\{U([a, b]; F) : F \subset [s, t], F \text{ finite}\}$. Fundamental properties of (sub)martingales are collected in the following proposition.

Proposition 2.2 *Let X be a submartingale with right-continuous paths. Then*
(i) For all $\lambda > 0$ and $0 \leq s \leq t$ one has

$$\lambda \mathbb{P}(\sup_{s \leq u \leq t} X_u \geq \lambda) \leq \mathbb{E}\, X_t^+ \tag{2.1}$$

and

$$\lambda \mathbb{P}(\inf_{s \leq u \leq t} X_u \leq -\lambda) \leq \mathbb{E}\, X_t^+ - \mathbb{E}\, X_s. \tag{2.2}$$

(ii) If X is a nonnegative submartingale and $p > 1$, then

$$|| \sup_{s \leq u \leq t} X_u ||_p \leq \frac{p}{p-1} ||X_t||_p, \tag{2.3}$$

where for random variables ξ we put $||\xi||_p = (\mathbb{E}\,|\xi|^p)^{1/p}$.
(iii) For the number of upcrossings $U([a, b]; [s, t])$ it holds that

$$\mathbb{E}\, U([a, b]; [s, t]) \leq \frac{\mathbb{E}\, X_t^+ + |a|}{b - a}.$$

(iv) Almost every path of X is bounded on compact intervals and has no discontinuities of the second kind. For almost each path the set of points at which this path is discontinuous is at most countable.

Proof. The proofs of these results are essentially the same as in the discrete time case. The basic argument to justify this claim is to consider X restricted to a finite set F in the interval $[s,t]$. With X restricted to such a set, the above inequalities in (i) and (ii) are valid as we know it from discrete time theory. Take then a sequence of such F, whose union is a dense subset of $[s,t]$, then the inequalities keep on being valid. By right-continuity of X we can extend the validity of the inequalities to the whole interval $[s,t]$. The same reasoning applies to (iii).
For (iv) we argue as follows. Combining the inequalities in (i), we see that $\mathbb{P}(\sup_{s\leq u\leq t}|X_u| = \infty) = 0$, hence almost all sample paths are bounded on intervals $[s,t]$. We have to show that almost all paths of X admit left limits everywhere. This follows as soon as we can show that for all n the set $\{\liminf_{s\uparrow t} X_s < \limsup_{s\uparrow t} X_s, \text{for some } t\in[0,n]\}$ has zero probability. This set is contained in $\cup_{a,b\in\mathbb{Q}}\{U([a,b],[0,n])=\infty\}$. But by (iii) this set has zero probability. It is a *fact* from analysis that the set of discontinuities of the first kind of a function is at most countable, which yields the last assertion. □

The last assertion of this proposition describes a regularity property of the sample paths of a right-continuous submartingale. The next theorem gives a sufficient and necessary condition that justifies the fact that we mostly restrict our attention to cadlag submartingales. This condition is trivially satisfied for martingales. We state the theorem without proof.

Theorem 2.3 *Let X be a submartingale and let the filtration $\mathbb{F}$ satisfy the usual conditions. Suppose that the function $t\mapsto \mathbb{E}\,X_t$ is right-continuous on $[0,\infty)$. Then there exists a modification Y of X that has cadlag paths and that is also a submartingale w.r.t. $\mathbb{F}$.*

2.2 Limit theorems and optional sampling

The following two theorems are the fundamental convergence theorems.

Theorem 2.4 *Let X be a right-continuous submartingale with $\sup_{t\geq 0}\mathbb{E}\,X_t^+ < \infty$. Then there exists a $\mathcal{F}_\infty$-measurable random variable X_∞ with $\mathbb{E}\,|X_\infty| < \infty$ such that $X_t \stackrel{a.s.}{\to} X_\infty$. If moreover X is uniformly integrable, then we also have $X_t \stackrel{L^1}{\to} X_\infty$ and $\mathbb{E}\,[X_\infty|\mathcal{F}_t] \geq X_t$ a.s. for all $t\geq 0$.*

Theorem 2.5 *Let X be a martingale. Then X is uniformly integrable iff there exists an integrable random variable Z such that $\mathbb{E}\,[Z|\mathcal{F}_t] = X_t$ a.s. for all $t\geq 0$. In this case we have $X_t \stackrel{a.s.}{\to} X_\infty$ and $X_t \stackrel{L^1}{\to} X_\infty$, where $X_\infty = \mathbb{E}\,[Z|\mathcal{F}_\infty]$.*

Proof. The proofs of these theorems are like in the discrete time case. □

We will frequently use the optional sampling theorem (theorem 2.7) for submartingales and martingales. In the proof of this theorem we use the following lemma.

Lemma 2.6 *Let $(\mathcal{G}_n)$ be a decreasing sequence of σ-algebras and let (Y_n) be a sequence of integrable random variables such that Y_n is a $\mathcal{G}_n$-measurable random variable for all n and such that*

$$\mathbb{E}\,[Y_m|\mathcal{G}_n] \geq Y_n, \forall n \geq m. \tag{2.4}$$

If the sequence of expectations $\mathbb{E}\,Y_n$ is bounded from below, then the collection $\{Y_n, n \geq 1\}$ is uniformly integrable.

Proof. Consider the chain of (in)equalities (where $n \geq m$)

$$\begin{aligned}
\mathbb{E}\,1_{|Y_n|>\lambda}|Y_n| &= \mathbb{E}\,1_{Y_n>\lambda}Y_n - \mathbb{E}\,1_{Y_n<-\lambda}Y_n \\
&= -\mathbb{E}\,Y_n + \mathbb{E}\,1_{Y_n>\lambda}Y_n + \mathbb{E}\,1_{Y_n\geq-\lambda}Y_n \\
&\leq -\mathbb{E}\,Y_n + \mathbb{E}\,1_{Y_n>\lambda}Y_m + \mathbb{E}\,1_{Y_n\geq-\lambda}Y_m \\
&\leq -\mathbb{E}\,Y_n + \mathbb{E}\,Y_m + \mathbb{E}\,1_{|Y_n|>\lambda}|Y_m|.
\end{aligned} \tag{2.5}$$

Since the sequence of expectations $(\mathbb{E}\,Y_n)$ has a limit and hence is Cauchy, we choose for given $\varepsilon > 0$ the integer m such that for all $n > m$ we have $-\mathbb{E}\,Y_n + \mathbb{E}\,Y_m < \varepsilon$. By the conditional version of Jensen's inequality we have $\mathbb{E}\,Y_n^+ \leq \mathbb{E}\,Y_1^+$. Since $\mathbb{E}\,Y_n \geq l$ for some finite l, we can conclude that $\mathbb{E}\,|Y_n| = 2\mathbb{E}\,Y_n^+ - \mathbb{E}\,Y_n \leq 2\mathbb{E}\,Y_1^+ - l$. This implies that $\mathbb{P}(|Y_n| > \lambda) \leq \frac{2\mathbb{E}\,Y_1^+ - l}{\lambda}$. Hence we can make the expression in (2.5) arbitrarily small for all n big enough and we can do this uniformly in n (fill in the details). □

Theorem 2.7 *Let X be a right-continuous submartingale with a last element X_∞ (i.e. $\mathbb{E}\,[X_\infty|\mathcal{F}_t] \geq X_t$, a.s. for every $t \geq 0$) and let S and T be two stopping times such that $S \leq T$. Then $X_S \leq \mathbb{E}\,[X_T|\mathcal{F}_S]$ a.s.*

Proof. Let $T_n = 2^{-n}[2^nT + 1]$ and $S_n = 2^{-n}[2^nS + 1]$. Then the T_n and S_n are stopping times (exercise 2.4) with $T_n \geq S_n$, and they form two decreasing sequences with limits T, respectively S. Notice that all T_n and S_n are at most countably valued. We can apply the optional sampling theorem for discrete time submartingales to get $\mathbb{E}\,[X_{T_n}|\mathcal{F}_{S_n}] \geq X_{S_n}$, from which we obtain that for each $A \in \mathcal{F}_S \subset \mathcal{F}_{S_n}$ it holds that

$$\mathbb{E}\,1_AX_{T_n} \geq \mathbb{E}\,1_AX_{S_n}. \tag{2.6}$$

It similarly follows, that $\mathbb{E}\,[X_{T_n}|\mathcal{F}_{T_{n+1}}] \geq X_{T_{n+1}}$. Notice that the expectations $\mathbb{E}\,X_{T_n}$ form a decreasing sequence with lower bound $\mathbb{E}\,X_0$. We can thus apply lemma 2.6 to conclude that the collection $\{X_{T_n} : n \geq 1\}$ is uniformly integrable. Of course the same holds for $\{X_{S_n} : n \geq 1\}$. By right-continuity of X we get $\lim_{n\to\infty} X_{T_n} = X_T$ a.s and $\lim_{n\to\infty} X_{S_n} = X_S$ a.s. Uniform integrability implies that we also have L^1-convergence, hence we have from equation (2.6) that $\mathbb{E}\,1_AX_T \geq \mathbb{E}\,1_AX_S$ for all $A \in \mathcal{F}_S$, which is what we had to prove. □

Remark 2.8. The condition that X has a last element can be replaced by restriction to bounded stopping times S and T.

Corollary 2.9 *Let X be a right-continuous submartingale and let $S \leq T$ be stopping times. Then the stopped process X^T is a submartingale as well and $\mathbb{E}\,[X_{T\wedge t}|\mathcal{F}_S] \geq X_{S\wedge t}$ a.s. for every $t \geq 0$.*

Proof. This is exercise 2.3. □

2.3 Doob-Meyer decomposition

It is instructive to formulate the discrete time analogues of what will be described below. This will be left as exercises.

Definition 2.10. (i) An adapted process A is called increasing if a.s. we have $A_0 = 0$, $t \to A_t(\omega)$ is a nondecreasing right-continuous function and if $\mathbb{E}\,A_t < \infty$ for all t. An increasing process is called integrable if $\mathbb{E}\,A_\infty < \infty$, where $A_\infty = \lim_{t\to\infty} A_t$.
(ii) An increasing process A is called natural if for every right-continuous and bounded martingale M we have

$$\mathbb{E} \int_{(0,t]} M_s\, dA_s = \mathbb{E} \int_{(0,t]} M_{s-}\, dA_s, \forall t \geq 0. \tag{2.7}$$

Remark 2.11. The integrals $\int_{(0,t]} M_s\, dA_s$ and $\int_{(0,t]} M_{s-}\, dA_s$ in equation (2.7) are defined pathwise as Lebesgue-Stieltjes integrals for all $t \geq 0$ and hence the corresponding processes are progressive. Furthermore, if the increasing process A is continuous, then it is natural. This follows from the fact the paths of M have only countably many discontinuities (proposition 2.2(iv)).

Definition 2.12. A right-continuous process is said to belong to class D if the collection of X_T, where T runs through the set of all finite stopping times, is uniformly integrable. A right-continuous process X is said to belong to class DL if for all $a > 0$ the collection X_T, where T runs through the set of all stopping times bounded by a, is uniformly integrable.

With these definitions we can state and prove the celebrated Doob-Meyer decomposition (theorem 2.14 below) of submartingales. In the proof of this theorem we use the following lemma. It is a rather deep result in functional analysis.

Lemma 2.13 *The collection $\{Z_n, n \geq 1\}$ of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a uniformly integrable family, iff there exist an increasing subsequence (n_k) and a random variable Z such that for all bounded random variables ζ one has $\lim_{k\to\infty} \mathbb{E}\,Z_{n_k}\zeta = \mathbb{E}\,Z\zeta$. Moreover, if $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$, then also $\lim_{k\to\infty} \mathbb{E}\,[\mathbb{E}\,[Z_{n_k}|\mathcal{G}]\zeta] = \mathbb{E}\,[\mathbb{E}\,[Z|\mathcal{G}]\zeta]$.*

Theorem 2.14 *Let the filtration $\mathbb{F}$ satisfy the usual conditions and let X be a right-continuous submartingale of class DL. Then X can be decomposed according to*

$$X = A + M, \tag{2.8}$$

where M is a right-continuous martingale and A a natural increasing process. The decomposition (2.8) *is unique up to indistinguishability. Under the stronger condition that X is of class D, M is uniformly integrable and A is integrable.*

Proof. Showing the uniqueness assertion is relatively easy. Suppose that we have next to (2.8) another decomposition $X = A' + M'$ of the same type. Then $B := A - A' = M' - M$ is a right-continuous martingale of bounded variation with $B_0 = 0$. Since the processes A and A' are natural, we have for any bounded right-continuous martingale ξ and for all $t \geq 0$ the equality (this is exercise 2.5)

$$\mathbb{E}\, B_t \xi_t = \mathbb{E} \int_{(0,t]} \xi_{s-}\, dB_s. \tag{2.9}$$

Take first ξ of the form $\xi = \sum_k \xi^{k-1} 1_{[t_{k-1},t_k)}$, with $0 = t_0 \leq \cdots \leq t_n = t$ and ξ^{k-1} $\mathcal{F}_{t_{k-1}}$-measurable. Then $\xi_- = \xi^0 1_{\{0\}} + \sum_k \xi^{k-1} 1_{(t_{k-1},t_k]}$ and the right hand side of (2.9) becomes $\mathbb{E} \sum_k \xi^{k-1}(B_{t_k} - B_{t_{k-1}}) = 0$, since B is a martingale. If ξ is a continuous time martingale we take a sequence of nested partitions of $[0,t]$ with points t_k whose mesh converges to zero. By taking $\xi^k = \xi_{t_{k-1}}$, we apply the previous result to obtain $\mathbb{E}\, \xi_t B_t = 0$.
Let ζ be a bounded random variable and put $\xi_t \equiv \mathbb{E}\,[\zeta|\mathcal{F}_t]$, then ξ is a martingale. We know from theorem 2.3 that ξ admits a right continuous modification and hence we get $\mathbb{E}\,(A_t - A'_t)\zeta = \mathbb{E}\, B_t \zeta = 0$. Choosing ζ in an appropriate way (how?) we conclude that $A_t = A'_t$ a.s. We can do this for any $t \geq 0$ and then right-continuity yields that A and A' are indistinguishable, and the same holds for M and M'.

We now establish the existence of the decomposition for $t \in [0,a]$ for an arbitrary $a > 0$. Let Y be the process defined by $Y_t := X_t - \mathbb{E}\,[X_a|\mathcal{F}_t]$. Since X is a submartingale $Y_t \leq 0$ for $t \in [0,a]$. Moreover the process Y is a submartingale itself that has a right-continuous modification (exercise 2.14), again denoted by Y. From now on we work with this modification. Consider the nested sequence of dyadic partitions Π^n of $[0,a]$, $\Pi^n \ni t^n_j = j2^{-n}a$, $j = 0, \ldots, 2^n$. With time restricted to Π^n, the process Y becomes a discrete time submartingale to which we apply the Doob decomposition (exercise 2.8), which results in $Y = A^n + M^n$, with $A^n_0 = 0$. Since $Y_a = 0$, we get

$$A^n_{t^n_j} = Y_{t^n_j} + \mathbb{E}\,[A^n_a|\mathcal{F}_{t^n_j}]. \tag{2.10}$$

We now make the following claim.

$$\text{The family } \{A^n_a : n \geq 1\} \text{ is uniformly integrable.} \tag{2.11}$$

Supposing that the claim holds true, we finish the proof as follows. In view of lemma 2.13 we can select from this family a subsequence, *again denoted by* A^n_a, and we can find a random variable A_a such that $\mathbb{E}\,\zeta A^n_a \to \mathbb{E}\, A_a \zeta$ for every bounded random variable ζ. We now define for $t \in [0,a]$ the random variable A_t as any version of

$$A_t = Y_t + \mathbb{E}\,[A_a|\mathcal{F}_t]. \tag{2.12}$$

The process A has a right-continuous modification on $[0,a]$, again denoted by A. We now show that A is increasing. Let $s,t \in \bigcup_n \Pi^n$. Then there is n_0 such that for all $n \geq n_0$ we have $s,t \in \Pi^n$. Using the second assertion of lemma 2.13, equations (2.10) and (2.12) we obtain $\mathbb{E}\,\zeta(A^n_t - A^n_s) \to \mathbb{E}\,\zeta(A_t - A_s)$. Since for each n the A^n is increasing we get that $\mathbb{E}\,\zeta(A_t - A_s) \geq 0$ as soon as $\zeta \geq 0$. Take now $\zeta = 1_{\{A_s > A_t\}}$ to conclude that $A_t \geq A_s$ a.s. Since A is right-continuous, we have that A is increasing on the whole interval $[0,a]$. It follows from the construction that $A_0 = 0$ a.s. (exercise 2.10).
Next we show that A is natural. Let ξ be a bounded and right-continuous martingale. The discrete time process A^n is predictable (and thus natural, exercise 2.9). Restricting time to Π^n and using the fact that Y and A as well as Y and A^n differ by a martingale (equations (2.10) and (2.12)), we get

$$
\begin{aligned}
\mathbb{E}\,\xi_a A^n_a &= \mathbb{E}\sum_j \xi_{t^n_{j-1}}(A^n_{t^n_j} - A^n_{t^n_{j-1}}) \\
&= \mathbb{E}\sum_j \xi_{t^n_{j-1}}(Y_{t^n_j} - Y_{t^n_{j-1}}) \\
&= \mathbb{E}\sum_j \xi_{t^n_{j-1}}(A_{t^n_j} - A_{t^n_{j-1}})
\end{aligned}
$$

Let $n \to \infty$ to conclude that

$$
\mathbb{E}\,\xi_a A_a = \mathbb{E}\int_{(0,a]} \xi_{s-}\, dA_s, \tag{2.13}
$$

by the definition of A_a and the dominated convergence theorem applied to the right hand side of the above string of equalities. Next we replace in (2.13) ξ with the, at the deterministic time $t \leq a$, stopped process ξ^t. A computation shows that we can conclude

$$
\mathbb{E}\,\xi_t A_t = \mathbb{E}\int_{(0,t]} \xi_{s-}\, dA_s, \forall t \leq a. \tag{2.14}
$$

In view of exercise 2.5 this shows that A is natural on $[0,a]$. The proof of the first assertion of the theorem is finished by setting $M_t = \mathbb{E}\,[X_a - A_a|\mathcal{F}_t]$. Clearly M is a martingale and $M_t = X_t - A_t$. Having shown that the decomposition is valid on each interval $[0,a]$, we invoke uniqueness to extend it to $[0,\infty)$.
If X belongs to class D, it has a last element and we can repeat the above proof with a replaced by ∞. What is left however, is the proof of the claim (2.11). This will be given now.
Let $\lambda > 0$ be fixed. Define $T^n = \inf\{t^n_{j-1} : A^n_{t^n_j} > \lambda, j = 1, \ldots, 2^n\} \wedge a$. One checks that the T^n are stopping times, bounded by a, that $\{T^n < a\} = \{A^n_a > \lambda\}$ and that we have $A^n_{T^n} \leq \lambda$ on this set. By the optional sampling theorem (in discrete time) applied to (2.10) one has $Y_{T^n} = A^n_{T^n} - \mathbb{E}\,[A^n_a|\mathcal{F}_{T^n}]$. Below we will often use that $Y \leq 0$. To show uniform integrability, one considers $\mathbb{E}\,1_{\{A^n_a > \lambda\}} A^n_a$ and with the just mentioned facts one gets

$$
\begin{aligned}
\mathbb{E}\,1_{\{A^n_a > \lambda\}} A^n_a &= \mathbb{E}\,1_{\{A^n_a > \lambda\}} A^n_{T^n} - \mathbb{E}\,1_{\{A^n_a > \lambda\}} Y_{T^n} \\
&\leq \lambda \mathbb{P}(T^n < a) - \mathbb{E}\,1_{\{T^n < a\}} Y_{T^n}. \qquad (2.15)
\end{aligned}
$$

Let $S^n = \inf\{t^n_{j-1} : A^n_{t^n_j} > \frac{1}{2}\lambda, j = 1, \ldots, 2^n\} \wedge a$. Then the S^n are bounded stopping times as well with $S^n \leq T^n$ and we have, like above, $\{S^n < a\} = \{A^n_a > \frac{1}{2}\lambda\}$ and $A^n_{S^n} \leq \frac{1}{2}\lambda$ on this set. Using $Y_{S^n} = A^n_{S_n} - \mathbb{E}\,[A^n_a|\mathcal{F}_{S^n}]$ we develop

$$
\begin{aligned}
-\mathbb{E}\,1_{\{S^n<a\}}Y_{S^n} &= -\mathbb{E}\,1_{\{S^n<a\}}A^n_{S_n} + \mathbb{E}\,1_{\{S^n<a\}}A^n_a \\
&= \mathbb{E}\,1_{\{S^n<a\}}(A^n_a - A^n_{S_n}) \\
&\geq \mathbb{E}\,1_{\{T^n<a\}}(A^n_a - A^n_{S_n}) \\
&\geq \frac{1}{2}\lambda\mathbb{P}(T^n < a).
\end{aligned}
$$

It follows that $\lambda\mathbb{P}(T^n < a) \leq -2\mathbb{E}\,1_{\{S^n<a\}}Y_{S^n}$. Inserting this estimate into inequality (2.15), we obtain

$$
\mathbb{E}\,1_{\{A^n_a>\lambda\}}A^n_a \leq -2\mathbb{E}\,1_{\{S^n<a\}}Y_{S^n} - \mathbb{E}\,1_{\{T^n<a\}}Y_{T^n}. \tag{2.16}
$$

The assumption that X belongs to class DL implies that both $\{Y_{T^n}, n \geq 1\}$ and $\{Y_{S^n}, n \geq 1\}$ are uniformly integrable, since Y and X differ for $t \in [0, a]$ by a uniformly integrable martingale. Then we can find for all $\varepsilon > 0$ a $\delta > 0$ such that $-\mathbb{E}\,1_F Y_{T^n}$ and $-\mathbb{E}\,1_F Y_{S^n}$ are smaller than ε for any set F with $\mathbb{P}(F) < \delta$ uniformly in n. Since $Y_a = 0$, we have $\mathbb{E}\,A^n_a = -\mathbb{E}\,M^n_a = -\mathbb{E}\,M^n_0 = -\mathbb{E}\,Y_0$. Hence we get $\mathbb{P}(T_n < a) \leq \mathbb{P}(S_n < a) = \mathbb{P}(A^n_a > \lambda/2) \leq -2\mathbb{E}\,Y_0/\lambda$. Hence for λ bigger than $-2\mathbb{E}\,Y_0/\delta$, we can make both probabilities $\mathbb{P}(T_n < a)$ and $\mathbb{P}(S_n < a)$ of the sets in (2.16) less than δ, uniformly in n. This shows that the family $\{A^n_a : n \geq 1\}$ is uniformly integrable. □

The general statement of theorem 2.14 only says that the processes A and M can be chosen to be right-continuous. If the process X is continuous, one would expect A and M to be continuous. This is true indeed, at least for nonnegative processes. However a stronger result can be proven.

Theorem 2.15 *Suppose that in addition to the conditions of theorem* 2.14 *the submartingale is such that* $\lim_{n\to\infty} \mathbb{E}\,X_{T^n} = \mathbb{E}\,X_T$ *for all bounded increasing sequences of stopping times* T^n *with limit* T. *Then the process* A *in equation* (2.8) *is continuous.*

Proof. Let $a > 0$ be an upper bound for T and the T^n. According to theorem 2.14 we have $X = A + M$. Hence, according to the optional sampling theorem we have $\mathbb{E}\,A_{T^n} = \mathbb{E}\,X_{T^n} - \mathbb{E}\,M_{T^n} = \mathbb{E}\,X_{T^n} - \mathbb{E}\,M_a$, which implies that

$$
\mathbb{E}\,A_{T^n} \uparrow \mathbb{E}\,A_T. \tag{2.17}
$$

By the monotone convergence theorem we conclude that $A_{T^n(\omega)}(\omega) \uparrow A_{T(\omega)}(\omega)$ for all ω outside a set of probability zero, which however in principle depends on the chosen sequence of stopping times.

We proceed as follows. Assume for the time being that A is a bounded process. As in the proof of theorem 2.14 we consider the dyadic partitions Π^n of the interval $[0, a]$. For every n and every $j = 0, \ldots, 2^n - 1$ we define $\xi^{n,j}$ to be the right-continuous modification of the martingale defined by $\mathbb{E}\,[A_{t^n_{j+1}}|\mathcal{F}_t]$. Then we define

$$\xi^n_t = \sum_{j=0}^{2^n-1} \xi^{n,j}_t 1_{(t^n_j,t^n_{j+1}]}(t).$$

The process ξ^n is on the whole interval $[0,a]$ right-continuous, except possibly at the points of the partition. Notice that we have $\xi^n_t \geq A_t$ a.s. for all $t \in [0,a]$ with equality at the points of the partition and $\xi^{n+1} \leq \xi^n$. Since A is a natural increasing process, we have for every n and j that

$$\mathbb{E} \int_{(t^n_j,t^n_{j+1}]} \xi^n_s \, dA_s = \mathbb{E} \int_{(t^n_j,t^n_{j+1}]} \xi^n_{s-} \, dA_s.$$

Summing this equality over all relevant j we get

$$\mathbb{E} \int_{(0,t]} \xi^n_s \, dA_s = \mathbb{E} \int_{(0,t]} \xi^n_{s-} \, dA_s, \text{ for all } t \in [0,a]. \tag{2.18}$$

We introduce the right-continuous nonnegative processes η^n defined by $\eta^n_t = (\xi^n_{t+} - A_t)1_{[0,a)}(t)$. For fixed $\varepsilon > 0$ we define the bounded stopping times (see proposition 1.10) $T^n = \inf\{t \in [0,a] : \eta^n_t > \varepsilon\} \wedge a$. Observe that also $T^n = \inf\{t \in [0,a] : \xi^n_t - A_t > \varepsilon\} \wedge a$ in view of the relation between η^n and ξ^n (exercise 2.11). Since the ξ^n_t are decreasing in n, we have that the sequence T^n is increasing and thus has a limit, T say. Let $\phi^n(t) = \sum_{j=0}^{2^n-1} t^n_{j+1} 1_{(t^n_j,t^n_{j+1}]}(t)$ and notice that for all n and $t \in [0,a]$ one has $a \geq \phi^n(t) \geq \phi^{n+1}(t) \geq t$. It follows that $\lim_{n\to\infty} \phi^n(T^n) = T$ a.s. From the optional sampling theorem applied to the martingales $\xi^{n,j}$ it follows that

$$\begin{aligned}
\mathbb{E}\,\xi^n_{T^n} &= \mathbb{E} \sum_j \mathbb{E}\,[A_{t^n_{j+1}}|\mathcal{F}_{T^n}] 1_{(t^n_j,t^n_{j+1}]}(T^n) \\
&= \mathbb{E} \sum_j \mathbb{E}\,[A_{t^n_{j+1}} 1_{(t^n_j,t^n_{j+1}]}(T^n)|\mathcal{F}_{T^n}] \\
&= \mathbb{E} \sum_j \mathbb{E}\,[A_{\phi^n(T^n)} 1_{(t^n_j,t^n_{j+1}]}(T^n)|\mathcal{F}_{T^n}] \\
&= \mathbb{E}\,A_{\phi^n(T^n)}.
\end{aligned}$$

And then

$$\begin{aligned}
\mathbb{E}\,(A_{\phi^n(T^n)} - A_{T^n}) &= \mathbb{E}\,(\xi^n_{T^n} - A_{T^n}) \\
&\geq \mathbb{E}\,1_{\{T^n<a\}}(\xi^n_{T^n} - A_{T^n}) \\
&\geq \varepsilon \mathbb{P}(T^n < a).
\end{aligned}$$

Thus

$$\mathbb{P}(T^n < a) \leq \frac{1}{\varepsilon}\mathbb{E}\,(A_{\phi^n(T^n)} - A_{T^n}) \to 0 \text{ for } n \to \infty, \tag{2.19}$$

by (2.17) and right-continuity and boundedness of A. But since $\{T^n < a\} = \{\sup_{t\in[0,a]} |\xi_t^n - A_t| > \varepsilon\}$ and since (2.19) holds for every $\varepsilon > 0$, we can find a subsequence (n_k) along which $\sup_{t\in[0,a]} |\xi_t^n - A_t| \stackrel{a.s.}{\to} 0$. Since both A and ξ are bounded, it follows from the dominated convergence theorem applied to equation (2.18) and the chosen subsequence that

$$\mathbb{E}\int_{(0,t]} A_s\, dA_s = \mathbb{E}\int_{(0,t]} A_{s-}\, dA_s, \text{ for all } t \in [0,a],$$

from which we conclude by monotonicity that

$$\mathbb{E}\int_{(0,t]} A_s\, dA_s = \mathbb{E}\int_{(0,t]} A_{s-}\, dA_s, \text{ for all } t \in [0,\infty). \tag{2.20}$$

Thus the nonnegative process $A - A_-$ must be zero $dA_t(\omega)\mathbb{P}(d\omega)$-a.e. and therefore indistinguishable from zero.
The result has been proved under the assumption that A is bounded. If this assumption doesn't hold, we can introduce the stopping times $S_m = \inf\{t \geq 0 : A_t > m\}$ and replace everywhere above A with the stopped process A^{S_m} which is bounded. The conclusion from the analogue of (2.20) will be that the processes $A1_{[0,S_m]}$ and $A_-1_{[0,S_m]}$ are indistinguishable for all m and the final result follows by letting $m \to \infty$. □

We close this section with the following proposition.

Proposition 2.16 *Let X satisfy the conditions of theorem 2.14 with Doob-Meyer decomposition $X = A + M$ and T a stopping time. Then also the stopped process X^T satisfies these conditions and its Doob-Meyer decomposition is given by*

$$X^T = A^T + M^T.$$

Proof. That X^T also satisfies the conditions of theorem 2.14 is straightforward. Of course, by stopping, we have $X^T = A^T + M^T$, so we only have to show that this indeed gives us the Doob-Meyer decomposition. By the optional sampling theorem and its corollary 2.9, the stopped process M^T is a martingale. That the process A^T is natural follows from the identity

$$\int_{(0,t]} \xi_s\, dA_s^T = \int_{(0,t]} \xi_s^T\, dA_s - \xi_T(A_t - A_{t\wedge T}),$$

a similar one for $\int_{(0,t]} \xi_{s-}\, dA_s^T$, valid for any bounded measurable process ξ and the fact that A is natural. The uniqueness of the Doob-Meyer decomposition then yields the result. □

2.4 Exercises

2.1 Let X be a supermartingale with constant expectation. Show that X is in fact a martingale. *Hint:* consider expectations over the set $\{\mathbb{E}[X_t|\mathcal{F}_s] > X_s\}$.

2.2 Why is $\sup_{s \le u \le t} X_u$ in formula (2.3) measurable?

2.3 Prove corollary 2.9.

2.4 Show that the random variables T_n in the proof of theorem 2.7 are stopping times and that they form a non-increasing sequence with pointwise limit T.

2.5 Show that an increasing process A is natural iff for all bounded right-continuous martingales one has $\mathbb{E} \int_{(0,t]} M_{s-}\, dA_s = \mathbb{E}\, M_t A_t$ for all $t \ge 0$. *Hint:* Consider first a partition $0 = t_0 \le t_1 \le \cdots \le t_n = t$ and take M constant on the $(t_{k-1}, t_k]$. Then you approximate M with martingales of the above type.

2.6 Show that a uniformly integrable martingale is of class D.

2.7 Let X be a right-continuous submartingale. Show that X is of class DL if X is nonnegative or if $X = M + A$, where A is an increasing process and M a martingale.

2.8 Let X be a discrete time process on some $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathbb{F} = \{\mathcal{F}_n\}_{n \ge 0}$ be a filtration. Assume that X is adapted, $X_0 = 0$ and that X_n is integrable for all n. Define the process M by $M_0 = 0$ and $M_n = M_{n-1} + X_n - \mathbb{E}\,[X_n|\mathcal{F}_{n-1}$ for $n \ge 1$. Show that M is a martingale. Define then $A = X - M$. Show that A_n is $\mathcal{F}_{n-1}$-measurable for all $n \ge 1$ (one says that A is predictable) and that $A_0 = 0$. Show that A is increasing iff X is a submartingale.

2.9 A discrete time process A is called *increasing* if it is adapted, $A_0 = 0$ a.s., $A_n - A_{n-1} \ge 0$ a.s. and $\mathbb{E}\, A_n < \infty$ for all $n \ge 1$. An increasing process is natural if for all bounded martingales M one has $\mathbb{E}\, A_n M_n = \mathbb{E} \sum_{k=1}^n M_{k-1}(A_k - A_{k-1})$.
(i) Show that a process A is natural iff for all bounded martingales one has $\mathbb{E} \sum_{k=1}^n A_k (M_k - M_{k-1}) = 0$.
(ii) Show that a predictable increasing process is natural.
(iii) Show that a natural process is a.s. predictable. *Hint:* you have to show that $A_n = \tilde{A}_n$ a.s., where $\tilde{A}_n$ is a version of $\mathbb{E}\,[A_n|\mathcal{F}_{n-1}]$ for each n, which you do along the following steps. First you show that for all n one has $\mathbb{E}\, M_n A_n = \mathbb{E}\, M_{n-1} A_n = \mathbb{E}\, M_n \tilde{A}_n$. Fix n, take $M_k = \mathbb{E}\,[\mathrm{sgn}(A_n - \tilde{A}_n)|\mathcal{F}_k]$ and finish the proof.

2.10 Show that the A_0 in the proof of the Doob-Meyer decomposition (theorem 2.14) is zero a.s.

2.11 Show that (in the proof of theorem 2.15) $T^n = \inf\{t \in [0,a] : \xi^n_t - A_t > \varepsilon\} \wedge a$. *Hint:* use that ξ^n is right-continuous except possibly at the t^n_j.

2.12 A continuous nonnegative submartingale satisfies the conditions of theorem 2.15. Show this.

2.13 Show (in the proof of theorem 2.15) the convergence $\phi^n(T^n) \to T$ a.s. and the convergence in (2.19).

2.14 Suppose that X is a submartingale with right-continuous paths. Show that $t \mapsto \mathbb{E}\, X_t$ is right-continuous (use lemma 2.6).

3 Square integrable martingales

In later sections we will use properties of (continuous) *square integrable martingales*. Throughout this section we work with a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\mathbb{F}$ that satisfies the usual conditions.

3.1 Structural properties

Definition 3.1. A right-continuous martingale X is called *square integrable* if $\mathbb{E}\, X_t^2 < \infty$ for all $t \geq 0$. By $\mathcal{M}^2$ we denote the class of all square integrable martingales starting at zero and by $\mathcal{M}_c^2$ its subclass of *a.s continuous* square integrable martingales.

To study properties of $\mathcal{M}^2$ and $\mathcal{M}_c^2$ we endow these spaces with a metric and under additional assumptions with a norm.

Definition 3.2. Let $X \in \mathcal{M}^2$. We define for each $t \geq 0$

$$||X||_t = (\mathbb{E}\, X_t^2)^{1/2}, \tag{3.1}$$

and $||X||_\infty = \sup ||X||_t$. X is called bounded in L^2 if $\sup ||X||_t < \infty$. For all $X \in \mathcal{M}^2$ we also define

$$||X|| = \sum_{n=1}^{\infty} 2^{-n}(||X||_n \wedge 1) \tag{3.2}$$

and for all $X, Y \in \mathcal{M}^2$

$$d(X, Y) = ||X - Y||. \tag{3.3}$$

Remark 3.3. If we identify processes that are indistinguishable, then $(\mathcal{M}^2, d)$ becomes a metric space and on the subclass of martingales bounded in L^2 the operator $||\cdot||_\infty$ becomes a norm (see exercise 3.1).

Proposition 3.4 *The metric space $(\mathcal{M}^2, d)$ is complete and $\mathcal{M}_c^2$ is a closed (w.r.t. to the metric d) subspace of $\mathcal{M}^2$ and thus complete as well.*

Proof. Let (X^m) be a Cauchy sequence in $(\mathcal{M}^2, d)$. Then for each fixed t the sequence (X_t^m) is Cauchy in the complete space $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ and thus has a limit, X_t say (which is the L^1-limit as well). We show that for $s \leq t$ one has $X_s = \mathbb{E}\,[X_t|\mathcal{F}_s]$. Consider for any $A \in \mathcal{F}_s$ the expectations $\mathbb{E}\, 1_A X_s$ and $\mathbb{E}\, 1_A X_t$. The former is the limit of $\mathbb{E}\, 1_A X_s^m$ and the latter the limit of $\mathbb{E}\, 1_A X_t^m$. But since each X^m is a martingale, one has $\mathbb{E}\, 1_A X_s^m = E 1_A X_t^m$, which yields the desired result. Choosing a right-continuous modification of the process X finishes the proof of the first assertion.
The proof of the second assertion is as follows. Let (X^m) be a sequence in $\mathcal{M}_c^2$, with limit $X \in \mathcal{M}^2$ say. We have to show that X is (almost surely) continuous. Using inequality (2.1), we have for every $\varepsilon > 0$ that

$$\mathbb{P}(\sup_{t \leq T} |X_t^m - X_t| > \varepsilon) \leq \frac{1}{\varepsilon^2} \mathbb{E}\,(X_T^m - X_T)^2 \to 0,\ m \to \infty.$$

Hence, for every $k \in \mathbb{N}$ there is m_k such that $\mathbb{P}(\sup_{t \leq T} |X_t^{m_k} - X_t| > \varepsilon) \leq 2^{-k}$. By the Borel-Cantelli lemma one has $\mathbb{P}(\liminf\{\sup_{t \leq T} |X_t^{m_k} - X_t| \leq \varepsilon\}) = 1$. Hence for all $T > 0$ and for almost all ω the functions $t \mapsto X^{m_k}(t, \omega) : [0, T] \to \mathbb{R}$ converge uniformly, which entails that limit functions $t \mapsto X(t, \omega)$ are continuous on each interval $[0, T]$. A.s. unicity of the limit on all these intervals with integer T yields a.s. continuity of X on $[0, \infty)$. $\square$

3.2 Quadratic variation

Of a martingale $X \in \mathcal{M}^2$ we know that X^2 is a nonnegative submartingale. Therefore we can apply the Doob-Meyer decomposition, see theorem 2.14 and exercise 2.7, to obtain

$$X^2 = A + M, \tag{3.4}$$

where A is a natural increasing process and M a martingale that starts in zero. We also know from theorem 2.15 that A and M are continuous if $X \in \mathcal{M}_c^2$.

Definition 3.5. For a process X in $\mathcal{M}^2$ the process A in the decomposition (3.4) is called the quadratic variation process and it is denoted by $\langle X \rangle$. So, $\langle X \rangle$ is the unique natural increasing process that makes $X^2 - \langle X \rangle$ a martingale.

Proposition 3.6 *Let X be a martingale in $\mathcal{M}^2$ and T a stopping time. Then the processes $\langle X^T \rangle$ and $\langle X \rangle^T$ are indistinguishable.*

Proof. This is an immediate consequence of proposition 2.16. $\square$

The term quadratic variation process for the process $\langle X \rangle$ will be explained for $X \in \mathcal{M}_c^2$ in proposition 3.8 below. We first introduce some notation. Let $\Pi = \{t_0, \ldots, t_n\}$ be a partition of the interval $[0, t]$ for some $t > 0$ with $0 = t_0 < \cdots < t_n = t$. For a process X we define

$$V_t(X, \Pi) = \sum_{j=1}^{n} (X_{t_j} - X_{t_{j-1}})^2.$$

Notice that $V_t(X, \Pi)$ is $\mathcal{F}_t$-measurable if X is adapted. The mesh $\mu(\Pi)$ of the partition is defined by $\mu(\Pi) = \max\{|t_j - t_{j-1}|, j = 1, \ldots, n\}$.

For any process X we denote by $m_T(X, \delta)$ the modulus of continuity:

$$m_T(X, \delta) = \max\{|X_t - X_s| : 0 \leq s, t \leq T, |s - t| < \delta\}.$$

If X is an a.s. continuous process, then for all $T > 0$ it holds that $m_T(X, \delta) \to 0$ a.s. for $\delta \to 0$.

The characterization of the quadratic variation process of proposition 3.8 below will be proved using the following lemma.

Lemma 3.7 *Let $M \in \mathcal{M}^2$ and $t > 0$. Then for all $t \geq s$ one has*

$$\mathbb{E}\,[(M_t - M_s)^2|\mathcal{F}_s] = \mathbb{E}\,[M_t^2 - M_s^2|\mathcal{F}_s] \tag{3.5}$$

and

$$\mathbb{E}\,[(M_t - M_s)^2 - \langle M\rangle_t + \langle M\rangle_s|\mathcal{F}_s] = 0 \tag{3.6}$$

For all $t_1 \leq \cdots \leq t_n$ it holds that

$$\mathbb{E}\,[\sum_{j=i+1}^{n}(M_{t_j} - M_{t_{j-1}})^2|\mathcal{F}_{t_i}] = \mathbb{E}\,[M_{t_n}^2 - M_{t_i}^2|\mathcal{F}_{t_i}]. \tag{3.7}$$

If M is bounded by a constant $K > 0$, then

$$\mathbb{E}\,[\sum_{j=i+1}^{n}(M_{t_j} - M_{t_{j-1}})^2|\mathcal{F}_{t_i}] \leq K^2. \tag{3.8}$$

Proof. Equation (3.5) follows by a simple computation and the martingale property of M. Then equation (3.7) follows simply by iteration and reconditioning. Equation (3.6) follows from (3.5) and the definition of the quadratic variation process. □

Proposition 3.8 *Let X be in $\mathcal{M}_c^2$. Then $V_t(X,\Pi)$ converges in probability to $\langle X\rangle_t$ as $\mu(\Pi) \to 0$: for all $\varepsilon > 0, \eta > 0$ there exists a $\delta > 0$ such that $\mathbb{P}(|V_t(X,\Pi) - \langle X\rangle_t| > \eta) < \varepsilon$ whenever $\mu(\Pi) < \delta$.*

Proof. Supposing that we had already proven the assertion for bounded continuous martingales with bounded quadratic variation, we argue as follows. Let X be an arbitrary element of $\mathcal{M}_c^2$ and define for each $n \in \mathbb{N}$ the stopping times $T^n = \inf\{t \geq 0 : |X_t| \geq n \text{ or } \langle X\rangle_t \geq n\}$. The T^n are stopping times in view of proposition 1.10. Then

$$\{|V_t(X,\Pi) - \langle X\rangle_t| > \eta\} \subset \{T^n \leq t\} \cup \{|V_t(X,\Pi) - \langle X\rangle_t| > \eta, T^n > t\}.$$

The probability on the first event on the right hand side obviously tends to zero for $n \to \infty$. The second event is contained in $\{|V_t(X^{T_n},\Pi) - \langle X\rangle_t^{T^n}| > \eta\}$. In view of proposition 3.6 we can rewrite it as $\{|V_t(X^{T_n},\Pi) - \langle X^{T^n}\rangle_t| > \eta\}$ and its probability can be made arbitrarily small, since the processes X^{T^n} and $\langle X\rangle^{T^n} = \langle X^{T^n}\rangle$ are both bounded, by what we supposed at the beginning of the proof.

We now show that the proposition holds for bounded continuous martingales with bounded quadratic variation. Actually we will show that we even have L^2-convergence and so we consider $\mathbb{E}\,(V_t(X,\Pi) - \langle X\rangle_t)^2$. In the second inequality below we use (3.6) to get rid of the expectation of the cross terms after expanding the square. We have

$$
\begin{aligned}
&\mathbb{E}\,(V_t(X,\Pi)-\langle X\rangle_t)^2 \\
&\quad= \mathbb{E}\,\Big(\sum_k (X_{t_k}-X_{t_{k-1}})^2-(\langle X\rangle_{t_k}-\langle X\rangle_{t_{k-1}})\Big)^2 \\
&\quad= \mathbb{E}\,\sum_k \big((X_{t_k}-X_{t_{k-1}})^2-(\langle X\rangle_{t_k}-\langle X\rangle_{t_{k-1}})\big)^2 \\
&\quad\leq 2\sum_k \big(\mathbb{E}\,(X_{t_k}-X_{t_{k-1}})^4+\mathbb{E}\,(\langle X\rangle_{t_k}-\langle X\rangle_{t_{k-1}})^2\big) \\
&\quad\leq 2\sum_k \mathbb{E}\,(X_{t_k}-X_{t_{k-1}})^4+\mathbb{E}\,\big(m_t(\langle X\rangle,\mu(\Pi))\langle X\rangle_t\big).
\end{aligned}
$$

The second term in the last expression goes to zero in view of the bounded convergence theorem and continuity of $\langle X\rangle$ when $\mu(\Pi)\to 0$. We henceforth concentrate on the first term. First we bound the sum by

$$
\mathbb{E}\,\sum_k \big((X_{t_k}-X_{t_{k-1}})^2 m_t(X,\mu(\Pi))^2\big)=\mathbb{E}\,(V_t(X,\Pi)m_t(X,\mu(\Pi)))^2),
$$

which is by the Schwartz inequality less then $\big(\mathbb{E}\,V_t(X,\Pi)^2\mathbb{E}\,m_t(X,\mu(\Pi))^4\big)^{1/2}$. By application of the dominated convergence theorem the last expectation tends to zero if $\mu(\Pi)\to 0$, so we have finished the proof as soon as we can show that $\mathbb{E}\,V_t(X,\Pi)^2$ stays bounded. Let $K>0$ be an upper bound for both X and $\langle X\rangle$. Then, a two fold application of (3.8) leads to the inequalities below,

$$
\begin{aligned}
&\mathbb{E}\,\Big(\sum_k (X_{t_k}-X_{t_{k-1}})^2\Big)^2 \\
&\quad= \mathbb{E}\,\sum_k (X_{t_k}-X_{t_{k-1}})^4+2\mathbb{E}\,\sum_{i<j}\mathbb{E}\,[(X_{t_j}-X_{t_{j-1}})^2(X_{t_i}-X_{t_{i-1}})^2|\mathcal{F}_{t_i}] \\
&\quad\leq 4K^2\mathbb{E}\,\sum_k (X_{t_k}-X_{t_{k-1}})^2+2K^2\mathbb{E}\,\sum_i (X_{t_i}-X_{t_{i-1}})^2 \\
&\quad\leq 6K^4.
\end{aligned}
$$

This finishes the proof. □

Having defined the quadratic variation (in definition 3.5) of a square integrable martingale, we can also define the quadratic covariation (also called cross-variation) between two square integrable martingales X and Y. It is the process $\langle X,Y\rangle$ defined through the polarization formula

$$
\langle X,Y\rangle=\frac{1}{4}(\langle X+Y\rangle-\langle X-Y\rangle). \tag{3.9}
$$

Notice that the quadratic covariation process is a process of a.s. bounded variation. Moreover

Proposition 3.9 *For all $X, Y \in \mathcal{M}^2$ the process $\langle X, Y\rangle$ is the unique process that can be written as the difference of two natural processes such that the difference of XY and such a process is a martingale. Moreover if $X, Y \in \mathcal{M}^2_c$ the process $\langle X, Y\rangle$ is the unique continuous process of bounded variation such that the difference of XY and such a process is a martingale.*

Proof. Exercise 3.4. □

3.3 Exercises

3.1 Let X be a martingale that is bounded in L^2. Then X_∞ exists as an a.s. limit of X_t, for $t \to \infty$. Show that $||X||_\infty = (\mathbb{E}\, X_\infty^2)^{1/2}$. Show also that $\{X \in \mathcal{M}^2 : ||X||_\infty < \infty\}$ is a vector space and that $||\cdot||_\infty$ is a norm on this space under the usual identification that two processes are "the same", if they are indistinguishable.

3.2 Give an example of a martingale (not continuous of course), for which the result of proposition 3.8 doesn't hold. *Hint:* embed a very simple discrete time martingale in continuous time, by defining it constant on intervals of the type $[n, n+1)$.

3.3 Show the following statements.
(a) $\langle X, Y\rangle = \frac{1}{2}(\langle X+Y\rangle - \langle X\rangle - \langle Y\rangle)$.
(b) The quadratic covariation is a bilinear form.
(c) The Schwartz inequality $\langle X, Y\rangle^2 \leq \langle X\rangle\langle Y\rangle$ holds.
(d) If $V(s,t]$ denotes the total variation of $\langle X, Y\rangle$ over the interval $(s,t]$, then a.s. for all $t \geq s$

$$V(s,t] \leq \frac{1}{2}(\langle X\rangle_t + \langle Y\rangle_t - \langle X\rangle_s - \langle Y\rangle_s). \tag{3.10}$$

3.4 Prove proposition 3.9

3.5 Let $X \in \mathcal{M}^2_c$ and let T be a stopping time (not necessarily finite). If $\langle X\rangle_T = 0$, then $\mathbb{P}(X_{t\wedge T} = 0, \forall t \geq 0) = 1$.

3.6 If X is a martingale w.r.t. some filtration, it is also a martingale w.r.t. $\mathbb{F}^X$. Let X and Y be independent processes on some $(\Omega, \mathcal{F}, \mathbb{P})$ and assume that both are right-continuous martingales. Let $\mathcal{F}^0_t = \sigma(X_s, Y_s, s \leq t)$ and let $\mathbb{F}$ be the filtration of σ-algebras $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}^0_s \vee \mathcal{N}$, where $\mathcal{N}$ is the collections of all $\mathbb{P}$-null sets of $\mathcal{F}$. Show that X and Y are martingales w.r.t. $\mathbb{F}$ and that $\langle X, Y\rangle = 0$.

3.7 Let X be a nonnegative continuous process and A a continuous increasing process such $\mathbb{E}\, X_T \leq \mathbb{E}\, A_T$ for all bounded stopping times T. Define the process X^* by $X^*_t = \sup_{0\leq s\leq t} X_s$. Let now T be any stopping time. Show that

$$\mathbb{P}(X^*_T \geq \varepsilon, A_T < \delta) \leq \frac{1}{\varepsilon}\mathbb{E}\,(\delta \wedge A_T).$$

Deduce that

$$\mathbb{P}(X_T^* \geq \varepsilon) \leq \frac{1}{\varepsilon}\mathbb{E}\,(\delta \wedge A_T) + \mathbb{P}(A_T \geq \delta).$$

Finally, let $X = M^2$, where $M \in \mathcal{M}_c^2$. What is a good candidate for the process A in this case?

3.8 Formulate the analogue of proposition 3.8 for the quadratic covariation process of two martingales in $\mathcal{M}_c^2$. Prove it by using the assertion of this proposition.

3.9 Let B be standard Brownian motion. Show that $\langle B \rangle_t = t$. Let X be the square integrable martingale defined by $X_t = B_t^2 - t$. Show (use proposition 3.8) that $\langle X \rangle_t = 4\int_0^t B_s^2\, ds$.

3.10 Let N be a Poisson process, with intensity λ. Let $M_t = N_t - \lambda t$. Show that $\langle M \rangle_t = \lambda t$ for all $t \geq 0$. Show also that $V_t(M, \Pi) \to N_t$ a.s., when $\mu(\Pi) \to 0$.

4 Local martingales

In proofs of results of previous sections we reduced a relatively difficult problem by stopping at judiciously chosen stopping times to an easier problem. This is a standard technique and it also opens the way to define wider classes of processes than the ones we have used thusfar and that still share many properties of the more restricted classes. In this section we assume that the underlying filtration satisfies the usual conditions.

4.1 Localizing sequences and local martingales

Definition 4.1. A sequence of stopping times T^n is called a *fundamental* or a *localizing* sequence if $T^n \geq T^{n-1}$ for all $n \geq 1$ and if $\lim_{n\to\infty} T^n = \infty$ a.s. If $\mathcal{C}$ is a class of processes satisfying a certain property, then (usually) by $\mathcal{C}^{loc}$ we denote the class of processes X for which there exists a fundamental sequence of stopping times T^n such that all stopped processes X^{T^n} belong to $\mathcal{C}$.

In the sequel we will take for $\mathcal{C}$ various classes consisting of martingales with a certain property. The first one is defined in

Definition 4.2. An adapted right-continuous process X is called a local martingale if there exists a localizing sequence of stopping times T^n such that for every n the process $X^{T^n} 1_{\{T^n>0\}} = \{X_{T^n\wedge t} 1_{\{T^n>0\}} : t \geq 0\}$ is a uniformly integrable martingale. The class of local martingales X with $X_0 = 0$ a.s. is denoted by $\mathcal{M}^{loc}$. The subclass of continuous local martingales X with $X_0 = 0$ a.s. is denoted by $\mathcal{M}^{loc}_c$.

Proposition 4.3 *Let X be an adapted right-continuous process such that $X_0 = 0$ a.s. Then the following statements are equivalent.*
(i) X is a local martingale.
(ii) There exists a localizing sequence (T^n) such that the processes X^{T^n} are uniformly integrable martingales.
(iii) There exists a localizing sequence (T^n) such that the processes X^{T^n} are martingales.

Proof. Exercise 4.1. □

Remark 4.4. By the optional stopping theorem, every martingale is a local martingale. This can be seen by choosing $T^n = n$ for all n.

4.2 Continuous local martingales

In the sequel we will deal mainly with continuous processes. The main results are as follows.

Proposition 4.5 *If X is a continuous local martingale with $X_0 = 0$ a.s., then there exist a localizing sequence of stopping times T^n such that the processes X^{T^n} are bounded martingales.*

Proof. Exercise 4.2. □

Recall that we called a (right-continuous) martingale square integrable, if $\mathbb{E}\, X_t^2 < \infty$ for all $t \geq 0$. Therefore we call a right-continuous adapted process X a locally square integrable martingale if $X_0 = 0$ and if there exists a localizing sequence of stopping times T^n such that the process X^{T^n} are all square integrable martingales. Obviously, one has that these processes are all local martingales. If we confine ourselves to continuous processes the difference disappears.

Proposition 4.6 *Let X be a continuous local martingale with $X_0 = 0$ a.s. Then it is also locally square integrable.*

Proof. This follows immediately from proposition 4.5. □

Some properties of local martingales are given in

Proposition 4.7 *(i) A local martingale of class* DL *is a martingale.*
(ii) Any nonnegative local martingale is a supermartingale.

Proof. Exercise 4.3. □

Although local martingales in general don't have a finite second moment, it is possible to define a quadratic variation process. We do this for continuous local martingales only and the whole procedure is based on the fact that for a localizing sequence (T^n) the processes X^{T^n} have a quadratic variation process in the sense of section 3.2.

Proposition 4.8 *Let $X \in \mathcal{M}_c^{loc}$. Then there exists a unique (up to indistinguishability) continuous process $\langle X \rangle$ with a.s. increasing paths such that $X^2 - \langle X \rangle \in \mathcal{M}_c^{loc}$.*

Proof. Choose stopping times T^n such that the stopped processes X^{T^n} are bounded martingales. This is possible in view of proposition 4.5. Then, for each n there exists a unique natural (even continuous) increasing process A^n such that $(X^{T^n})^2 - A^n$ is a martingale. For $n > m$ we have that $(X^{T^n})^{T^m} = X^{T^m}$. Hence, by the uniqueness of the Doob-Meyer decomposition, one has $(A^n)^{T^m} = A^m$. So we can unambiguously define $\langle X \rangle$ by setting $\langle X \rangle_t = A_t^n$ if $t < T^n$. Moreover, we have

$$X_{t \wedge T^n}^2 - \langle X \rangle_{t \wedge T^n} = (X_t^{T^n})^2 - A_t^n,$$

which shows that for all n the process $(X^2)^{T^n} - \langle X \rangle^{T^n}$ is a martingale, and thus that $X^2 - \langle X \rangle$ is a (continuous) local martingale. □

Corollary 4.9 *If X and Y belong to $\mathcal{M}_c^{loc}$, then there exists a unique (up to indistinguishability) continuous process $\langle X, Y \rangle$ with paths of bounded variation a.s. such that $XY - \langle X, Y \rangle \in \mathcal{M}_c^{loc}$.*

Proof. Exercise 4.4. □

4.3 Exercises

4.1 Prove proposition 4.3.

4.2 Prove proposition 4.5.

4.3 Prove proposition 4.7.

4.4 Prove corollary 4.9.

4.5 Let $M \in \mathcal{M}_c^{loc}$ and let S be a stopping time. Put $X_t = M_t^2$ and define $X_\infty = \liminf_{t\to\infty} X_t$ (and $\langle M \rangle_\infty = \lim_{t\to\infty} \langle M \rangle_t$). Show that $\mathbb{E}\, X_S \leq \mathbb{E}\, \langle M \rangle_S$.

4.6 Let $M \in \mathcal{M}_c^2$ satisfy the property $\mathbb{E}\, \langle M \rangle_\infty < \infty$. Deduce from exercise 4.5 that $\{M_S : S \text{ a finite stopping time}\}$ is uniformly integrable and hence that M_∞ is well defined as a.s. limit of M_t. Show that $\mathbb{E}\, M_\infty^2 = \mathbb{E}\, \langle M \rangle_\infty$.

4.7 Let $M \in \mathcal{M}^2$ have independent and stationary increments, the latter meaning that $M_{t+h} - M_t$ has the same distribution as $M_{s+h} - M_s$ for all $s, t, h > 0$. Show that $\langle M \rangle_t = t\mathbb{E}\, M_1^2$.

4.8 Let X be an adapted process and T a stopping time. Show that $X^T 1_{\{T>0\}}$ is a uniformly integrable martingale iff $X_0 1_{\{T>0\}}$ is integrable and $X^T - X_0$ is a uniformly integrable martingale. (*The latter property for a fundamental sequence of stopping times is also used as definition of local martingale.*)

5 Spaces of progressive processes

Throughout this section we work with a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that is endowed with a filtration $\mathbb{F}$ that satisfies the usual conditions. All properties below that are defined relative to a filtration (such as adaptedness) are assumed to be defined in terms of this filtration.

5.1 Doléans measure

In this section we often work with (Lebesgue-Stieltjes) integrals w.r.t. a process of finite variation A that satisfy $A_0 = 0$. So, for an arbitrary process X with the right measurability properties we will look at integrals of the form

$$\int_{[0,T]} X_t(\omega)\, dA_t(\omega), \tag{5.1}$$

with t the integration variable and where this integral has to be evaluated ω-wise. We follow the usual convention for random variables by omitting the variable ω. But in many cases we also omit the integration variable t and hence an expression like (5.1) will often be denoted by

$$\int_{[0,T]} X\, dA.$$

We also use the notation $\int_0^T$ instead of $\int_{[0,T]}$ if the process A has a.s. continuous paths. Let M be a square integrable continuous martingale ($M \in \mathcal{M}_c^2$). Recall that $\langle M \rangle$ is the unique continuous increasing process such that $M^2 - \langle M \rangle$ is a martingale.

Definition 5.1. The Doléans measure μ_M on $([0,\infty) \times \Omega, \mathcal{B}([0,\infty)) \times \mathcal{F})$ is defined by

$$\mu_M(A) = \mathbb{E} \int_0^\infty 1_A(t,\omega)\, d\langle M \rangle_t(\omega). \tag{5.2}$$

For a measurable, adapted process X we define for every $T \in [0,\infty)$

$$||X||_T = (\mathbb{E} \int_0^T X_t^2\, d\langle M \rangle_t)^{1/2} = (\int X^2 1_{[0,T]\times\Omega}\, d\mu_M)^{1/2} \tag{5.3}$$

and

$$||X|| = \sum_{n=1}^\infty 2^{-n}(1 \wedge ||X||_n). \tag{5.4}$$

We will also use

$$||X||_\infty = (\mathbb{E} \int_0^\infty X_t^2\, d\langle M \rangle_t)^{1/2} = (\int X^2\, d\mu_M)^{1/2} \tag{5.5}$$

Two measurable, adapted processes X and Y are called (M-)equivalent if $||X - Y|| = 0$. By $\mathcal{P}$ we denote the class of progressive processes X for which $||X||_T < \infty$ for all $T \in [0,\infty)$.

Remark 5.2. Notice that, with time restricted to $[0, T]$, the function $|| \cdot ||_T$ defines an L^2-norm on the space of measurable, adapted processes if we identify equivalent processes. Similarly, $d(X, Y) := ||X - Y||$ defines a metric. Notice also that X and Y are equivalent iff $\int_0^T (X-Y)^2 \, d\langle M \rangle = 0$ a.s. for all $T \in [0, \infty)$.

In addition to the class $\mathcal{P}$ introduced above we also need the classes $\mathcal{P}_T$ for $T \in [0, \infty]$. These are defined in

Definition 5.3. For $T < \infty$ the class $\mathcal{P}_T$ is the set of processes X in $\mathcal{P}$ for which $X_t = 0$ if $t > T$. The class $\mathcal{P}_\infty$ is the subclass of processes $X \in \mathcal{P}$ for which $\mathbb{E} \int_0^\infty X^2 \, d\langle M \rangle < \infty$.

Remark 5.4. A process X belongs to $\mathcal{P}_T$ iff $X_t = 0$ for $t > T$ and $||X||_T < \infty$.

Remark 5.5. All the classes, norms and metrics above depend on the martingale M. When other martingales than M play a role in a certain context, we emphasize this dependence by e.g. writing $\mathcal{P}_T(M)$, $\mathcal{P}(M)$, etc.

Proposition 5.6 *For $T \leq \infty$ the class $\mathcal{P}_T$ is a Hilbert space with inner product given by*

$$(X, Y)_T = \mathbb{E} \int_0^T XY \, d\langle M \rangle,$$

if we identify two processes X and Y that satisfy $||X - Y||_T = 0$.

Proof. Let $T < \infty$ and let (X^n) be a Cauchy sequence in $\mathcal{P}_T$. Since $\mathcal{P}_T$ is a subset of the Hilbert space $L^2([0, T] \times \Omega, \mathcal{B}([0, T]) \times \mathcal{F}_T, \mu_M)$, the sequence (X^n) has a limit X in this space. We have to show that $X \in \mathcal{P}_T$, but it is not clear that the limit process X is progressive (a priori we can only be sure that X is $\mathcal{B}([0, T]) \times \mathcal{F}_T$-measurable). We will replace X with an equivalent process as follows. First we select a subsequence X^{n_k} that converges to X almost everywhere w.r.t. μ_M. We set $Y_t(\omega) = \limsup_{k \to \infty} X_t^{n_k}(\omega)$. Then Y is a progressive process, since the X^n are progressive and $||X - Y||_T = 0$, so Y is equivalent to X and $||X_n - Y||_T \to 0$. For $T = \infty$ the proof is similar. $\square$

5.2 Local variants

In this subsection we enlarge the classes $\mathcal{P}_T$ by dropping the requirement that the expectations are finite and by relaxing the condition that $M \in \mathcal{M}_c^2$. We have

Definition 5.7. Let $M \in \mathcal{M}_c^{loc}$. The class $\mathcal{P}^*$ is the equivalence class of progressive processes X such that $\int_0^T X^2 \, d\langle M \rangle < \infty$ a.s. for every $T \in [0, \infty)$.

Remark 5.8. If $M \in \mathcal{M}_c^{loc}$, there exist a fundamental sequence of stopping times R^n such that $M^{R^n} \in \mathcal{M}_c^2$. If we take $X \in \mathcal{P}^*$, then the bounded stopping times $S^n = n \wedge \inf\{t \geq 0 : \int_0^t X^2 \, d\langle M \rangle \geq n\}$ also form a fundamental sequence. Consider then the stopping times $T^n = R^n \wedge S^n$. These form a fundamental sequence as well. Moreover, $M^{T^n} \in \mathcal{M}_c^2$ and the processes X^n defined by $X_t^n = X_t 1_{\{t \leq T^n\}}$ belong to $\mathcal{P}(M^{T^n})$.

5.3 Simple processes

We start with a definition.

Definition 5.9. A process X is called simple if there exists a strictly increasing sequence of real numbers t_n with $t_0 = 0$ and $t_n \to \infty$, a uniformly bounded sequence of random variables ξ_n with the property that ξ_n is $\mathcal{F}_{t_n}$-measurable for each n such that

$$X_t = \xi_0 1_{\{0\}}(t) + \sum_{n=0}^{\infty} \xi_n 1_{(t_n,t_{n+1}]}(t),\ t \geq 0.$$

The class of simple processes is denoted by $\mathcal{S}$.

Remark 5.10. Notice that simple processes are progressive and bounded. If $M \in \mathcal{M}_c^2$, then a simple process X belongs to $\mathcal{P} = \mathcal{P}(M)$.

The following lemma is crucial for the construction of the stochastic integral.

Lemma 5.11 *Let X be a bounded progressive process. Then there exists a sequence of simple processes X^n such that for all $T \in [0, \infty)$ one has*

$$\lim_{n\to\infty} \mathbb{E} \int_0^T (X_t^n - X_t)^2 \, dt = 0. \tag{5.6}$$

Proof. Suppose that we had found for each $T \in [0, \infty)$ a sequence of simple processes $X^{n,T}$ (depending on T) such that (5.6) holds. Then for all integers n there exist integers m_n such that

$$\mathbb{E} \int_0^n (X_t^{m_n,n} - X_t)^2 \, dt \leq \frac{1}{n}.$$

One verifies that the sequence with elements $X^n = X^{m_n,n}$ then has the asserted property. Therefore, we will keep T fixed in the remainder of the proof and construct a sequence of simple processes X^n for which (5.6) holds. This is relatively easy if X is continuous. Consider the sequence of approximating processes that we used in the proof of 1.7. This sequence has the desired property in view of the bounded convergence theorem.
If X is merely progressive (but bounded), we proceed by first approximating it with continuous processes to which we apply the preceding result. For $t \leq T$ we define the bounded continuous ('primitive') process F by $F_t(\omega) = \int_0^t X_s(\omega)\, ds$ for $t \geq 0$ and $F_t(\omega) = 0$ for $t < 0$ and for each integer m

$$Y_t^m(\omega) = m(F_t(\omega) - F_{t-1/m}(\omega)).$$

By the fundamental theorem of calculus, for each ω, the set of t for which $Y_t^m(\omega)$ does not converge to $X_t(\omega)$ for $m \to \infty$ has Lebesgue measure zero. Hence, by dominated convergence, we also have $\mathbb{E} \int_0^T (Y_t^m - X_t)^2 \, dt \to 0$. By the preceding result we can approximate each of the continuous Y^m by simple processes $Y^{m,n}$ in the sense that $\mathbb{E} \int_0^T (Y_t^m - Y_t^{m,n})^2 \, dt \to 0$. But then we can also find a sequence of simple processes $X^m = Y^{m,n_m}$ for which $\mathbb{E} \int_0^T (X_t^m - X_t)^2 \, dt \to 0$. $\square$

The preceding lemma enables us to prove

Proposition 5.12 *Let M be in $\mathcal{M}_c^2$ and assume that there exists a progressive nonnegative process f such that $\langle M \rangle$ is indistinguishable from $\int_0^\cdot f_s \, ds$ (the process $\langle M \rangle$ is said to be a.s. absolutely continuous). Then the set $\mathcal{S}$ of simple processes is dense in $\mathcal{P}$ with respect to the metric d defined in remark 5.2.*

Proof. Let $X \in \mathcal{P}$ and assume that X is bounded. By lemma 5.11 we can find simple processes X^n such that for all $T > 0$ it holds that $\mathbb{E} \int_0^T (X^n - X)^2 \, dt \to 0$. But then we can select a subsequence (X^{n_k}) such that $X^{n_k} \to X$ for $dt \times \mathbb{P}$-almost all (t, ω). By the dominated convergence theorem we then also have for all $T > 0$ that $\mathbb{E} \int_0^T (X^{n_k} - X)^2 f_t \, dt \to 0$.
If X is not bounded we truncate it and introduce the processes $X^n = X 1_{\{|X| \leq n\}}$. Each of these can be approximated by simple processes in view of the previous case. The result then follows upon noticing that $\mathbb{E} \int_0^T (X^n - X)^2 \, d\langle M \rangle = \mathbb{E} \int_0^T X^2 1_{\{|X| > n\}} \, d\langle M \rangle \to 0$. $\square$

Remark 5.13. The approximation results can be strengthened. For instance, in the previous lemma we didn't use progressive measurability. The space $\mathcal{S}$ is also dense in the set of measurable processes. Furthermore, if we drop the requirement that the process $\langle M \rangle$ is a.s. absolutely continuous, the assertion of proposition 5.12 is still true, but the proof is much more complicated. For most, if not all, of our purposes the present version is sufficient.

5.4 Exercises

5.1 Let X be a simple process given by $X_t = \sum_{k=1}^\infty \xi_{k-1} 1_{(t_{k-1}, t_k]}(t)$ and let M be an element of $\mathcal{M}_c^2$. Consider the discrete time process V defined by $V_n = \sum_{k=1}^n \xi_{k-1}(M_{t_k} - M_{t_{k-1}})$. Show that V is a martingale w.r.t. an appropriate filtration $\mathbb{G} = (\mathcal{G}_n)$ in discrete time. Compute $\sum_{k=1}^n \mathbb{E}\,[(V_k - V_{k-1})^2 | \mathcal{G}_{k-1}]$. Compute also $\int_0^{t_n} X_t^2 \, d\langle M \rangle_t$.

5.2 Let W be a Brownian motion and let for each n a partition $\Pi^n = \{0 = t_0^n, \ldots, t_n^n = T\}$ of $[0, T]$ be given with $\mu(\Pi^n) \to 0$ for $n \to \infty$. Let $h_n : \mathbb{R} \to \mathbb{R}$ be defined by $h_n(x) = x 1_{[-n,n]}(x)$ and put

$$W_t^n = \sum_{j=1}^n h_n(W_{t_{j-1}^n}) 1_{(t_{j-1}^n, t_j^n]}(t).$$

Then $W^n \in \mathcal{S}$ for all n. Show that $W^n \to W$ in $\mathcal{P}_T(W)$.

6 Stochastic Integral

In previous sections we have already encountered integrals where both the integrand and the integrator were stochastic processes, e.g. in the definition of a natural process. In all these cases the integrator was an increasing process or, more general, a process with paths of bounded variation over finite intervals. In the present section we will consider integrals where the integrator is a continuous martingale. Except for trivial exceptions, these have paths of unbounded variation so that a pathwise definition of these integrals in the Lebesgue-Stieltjes sense cannot be given. As a matter of fact, if one aims at a sensible pathwise definition of such an integral, one finds himself in a (seemingly) hopeless position in view of proposition 6.1 below. For integrals, over the interval $[0, 1]$ say, defined in the Stieltjes sense we know that the sums

$$S_\Pi h = \sum h(t_k)(g(t_{k+1}) - g(t_k)) \tag{6.1}$$

converge for continuous h and g of bounded variation, if we sum over the elements of partitions $\Pi = \{0 = t_0 < \cdots < t_n = 1\}$ whose mesh $\mu(\Pi)$ tends to zero.

Proposition 6.1 *Suppose that the fixed function g is such that for all continuous functions h one has that $S_\Pi h$ converges, if $\mu(\Pi) \to 0$. Then g is of bounded variation.*

Proof. We view the S_Π as linear operators on the Banach space of continuous functions on $[0, 1]$ endowed with the sup-norm $||\cdot||$. Notice that $|S_\Pi h| \leq ||h|| \sum |g(t_{k+1}) - g(t_k)| = ||h|| V^1(g; \Pi)$, where $V^1(g; \Pi)$ denotes the variation of g over the partition Π. Hence the operator norm $||S_\Pi||$ is less then $V^1(g; \Pi)$. For any partition $\Pi = \{0 = t_0 < \cdots < t_n = 1\}$ we can find (by linear interpolation) a continuous function h_Π (bounded by 1) such that $h_\Pi(t_k) = \mathrm{sgn}(g(t_{k+1}) - g(t_k))$. Then we have $S_\Pi h_\Pi = V^1(g; \Pi)$. It follows that $||S_\Pi|| = V^1(g; \Pi)$. By assumption, for any h we have that the sums $S_\Pi h$ converge if $\mu(\Pi) \to 0$, so that for any h the set with elements $|S_\Pi h|$ (for such Π) is bounded. By the Banach-Steinhaus theorem, also the $||S_\Pi||$ form a bounded set. The result follows since we had already observed that $||S_\Pi|| = V^1(g; \Pi)$. □

The function h in the above proof evaluated at points t_k uses the value of g at a 'future' point t_{k+1}. Excluding functions that use 'future' information, one also says that such functions are anticipating, is one of the ingredients that allow us to nevertheless finding a coherent notion of the (stochastic) integral with martingales as integrator.

6.1 Construction

The basic formula for the construction of the stochastic integral is formula (6.2) below. We consider a process $X \in \mathcal{S}$ as in definition 5.9. The stochastic integral of X w.r.t. $M \in \mathcal{M}_c^2$ is a stochastic process denoted by $I(X)$ (or by

I, or $I(X;M)$, in sections further down also by $X \cdot M$) that at each time t is defined to be the random variable

$$I_t(X) = \sum_{i=0}^{\infty} \xi_i (M_{t\wedge t_{i+1}} - M_{t\wedge t_i}). \tag{6.2}$$

For each $t \in [0,\infty)$ there is a unique $n = n(t)$ such that $t_n \leq t < t_{n+1}$. Hence equation (6.2) then takes the form

$$I_t(X) = \sum_{i=0}^{n-1} \xi_i (M_{t_{i+1}} - M_{t_i}) + \xi_n (M_t - M_{t_n}). \tag{6.3}$$

As in classical Lebesgue integration theory one can show that the given expression for $I_t(X)$ is independent of the chosen representation of X. Notice that (6.3) expresses $I_t(X)$ as a *martingale transform*. With this observation we now present the first properties of the stochastic integral.

Lemma 6.2 *Let $X \in \mathcal{S}$, $M \in \mathcal{M}_c^2$ and $0 \leq s \leq t < \infty$. The following identities are valid.*

$$\begin{aligned} I_0(X) &= 0, \\ \mathbb{E}\,[I_t(X)|\mathcal{F}_s] &= I_s(X)\, a.s., & (6.4) \\ \mathbb{E}\,[(I_t(X) - I_s(X))^2]|\mathcal{F}_s] &= \mathbb{E}\,[\int_s^t X_u^2 \, d\langle M\rangle_u |\mathcal{F}_s]\, a.s. & (6.5) \end{aligned}$$

Proof. The first identity is obvious. To prove equations (6.4) and (6.5) we assume without loss of generality that $t = t_n$ and $s = t_m$. Then $I_t(X) - I_s(X) = \sum_{i=m}^{n-1} \xi_i (M_{t_{i+1}} - M_{t_i})$. Since $\mathbb{E}\,[\xi_i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_{t_i}] = 0$, equation (6.4) follows by re-conditioning. Similarly, we have

$$\begin{aligned} \mathbb{E}\,[(\xi_i(M_{t_{i+1}} - M_{t_i}))^2|\mathcal{F}_{t_i}] &= \xi_i^2 \mathbb{E}\,[(M_{t_{i+1}} - M_{t_i})^2|\mathcal{F}_{t_i}] \\ &= \xi_i^2 \mathbb{E}\,[\langle M\rangle_{t_{i+1}} - \langle M\rangle_{t_i}|\mathcal{F}_{t_i}] \\ &= \mathbb{E}\,[\int_{t_i}^{t_{i+1}} X_u^2 \, d\langle M\rangle_u |\mathcal{F}_{t_i}], \end{aligned}$$

from which (6.5) follows. □

Proposition 6.3 *For $X \in \mathcal{S}$ and $M \in \mathcal{M}_c^2$ the process $I(X)$ is a continuous square integrable martingale with quadratic variation process*

$$\langle I(X)\rangle = \int_0^{\cdot} X_u^2 \, d\langle M\rangle_u, \tag{6.6}$$

and,

$$\mathbb{E}\, I_t(X)^2 = \mathbb{E} \int_0^t X_u^2 \, d\langle M\rangle_u. \tag{6.7}$$

Thus, for each fixed t, we can view $I_t(\cdot)$ as a linear operator on the space of simple processes that are annihilated after t with values in $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ It follows that $||I(X)||_t = ||X||_t$, where the first $||\cdot||_t$ is as in definition 3.2 and the second one as in definition 5.1. Hence I_t is an isometry. With the pair of norms $||\cdot||$ in the same definitions, we also have $||I(X)|| = ||X||$.

Proof. From (6.2) it follows that $I(X)$ is a continuous process and from lemma 6.2 we then obtain that $I(X) \in \mathcal{M}_c^2$ and the expression (6.6) for its quadratic variation process. Then we also have (6.7) and the equality $||I(X)||_t = ||X||_t$ immediately follows, as well as $||I(X)|| = ||X||$. Linearity of I_t can be proved as in Lebesgue integration theory. □

Theorem 6.4 *Let $M \in \mathcal{M}_c^2$. For all $X \in \mathcal{P}$, there exists a unique (up to indistinguishability) process $I(X) \in \mathcal{M}_c^2$ with the property $||I(X^n) - I(X)|| \to 0$ for every sequence (X^n) in $\mathcal{S}$ such that $||X^n - X|| \to 0$. Moreover, its quadratic variation is given by (6.6). This process is called the stochastic integral of X w.r.t. M.*

Proof. First we show existence. Let $X \in \mathcal{P}$. From proposition 5.12 and remark 5.13 we know that there is a sequence of $X^n \in \mathcal{S}$ such that $||X - X^n|| \to 0$. By proposition 6.3 we have for each t that $\mathbb{E}\,(I(X^m)_t - I(X^n)_t)^2 = \mathbb{E}\int_0^t (X^m_s - X^n_s)^2\, d\langle M\rangle_s$ and hence $||I(X^m) - I(X^n)|| = ||X^m - X^n||$. This shows that the sequence $I(X^n)$ is Cauchy in the complete space $\mathcal{M}_c^2$ (proposition 3.4) and thus has a limit in this space. We call it $I(X)$. The limit can be seen to be independent of the particular sequence (X^n) by the following familiar trick. Let (Y^n) be another sequence in $\mathcal{S}$ converging to X. Mix the two sequences as follows: $X^1, Y^1, X^2, Y^2, \ldots$. Also this sequence converges to X. Consider the sequence of corresponding stochastic integrals $I(X^1), I(Y^1), I(X^2), I(Y^2), \ldots$. This sequence has a unique limit in $\mathcal{M}_c^2$ and hence its subsequences $(I(X^n))$ and $(I(Y^n))$ must converge to the same limit, which must then be $I(X)$. The proof of (6.6) is left as exercise 6.7. □

We will frequently need the following extension of proposition 6.3.

Lemma 6.5 *The mapping $I : \mathcal{P} \to \mathcal{M}_c^2$ is linear and $I_T : (\mathcal{P}_T, ||\cdot||) \to L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ an isometry.*

Proof. Exercise 6.8 □

Proposition 6.6 *Let $X \in \mathcal{P}$, $M \in \mathcal{M}_c^2$ and T a stopping time. Then*

$$I(X; M)^T = I(X; M^T)$$

and

$$I(X; M)^T = I(1_{[0,T]}X; M).$$

Proof. If $X \in \mathcal{S}$, the result is trivial. For $X \in \mathcal{P}$, we take $X^n \in \mathcal{S}$ such that $||X^n - X|| \to 0$. Using the assertion for processes in $\mathcal{S}$ and the linearity property of lemma 6.5 we can write

$$I(X; M)^T - I(X; M^T) = I(X - X^n; M)^T - I(X - X^n; M^T).$$

For any $t > 0$ we have $\mathbb{E}\,(I(X - X^n; M)^T_t)^2 = \mathbb{E}\int_0^{t\wedge T}(X_u - X^n_u)^2\, d\langle M\rangle_u \leq \mathbb{E}\int_0^t (X_u - X^n_u)^2\, d\langle M\rangle_u \to 0$. Similarly, $\mathbb{E}\,I(X - X^n; M^T)^2_t = \mathbb{E}\int_0^t (X_u - X^n_u)^2\, d\langle M^T\rangle_u = \mathbb{E}\int_0^t (X_u - X^n_u)^2\, d\langle M\rangle^T_u \leq \mathbb{E}\int_0^t (X_u - X^n_u)^2\, d\langle M\rangle_u \to 0$. This shows the first equality. To prove the second one we write

$$\begin{aligned} &I(X; M)^T_t - I(1_{[0,T]}X; M)_t = \\ &\qquad I(X - 1_{[0,T]}X; M)^T_t - (I(1_{[0,T]}X; M)_t - I(1_{[0,T]}X; M)_{t\wedge T}). \end{aligned}$$

The first term here is equal to $I(1_{(T,\infty]}X; M)^T_t$ which has second moment $\mathbb{E}\int_0^{t\wedge T} 1_{(T,\infty]}X^2\, d\langle M\rangle = 0$. Therefore the first term vanishes for all t. The term in parentheses has second moment $\mathbb{E}\int_{t\wedge T}^t 1_{[0,T]}X^2\, d\langle M\rangle$ (why?), which is zero as well. □

Proposition 6.7 *Let $X, Y \in \mathcal{P}$, $M \in \mathcal{M}^2_c$ and $S \leq T$ be stopping times. Then*

$$\mathbb{E}\,[I(X)_{T\wedge t}|\mathcal{F}_S] = I(X)_{S\wedge t}$$

and

$$\begin{aligned} &\mathbb{E}\,[(I(X)_{T\wedge t} - I(X)_{S\wedge t})(I(Y)_{T\wedge t} - I(Y)_{S\wedge t})|\mathcal{F}_S] = \\ &\qquad \mathbb{E}\,[\int_{S\wedge t}^{T\wedge t} X_u Y_u\, d\langle M\rangle_u|\mathcal{F}_S]. \end{aligned}$$

Proof. The first property follows from corollary 2.9. The second property is first proved for $X = Y$ by applying corollary 2.9 to the martingale $I(X)^2 - \int_0^{\cdot} X^2\, d\langle M\rangle$ and then by polarization. □

6.2 Characterizations and further properties

One of the aims of this section is the computation of the quadratic covariation between the martingales $I(X; M)$ and $I(Y; N)$, where $X \in \mathcal{P}(M)$, $Y \in \mathcal{P}(N)$ and $M, N \in \mathcal{M}^2_c$. For $X, Y \in \mathcal{S}$ this is (relatively) straightforward (exercise 6.6) since the integrals become sums and the result is

$$\langle I(X; M), I(Y; N)\rangle = \int_0^{\cdot} XY\, d\langle M, N\rangle. \tag{6.8}$$

The extension to more general X and Y will be established in a number of steps. The first step is a result known as the Kunita-Watanabe inequality.

Proposition 6.8 *Let $M, N \in \mathcal{M}_c^2$, $X \in \mathcal{P}(M)$ and $Y \in \mathcal{P}(N)$. Let V be the total variation process of the process $\langle M, N\rangle$. Then*

$$\int_0^t |XY|\, dV \leq \Big(\int_0^t |X|^2\, d\langle M\rangle\Big)^{1/2}\Big(\int_0^t |Y|^2\, d\langle N\rangle\Big)^{1/2} \; a.s.\, (t \geq 0). \tag{6.9}$$

Proof. Since $\langle M, N\rangle$ is absolutely continuous w.r.t. V, it is sufficient to prove that $|\int_0^t XY\, d\langle M, N\rangle|$ is a.s. less than the right hand side of (6.9). Indeed, since the density process $d\langle M, N\rangle/dV$ takes its values in $\{-1, 1\}$, we can apply the latter result by noticing that $\int_0^t |XY|\, dV = \int_0^t XY \mathrm{sgn}(XY) \frac{d\langle M,N\rangle}{dV}\, d\langle M, N\rangle$.
As a first result we have that $(\langle M, N\rangle_t - \langle M, N\rangle_s)^2 \leq (\langle M\rangle_t - \langle M\rangle_s)(\langle N\rangle_t - \langle N\rangle_s)$ a.s. (exercise 6.4), which is used to obtain the second inequality in the displayed formulas below. First we assume that X and Y are simple. The final result then follows by taking limits. The process X we can represent on $(0, t]$ as $\sum_k x_k 1_{(t_k, t_{k+1}]}$ with the last $t_n = t$. For Y we similarly have $\sum_k y_k 1_{(t_k, t_{k+1}]}$. It follows that

$$\begin{aligned}
&|\int_0^t XY\, d\langle M, N\rangle| \\
&\leq \sum |x_k||y_k||\langle M, N\rangle_{t_{k+1}} - \langle M, N\rangle_{t_k}| \\
&\leq \sum_k |x_k||y_k|\big((\langle M\rangle_{t_{k+1}} - \langle M\rangle_{t_k})(\langle N\rangle_{t_{k+1}} - \langle N\rangle_{t_k})\big)^{1/2} \\
&\leq \Big(\sum_k x_k^2(\langle M\rangle_{t_{k+1}} - \langle M\rangle_{t_k})\Big)^{1/2}\Big(\sum_k y_k^2(\langle N\rangle_{t_{k+1}} - \langle N\rangle_{t_k})\Big)^{1/2} \\
&= \Big(\int_0^t |X|^2\, d\langle M\rangle\Big)^{1/2}\Big(\int_0^t |Y|^2\, d\langle N\rangle\Big)^{1/2}.
\end{aligned}$$

The rest of the proof is left as exercise 6.14. □

Lemma 6.9 *Let $M, N \in \mathcal{M}_c^2$ and $X \in \mathcal{P}(M)$. Then*

$$\langle I(X; M), N\rangle = \int_0^\cdot X\, d\langle M, N\rangle$$

or, in short hand notation,

$$d\,\langle I(X; M), N\rangle = X\, d\langle M, N\rangle.$$

Proof. Choose $X^n \in \mathcal{S}$ such that $||X^n - X|| \to 0$. Then we can find for every $T > 0$ a subsequence, again denoted by X^n, such that $\int_0^T (X^n - X)^2\, d\langle M\rangle \to 0$ a.s. But then $\langle I(X^n; M) - I(X; M), N\rangle_T^2 \leq \langle I(X^n; M) - I(X; M)\rangle_T \langle N\rangle_T \to 0$ a.s. But for the simple X^n we easily obtain $\langle I(X^n; M), N\rangle_T = \int_0^T X^n\, d\langle M, N\rangle$. Application of the Kunita-Watanabe inequality to $|\int_0^T (X^n - X)\, d\langle M, N\rangle|$ yields the result. □

Proposition 6.10 *Let $M, N \in \mathcal{M}_c^2$, $X \in \mathcal{P}(M)$ and $Y \in \mathcal{P}(N)$. Then equation (6.8) holds.*

Proof. We apply lemma 6.9 twice and get

$$\langle I(X;M), I(Y;N)\rangle = \int_0^{\cdot} Y\, d\langle I(X;M), N\rangle = \int_0^{\cdot} XY\, d\langle M, N\rangle,$$

which is the desired equality. □

We are now in the position to state an important characterization of the stochastic integral. In certain books it is taken as a definition.

Theorem 6.11 *Let $M \in \mathcal{M}_c^2$ and $X \in \mathcal{P}(M)$. Let $I \in \mathcal{M}_c^2$ be such that $\langle I, N\rangle = \int_0^{\cdot} X\, d\langle M, N\rangle$ for all $N \in \mathcal{M}_c^2$. Then I is indistinguishable from $I(X;M)$.*

Proof. Since $\langle I(X;M), N\rangle = \int_0^{\cdot} X\, d\langle M, N\rangle$ we get $\langle I - I(X;M), N\rangle = 0$ for all $N \in \mathcal{M}_c^2$ by subtraction. In particular for $N = I - I(X;M)$. The result follows. □

The characterization is a useful tool in the proof of the following 'chain rule'.

Proposition 6.12 *Let $M \in \mathcal{M}_c^2$, $X \in \mathcal{P}(M)$ and $Y \in \mathcal{P}(I(X;M))$. Then $XY \in \mathcal{P}(M)$ and $I(Y; I(X;M)) = I(XY;M)$ up to indistinguishability.*

Proof. Since $\langle I(X;M)\rangle = \int_0^{\cdot} X^2\, d\langle M\rangle$, it immediately follows that $XY \in \mathcal{P}(M)$. Furthermore, for any martingale $N \in \mathcal{M}_c^2$ we have

$$\begin{aligned}
\langle I(XY;M), N\rangle &= \int_0^{\cdot} XY\, d\langle M, N\rangle \\
&= \int_0^{\cdot} Y\, d\langle I(X;M), N\rangle \\
&= \langle I(Y; I(X;M)), N\rangle.
\end{aligned}$$

It follows from theorem 6.11 that $I(Y; I(X;M)) = I(XY;M)$. □

The construction of the stochastic integral that we have developed here is founded on 'L^2-theory'. We have not defined the stochastic integral w.r.t. a continuous martingale in a pathwise way. Nevertheless, there exists a 'pathwise-uniqueness result'.

Proposition 6.13 *Let $M_1, M_2 \in \mathcal{M}_c^2$, $X_1 \in \mathcal{P}(M_1)$ and $X_2 \in \mathcal{P}(M_2)$. Let T be a stopping time and suppose that M_1^T and M_2^T as well as X_1^T and X_2^T are indistinguishable. Then the same holds for $I(X_1;M_1)^T$ and $I(X_2;M_2)^T$.*

Proof. For any $N \in \mathcal{M}_c^2$ we have that $\langle M_1 - M_2, N\rangle^T = 0$. Hence,

$$\begin{aligned}
\langle I(X_1;M_1)^T - I(X_2;M_2)^T, N\rangle &= \int_0^{\cdot} X_1\, d\langle M_1^T, N\rangle - \int_0^{\cdot} X_2\, d\langle M_2^T, N\rangle \\
&= \int_0^{\cdot} (X_1 - X_2)\, d\langle M_1^T, N\rangle \\
&= \int_0^{\cdot} (X_1 - X_2)\, d\langle M_1, N\rangle^T
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^{\cdot} (X_1 - X_2) 1_{[0,T]} \, d\langle M_1, N\rangle \\
&= 0.
\end{aligned}
$$

The assertion follows by application of theorem 6.11. □

6.3 Integration w.r.t. local martingales

In this section we extend the definition of stochastic integral into two directions. In the first place we relax the condition that the integrator is a martingale (in $\mathcal{M}_c^2$) and in the second place, we put less severe restrictions on the integrand.

In this section M will be a continuous local martingale. We have

Definition 6.14. For $M \in \mathcal{M}_c^{loc}$ the class $\mathcal{P}^* = \mathcal{P}^*(M)$ is defined as the collection of progressive processes X with the property that $\int_0^T X^2 \, d\langle M\rangle < \infty$ a.s. for all $T \geq 0$.

Recall that for local martingales the quadratic (co-)variation processes exist.

Theorem 6.15 *Let $M \in \mathcal{M}_c^{loc}$ and $X \in \mathcal{P}^*(M)$. Then there exists a unique local martingale, denoted by $I(X;M)$, such that for all $N \in \mathcal{M}_c^{loc}$ it holds that*

$$
\langle I(X;M), N\rangle = \int_0^{\cdot} X \, d\langle M, N\rangle. \tag{6.10}
$$

This local martingale is called the stochastic integral of X w.r.t. M. If furthermore $Y \in \mathcal{P}^*(N)$, then equality (6.8) is still valid.

Proof. Define the stopping times S^n as a localizing sequence for M and $T^n = \inf\{t \geq 0 : M_t^2 + \int_0^t X^2 \, d\langle M\rangle \geq n\} \wedge S^n$. Then the T^n also form a localizing sequence, $M^{T^n} \leq n$ and $X^{T^n} \in \mathcal{P}(M^{T^n})$. Therefore the stochastic integrals $I^n := I(X^{T^n}; M^{T^n})$ can be defined as before. It follows from e.g. proposition 6.13 that I^{n+1} and I^n coincide on $[0, T^n]$. Hence we can unambiguously define $I(X;M)_t$ as I_t^n for any n such that $T^n \geq t$. Since $I(X;M)^{T^n} = I^n$ is a martingale, $I(X;M)$ is a local martingale. Furthermore, for any $N \in \mathcal{M}_c^{loc}$ we have $\langle I(X;M), N\rangle^{T^n} = \langle I^n; N\rangle = \int_0^{\cdot} X \, d\langle M, N\rangle^{T^n}$. By letting $n \to \infty$ we obtain (6.10). The uniqueness follows as in the proof of theorem 6.11. □

6.4 Exercises

6.1 Let (X^n) be a sequence in $\mathcal{P}$ that converges to X w.r.t. the metric of definition 5.1. Show that the stochastic integrals $I(X^n)$ converge to $I(X;M)$ w.r.t. the metric of definition 3.2 and also that $\sup_{s \leq t} |I(X^n)_s - I(X)_s| \xrightarrow{\mathbb{P}} 0$. (The latter convergence is called uniform convergence on compacts in probability, abbreviated by ucp-convergence).

6.2 Show, not referring to proposition 6.6, that $I(1_{[0,T]}X; M) = I(X;M)^T$ for any finite stopping time T, $M \in \mathcal{M}_c^2$ and $X \in \mathcal{P}(M)$.

6.3 Show that $I(\zeta 1_{(S,T]}; M)_t = \zeta(M_{T\wedge t} - M_{S\wedge t})$ if S and T are stopping times such that $S \leq T$ and ζ is a bounded $\mathcal{F}_S$-measurable random variable.

6.4 Show that $(\langle M,N\rangle_t - \langle M,N\rangle_s)^2 \leq (\langle M\rangle_t - \langle M\rangle_s)(\langle N\rangle_t - \langle N\rangle_s)$ a.s. for $M,N \in \mathcal{M}^2$. *Hint:* It holds on a set with probability one that for all rational a and b one has $\langle aM + bN\rangle_t - \langle aM + bN\rangle_s \geq 0$. Write this difference termwise and show that we may also take a and b real and use then that this defines a nonnegative quadratic form.

6.5 Prove the second assertion of proposition 6.7.

6.6 Show the equality (6.8) for X and Y in $\mathcal{S}$.

6.7 Show that the quadratic variation of $I(X;M)$ is given by (6.6) if $M \in \mathcal{M}_c^2$ and $X \in \mathcal{P}(M)$.

6.8 Prove lemma 6.5, so show that $I(X+Y;M)$ and $I(X;M)+I(Y;M)$ are indistinguishable if $M \in \mathcal{M}_c^2$ and $X,Y \in \mathcal{P}(M)$.

6.9 Let W be standard Brownian motion. Find a sequence of piecewise constant processes W^n such that $\mathbb{E}\int_0^T |W_t^n - W_t|^2 dt \to 0$. Compute $\int_0^T W_t^n \, dW_t$ and show that it 'converges' (in what sense?) to $\frac{1}{2}(W_T^2 - T)$, if we consider smaller and smaller intervals of constancy. Deduce that $\int_0^T W_t \, dW_t = \frac{1}{2}(W_T^2 - T)$.

6.10 Let $M \in \mathcal{M}_c^{loc}$, $X,Y \in \mathcal{P}^*(M)$ and $a,b \in \mathbb{R}$. Show that $I(aX+bY;M)$ and $aI(X;M)+bI(Y;M)$ are indistinguishable.

6.11 Let W be Brownian motion and T a stopping time with $\mathbb{E}\,T < \infty$. Show that $\mathbb{E}\,W_T = 0$ and $\mathbb{E}\,W_T^2 = \mathbb{E}\,T$.

6.12 Define for $M \in \mathcal{M}_c^2$ and $X \in \mathcal{P}(M)$ for $0 \leq s < t \leq T$ the random variable $\int_s^t X\,dM$ as $I(X;M)_t - I(X;M)_s$. Show that $\int_s^t X\,dM = I(1_{(s,t]}X;M)_T$.

6.13 Let $M,N \in \mathcal{M}_c^2$ and $X \in \mathcal{P}(M) \cap \mathcal{P}(N)$. Show that $I(X;M)+I(X;N)$ and $I(X;M+N)$ are indistinguishable.

6.14 Finish the proof of proposition 6.8 as follows. Show that we can deduce from the given proof that inequality (6.9) holds for all *bounded* processes $X \in \mathcal{P}(M)$ and $Y \in \mathcal{P}(N)$ and then for all $X \in \mathcal{P}(M)$ and $Y \in \mathcal{P}(N)$.

7 The Itô formula

As almost always we assume also in this section that the filtration $\mathbb{F}$ on the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfies the usual conditions.

7.1 Semimartingales

Definition 7.1. A process X is called a (continuous) semimartingale if it admits a decomposition

$$X = X_0 + A + M, \tag{7.1}$$

where $M \in \mathcal{M}_c^{loc}$ and A is a continuous process with paths of bounded variation (over finite intervals) and $A_0 = 0$ a.s.

Remark 7.2. The decomposition (7.1) is unique up to indistinguishability. This follows from the fact that continuous local martingales that are of bounded variation are a.s. constant in time.

Every continuous submartingale is a semimartingale and its semimartingale decomposition of definition 7.1 coincides with the Doob-Meyer decomposition.

Definition 7.3. If X and Y are semimartingales with semimartingale decompositions $X = X_0 + A + M$ and $Y = Y_0 + B + N$ where M and N are local martingales and A and B processes of bounded variation, then we define their quadratic covariation process $\langle X, Y\rangle$ as $\langle M, N\rangle$.

Proposition 7.4 *The given definition of $\langle X, Y\rangle$ for semimartingales coincides with our intuitive understanding of quadratic covariation. If (Π^n) is a sequence of partitions of $[0, t]$ whose meshes $\mu(\Pi^n)$ tend to zero, then for every $T > 0$ we have*

$$\sup_{t\leq T} |V_t(X, Y; \Pi^n) - \langle X, Y\rangle_t| \xrightarrow{\mathbb{P}} 0,$$

where $V_t(X, Y; \Pi) = \sum (X_{t_{k+1}} - X_{t_k})(Y_{t_{k+1}} - Y_{t_k})$ with the summation over the $t_k \in \Pi$.

Proof. It is sufficient to prove this for $X = Y$ with $X_0 = 0$. Let $X = A + M$ according to equation (7.1). Write $V_t(X, X; \Pi^n) = V_t(M, M; \Pi^n) + 2V_t(A, M; \Pi^n) + V_t(A, A; \Pi^n)$. Since A has paths of bounded variation and A and M have continuous paths, the last two terms tend to zero a.s. We concentrate henceforth on $V_t(M, M; \Pi^n)$. Let (T^m) be a localizing sequence for M such that M^{T^m} is bounded by m as well as $\langle M^{T^m}\rangle$. For given $\varepsilon, \delta, T > 0$, we can choose T^m such that $\mathbb{P}(T^m \leq T) < \delta$. We thus have

$$\begin{aligned}\mathbb{P}(\sup_{t\leq T} |V_t(M, M; \Pi^n) - \langle M\rangle_t| > \varepsilon) \leq \\ \delta + \mathbb{P}(\sup_{t\leq T} |V_t(M^{T^m}, M^{T^m}; \Pi^n) - \langle M^{T^m}\rangle_t| > \varepsilon, T < T^m).\end{aligned} \tag{7.2}$$

Realizing that $V(M^{T^m}, M^{T^m}; \Pi^n) - \langle M^{T^m} \rangle$ is a bounded martingale, we apply Doob's inequality to show that the second term in (7.2) tends to zero when the mesh $\mu(\Pi^n)$ tends to zero. Actually, as in the proof of proposition 3.8, we show L^2-convergence. Then let $m \to \infty$. (We thus obtain by almost the same techniques an improvement of proposition 3.8). $\square$

Definition 7.5. A progressive process Y is called locally bounded if there exists a localizing sequence (T^n) such that the stopped processes Y^{T^n} are bounded.

Clearly, all continuous processes are locally bounded and locally bounded processes belong to $\mathcal{P}^*(M)$ for any continuous local martingale M. Moreover, for locally bounded processes Y and continuous processes A that are of bounded variation, the pathwise Stieltjes integrals $\int_0^t Y_s(\omega)\, dA_s(\omega)$ are (a.s.) defined for every $t > 0$.

Definition 7.6. Let Y be a locally bounded process and X a semimartingale with decomposition (7.1). Then the stochastic integral of Y w.r.t. X is defined as $I(Y;M) + \int_0^{\cdot} Y\, dA$. Here the first integral is the stochastic integral of theorem 6.15 and the second one is the Stieltjes integral that we just mentioned. From now on we will use the notation $\int_0^{\cdot} Y\, dX$ or $Y \cdot X$ to denote the stochastic integral of Y w.r.t. X.

Proposition 7.7 *Let Y, U be locally bounded processes and X a semimartingale. Then the product YU is locally bounded as well, $Y \cdot X$ is a semimartingale and $U \cdot (Y \cdot X) = (YU) \cdot X$. If T is a stopping time, then $(Y \cdot X)^T = Y \cdot X^T = (1_{[0,T]} Y) \cdot X$.*

Proof. We only have to take care of the (local) martingale parts. But for these we have propositions 6.6 and 6.12, that we combine with an appropriate stopping argument. $\square$

Proposition 7.8 *Let X be a continuous semimartingale and (Y^n) a sequence of locally bounded progressive processes that converge to zero pointwise. If there exists a locally bounded progressive process Y that dominates all the Y^n, then $Y^n \cdot X$ converges to zero uniformly in probability on compact sets (notation: $Y^n \cdot X \overset{ucp}{\to} 0$), meaning that $\sup_{s \le t} |\int_0^s Y^n\, dX|$ tends to zero in probability for all $t \geq 0$.*

Proof. Let $X = A + M$. Then $\sup_{s \le t} |\int_0^s Y^n\, dA|$ tends to zero pointwise in view of Lebesgue's dominated convergence theorem. Therefore we prove the proposition for $X = M$, a continuous local martingale. Choose stopping times T^m such that each M^{T^m} is a square integrable martingale and Y^{T^m} bounded. Then for all $T > 0$, $\mathbb{E} \int_0^T (Y^n)^2\, d\langle M^{T^m} \rangle \to 0$ for $n \to \infty$ and by the first construction of the stochastic integral we also have for all $T > 0$ that $\mathbb{E}\,(Y^n \cdot M^{T^m})_T^2 \to 0$ and consequently, see the proof of proposition 7.4, we have $Y^n \cdot M^{T^m} \overset{ucp}{\to} 0$ for $n \to \infty$. The dependence on T^m can be removed in the same way as in the proof of proposition 7.4. $\square$

Corollary 7.9 *Let Y be a continuous adapted process and X a semimartingale. Let Π^n be a sequence of partitions $\{t_0^n, \dots, t_{k^n}^n\}$ of $[0,t]$ whose meshes tend to zero. Let $Y^n = \sum_k Y_{t_k^n} 1_{(t_k^n, t_{k+1}^n]}$. Then $Y^n \cdot X \stackrel{ucp}{\to} Y \cdot X$.*

Proof. Since Y is locally bounded, we can find stopping times T^m such that $|Y^{T^m}| \leq m$ and hence $\sup_t |(Y^n)_t^{T^m}| \leq m$. We can therefore apply the preceding proposition to the sequence $((Y^n)^{T^m})_{n\geq 1}$ that converges pointwise to Y^{T^m} and with X replaced with X^{T^m}. We thus obtain $Y^n \cdot (X^{T^m}) \stackrel{ucp}{\to} Y \cdot (X^{T^m})$, which is nothing else but $\sup_{t\leq T} |\int_0^{t\wedge T^m} Y^n \, dX - \int_0^{t\wedge T^m} Y \, dX| \stackrel{\mathbb{P}}{\to} 0$ for all $T > 0$. Finally, since for each t the probability $\mathbb{P}(t \leq T^m) \to 1$, we can remove the stopping time in the last ucp convergence. □

7.2 Integration by parts

The following (first) stochastic calculus rule is the foundation for the Itô formula of the next section.

Proposition 7.10 *Let X and Y be (continuous) semimartingales. Then*

$$X_t Y_t = X_0 Y_0 + \int_0^t X \, dY + \int_0^t Y \, dX + \langle X, Y\rangle_t \quad a.s. \, (t \geq 0). \tag{7.3}$$

A special case occurs when $Y = X$ in which case (7.3) becomes

$$X_t^2 = X_0^2 + 2\int_0^t X \, dX + \langle X\rangle_t \quad a.s. \, (t \geq 0). \tag{7.4}$$

Proof. It is sufficient to prove (7.4), because then (7.3) follows by polarization. Let then Π be a subdivision of $[0,t]$. Then, summing over the elements of the subdivision, we have

$$X_t^2 - X_0^2 = 2\sum X_{t_k}(X_{t_{k+1}} - X_{t_k}) + \sum (X_{t_{k+1}} - X_{t_k})^2 \, a.s. (t \geq 0).$$

To the first term on the right we apply corollary 7.9 and for the second term we use proposition 7.4. This yields the assertion. □

Notice that a consequence of proposition 7.10 is that we can use equation (7.3) to *define* the quadratic covariation between two semimartingales. Indeed, some authors take as their point of view.

7.3 Itô's formula

Theorem 7.11 below contains the celebrated Itô formula (7.5), perhaps the most famous and a certainly not to be underestimated result in stochastic analysis.

Theorem 7.11 *Let X be a continuous semimartingale and f a twice continuously differentiable function on $\mathbb{R}$. Then $f(X)$ is a continuous semimartingale as well and it holds that*

$$f(X_t) = f(X_0) + \int_0^t f'(X) \, dX + \frac{1}{2}\int_0^t f''(X) \, d\langle X\rangle, a.s. \, (t \geq 0). \tag{7.5}$$

Before giving the proof of theorem 7.11, we comment on the integrals in equation (7.5). The first integral we have to understand as a sum as in definition 7.6. With M the local martingale part of X we therefore have to consider the integral $f'(X) \cdot M$, which is well defined since $f'(X)$ is continuous and thus locally bounded. Moreover it is a continuous local martingale. With A the finite variation part of X, the integral $\int_0^{\cdot} f'(X)\, dA$ has to be understood in the (pathwise) Lebesgue-Stieltjes sense and thus it becomes a process of finite variation, as is the case for the integral $\int_0^{\cdot} f''(X)\, d\langle X\rangle$. Hence, $f(X)$ is a continuous semimartingale and its local martingale part is $f'(X) \cdot M$.

Proof. The theorem is obviously true for affine functions and for f given by $f(x) = x^2$ equation (7.5) reduces to (7.4). We show by induction that (7.5) is true for any monomial and hence, by linearity, for every polynomial. The general case follows at the end.
Let $f(x) = x^n = x^{n-1}x$. We apply the integration by parts formula (7.3) with $Y_t = X_t^{n-1}$ and assume that (7.5) is true for $f(x) = x^{n-1}$. We obtain

$$X_t^n = X_0^n + \int_0^t X^{n-1}\, dX + \int_0^t X\, dX^{n-1} + \langle X, X^{n-1}\rangle_t. \tag{7.6}$$

By assumption we have

$$X_t^{n-1} = X_0^{n-1} + \int_0^t (n-1)X^{n-2}\, dX + \frac{1}{2}\int_0^t (n-1)(n-2)X^{n-3}\, d\langle X\rangle. \tag{7.7}$$

We obtain from this equation that $\langle X^{n-1}, X\rangle$ (remember that the quadratic covariation between two semimartingales is determined by their (local) martingale parts) is given by $(n-1)\int_0^{\cdot} X^{n-2}\, d\langle X\rangle$. Inserting this result as well as (7.7) into (7.6) and using proposition 7.7 we get the result for $f(x) = x^n$ and hence for f equal to an arbitrary polynomial.
Suppose now that f is twice continuously differentiable and that X is bounded, with values in $[-K, K]$, say. Then, since f'' is continuous, we can (by the Weierstraß approximation theorem) view it on $[-K, K]$ as the uniform limit of a sequence of polynomials, p_n'' say: for all $\varepsilon > 0$ there is n_0 such that $\sup_{[-K,K]} |f''(x) - p_n''(x)| < \varepsilon$, if $n > n_0$. But then f' is the uniform limit of p_n' defined by $p_n'(x) = f'(-K) + \int_{-K}^x p_n''(u)\, du$ and f as the uniform limit of the polynomials p_n defined by $p_n(x) = f(-K) + \int_{-K}^x p_n'(u)\, du$. For the polynomials p_n we already know that (7.5) holds true. Write R for the difference of the left hand side of (7.5) minus its right hand side. Then

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

The first two differences in this equation can be made arbitrarily small by the definition of the p_n. To the first (stochastic) integral we apply proposition 7.8 and the last integral has absolute value less than $\varepsilon\langle X\rangle_t$ if $n > n_0$.
Let now f be arbitrary and let $T^K = \inf\{t \geq 0 : |X_t| > K\}$. Then, certainly

X^{T^K} is bounded by K and we can apply the result of the previous step. We have

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

which can be rewritten as

$$f(X_{t\wedge T^K}) = f(X_0) + \int_0^{t\wedge T^K} f'(X)\, dX + \frac{1}{2}\int_0^{t\wedge T^K} f''(X)\, d\langle X\rangle.$$

We trivially have $f(X_{t\wedge T^K}) \to f(X_t)$ a.s. and the right hand side of the previous equation is on $\{t < T^K\}$ (whose probability tends to 1) equal to the right hand side of (7.5). The theorem has been proved. □

Remark 7.12. Formula (7.5) is often represented in *differential notation*, a short hand way of writing the formula down without integrals. We write

$$df(X_t) = f'(X_t)\, dX_t + \frac{1}{2} f''(X_t)\, d\langle X\rangle_t,$$

or merely

$$df(X) = f'(X)\, dX + \frac{1}{2} f''(X)\, d\langle X\rangle.$$

With minor changes in the proof one can show that also the following multivariate extension of the Itô formula (7.5) holds true. If $X = (X^1, \ldots, X^d)$ is a d-dimensional vector of semimartingales and $f : \mathbb{R}^d \to \mathbb{R}$ is twice continuously differentiable in all its arguments, then

$$f(X_t) = f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X)\, dX^i + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X)\, d\langle X^i, X^j\rangle. \quad (7.8)$$

Notice that in this expression we only need second order derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}$, if the corresponding components X^i and X^j of X both have a non-vanishing martingale part.

Remark 7.13. The integration by parts formula (7.3) is a special case of (7.8).

7.4 Applications of Itô's formula

Let X be a (continuous) semimartingale and define the process Z by $Z_t = Z_0 \exp(X_t - \frac{1}{2}\langle X\rangle_t)$, where Z_0 is an $\mathcal{F}_0$-measurable random variable. Application of Itô's formula gives

$$Z_t = Z_0 + \int_0^t Z\, dX. \quad (7.9)$$

This is a linear (stochastic) integral equation, that doesn't have the usual exponential solution as in ordinary calculus (unless $\langle X\rangle = 0$). The process Z with

$Z_0 = 1$ is called the Doléans exponential of X and we have in this case the special notation

$$Z = \mathcal{E}(X). \tag{7.10}$$

Notice that if X is a local martingale, Z is local martingale as well. Later on we will give conditions on X that ensure that Z becomes a martingale.

An application of Itô's formula we present in the proof of Lévy's characterization of Brownian motion, proposition 7.14.

Proposition 7.14 *Let M be a continuous local martingale (w.r.t. to the filtration $\mathbb{F}$) with $M_0 = 0$ and $\langle M \rangle_t \equiv t$. Then M is a Brownian motion (w.r.t. $\mathbb{F}$).*

Proof. By splitting into real and imaginary part one can show that Itô's formula also holds for complex valued semimartingales (here, by definition both the real and the imaginary part are semimartingales). Let $u \in \mathbb{R}$ be arbitrary and define the process Y by $Y_t = \exp(iuM_t + \frac{1}{2}u^2 t)$. Applying Itô's formula, we obtain

$$Y_t = 1 + iu \int_0^t Y_s \, dM_s.$$

It follows that Y is a complex valued local martingale. We stop Y at the fixed time point t_0. Then, the stopped process Y^{t_0} is bounded and thus a martingale and since $Y > 0$ we get for all $s < t < t_0$

$$\mathbb{E}\,[\frac{Y_t}{Y_s}|\mathcal{F}_s] = 1.$$

This identity in explicit form is equal to

$$\mathbb{E}\,[\exp(iu(M_t - M_s) + \frac{1}{2}u^2(t-s))|\mathcal{F}_s] = 1,$$

which is valid for all $t > s$, since t_0 is arbitrary. Rewriting this as

$$\mathbb{E}\,[\exp(iu(M_t - M_s)|\mathcal{F}_s] = \exp(-\frac{1}{2}u^2(t-s)),$$

we conclude that $M_t - M_s$ is independent of $\mathcal{F}_s$ and has a normal distribution with zero mean and variance $t - s$. Since this is true for all $t > s$ we conclude that M is a Brownian motion w.r.t. $\mathbb{F}$. □

7.5 Exercises

7.1 The Hermite polynomials h_n are defined as

$$h_n(x) = (-1)^n \exp(\frac{1}{2}x^2) \frac{d^n}{dx^n} \exp(-\frac{1}{2}x^2).$$

Let $H_n(x,y) = y^{n/2} h_n(x/\sqrt{y})$. Show that $\frac{\partial}{\partial x} H_n(x,y) = nH_{n-1}(x,y)$, for which you could first prove that

$$\sum_{n\geq 0} \frac{u^n}{n!} h_n(x) = \exp(ux - \frac{1}{2}u^2).$$

Show also that $\frac{\partial}{\partial y} H_n(x,y) + \frac{1}{2}\frac{\partial^2}{\partial x^2} H_n(x,y) = 0$. Show finally that for a Brownian motion W it holds that $H_n(W_t,t) = n\int_0^t H_{n-1}(W_s,s)\,dW_s$.

7.2 Let W be Brownian motion and $X \in \mathcal{S}$. Let $M = X \cdot W$ and $Z = \mathcal{E}(M)$. Show that M and Z are martingales.

7.3 Let X and Y be (continuous) semimartingales. Show that $\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X + Y + \langle X, Y\rangle)$.

7.4 Let X be a strictly positive continuous semimartingale with $X_0 = 1$ and define the process Y by

$$Y_t = \int_0^t \frac{1}{X}\,dX - \frac{1}{2}\int_0^t \frac{1}{X^2}\,d\langle X\rangle.$$

Let the process Z be given by $Z_t = e^{Y_t}$. Compute dZ_t and show that $Z = X$.

7.5 Let W^1, W^2, W^3 be three independent Brownian motions. Let

$$M_t = \big((W_t^1)^2 + (W_t^2)^2 + (W_t^3)^2\big)^{-1/2}.$$

Show that M is a local martingale with $\sup_t \mathbb{E}\, M_t^p < \infty$ if $p < 3$ and that M is *not* a martingale.

7.6 Let $f : \mathbb{R}^2 \to \mathbb{R}$ be twice continuously differentiable in the first variable and continuously differentiable in the second variable. Let X be a continuous semimartingale and B a continuous process of finite variation over finite intervals. Show that for $t \geq 0$

$$\begin{aligned} f(X_t, B_t) &= f(X_0, B_0) + \int_0^t f_x(X,B)\,dX \\ &\quad + \int_0^t f_y(X,B)\,dB + \frac{1}{2}\int_0^t f_{xx}(X,B)\,d\langle X\rangle, \text{ a.s.} \end{aligned}$$

8 Integral representations

We assume that the underlying filtration satisfies the usual conditions. However, in section 8.2 we will encounter a complication and see how to repair this.

8.1 First representation

The representation result below explains a way how to view any continuous local martingale as a stochastic integral w.r.t. a suitably defined process W that is a Brownian motion.

Proposition 8.1 *Let M be a continuous local martingale with $M_0 = 0$ whose quadratic variation process $\langle M\rangle$ is almost surely absolutely continuous. Then there exists, possibly defined on an extended probability space (again denoted by $(\Omega, \mathcal{F}, \mathbb{P})$), an $\mathbb{F}$-Brownian motion W and process X in $\mathcal{P}^*(W)$ such that $M = X \cdot W$ (up to indistinguishability).*

Proof. Let X^n for $n \in \mathbb{N}$ be defined by $X^n_t = n(\langle M\rangle_t - \langle M\rangle_{t-1/n})$ (with $\langle M\rangle_s = 0$ for $s < 0$). Then all the X^n are progressive processes and thanks to the fundamental theorem of calculus their limit X for $n \to \infty$ exists for each $t > 0$ a.s. by assumption, and is progressive as well. Suppose that $X_t > 0$ a.s. for all $t > 0$. Then we define $f_t = X_t^{-1/2}$ and $W = f \cdot M$ and we notice that $f \in \mathcal{P}^*(M)$. From the calculus rules for computing the quadratic variation of a stochastic integral we see that $\langle W\rangle_t \equiv t$. Hence, by Lévy's characterization (proposition 7.14), W (which is clearly adapted) is a Brownian motion. By the chain rule for stochastic integrals (proposition 6.12) we obtain that $X^{1/2} \cdot W$ is indistinguishable from M.
If X_t assumes for some t the value zero with positive probability, we cannot define the process f as we did above. Let in this case $(\Omega', \mathbb{F}', \mathbb{P}')$ be another probability space that is rich enough to support a Brownian motion B. We consider now the product space of $(\Omega, \mathcal{F}, \mathbb{P})$ with this space and define in the obvious way M and B (as well as the other processes that we need below) on this product space. Notice that everything defined on the original space now becomes independent of B. We define in this case the process W by

$$W = 1_{\{X>0\}} X^{-1/2} \cdot M + 1_{\{X=0\}} \cdot B.$$

The two (local) martingales that sum to W have zero quadratic covariation (exercise 3.6) and hence $\langle W\rangle_t = \int_0^t 1_{\{X>0\}} X^{-1}\, d\langle M\rangle + \int_0^t 1_{\{X=0\}} d\langle B\rangle = t$. Hence W is also a Brownian motion in this case. Finally, again by the chain rule for stochastic integrals, $X^{1/2} \cdot W = 1_{\{X>0\}} \cdot M + 1_{\{X=0\}} X^{1/2} \cdot B = 1_{\{X>0\}} \cdot M$ and $M - X^{1/2} \cdot W = 1_{\{X=0\}} \cdot M$ has quadratic variation identically zero and is thus indistinguishable from the zero martingale. □

8.2 Representation of Brownian local martingales

Suppose that on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we can define a Brownian motion W. Let $\mathbb{F}^W$ be the filtration generated by this process. Application of the preceding theory to situations that involve this filtration is not always justified since this filtration doesn't satisfy the usual conditions (one can show that it is not right-continuous, see exercise 8.1). However, we have

Proposition 8.2 *Let $\mathbb{F}$ be the filtration with σ-algebras $\mathcal{F}_t = \mathcal{F}_t^W \vee \mathcal{N}$, where $\mathcal{N}$ is the collection of sets that are contained in sets of $\mathcal{F}_\infty^W$ with zero probability. Then $\mathbb{F}$ satisfies the usual conditions and W is also a Brownian motion w.r.t. this filtration.*

Proof. The proof of right-continuity involves arguments that use the Markovian character of Brownian motion (the result can be extended to hold for more general Markov processes) and will not be given here. Clearly, by adding null sets to the original filtration the distributional properties of W don't change. □

Remark 8.3. It is remarkable that the addition of the null sets to the σ-algebras of the given filtration renders the filtration right-continuous.

Any process that is a martingale w.r.t. the filtration of proposition 8.2 (it will be referred to as the augmented Brownian filtration) is called a Brownian martingale. Below we sharpen the result of proposition 8.1 in the sense that the Brownian motion is now given and not constructed and that moreover the integrand process X is progressive w.r.t. the augmented Brownian filtration.

Lemma 8.4 *Let $T > 0$ and $\mathcal{R}_T$ be the subset of $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ consisting of the random variables $(X \cdot W)_T$ for $X \in \mathcal{P}_T(W)$ and let $\mathcal{R}$ be the class of stochastic integrals $X \cdot W$, where X ranges through $\mathcal{P}(W)$. Then $\mathcal{R}_T$ is a closed subspace of $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. Moreover, every martingale M in $\mathcal{M}^2$ can be uniquely (up to indistinguishability) written as the sum $M = N + Z$, where $N \in \mathcal{R}$ and $Z \in \mathcal{M}^2$ that is such that $\langle Z, N' \rangle = 0$ for every $N' \in \mathcal{R}$.*

Proof. Let $((X^n \cdot W)_T)$ be a Cauchy sequence in $\mathcal{R}_T$. By the construction of the stochastic integrals (the isometry property in particular), we have that (X^n) is Cauchy in $\mathcal{P}_T(W)$. But this is a complete space (proposition 5.6) and thus has a limit X in this space. By the isometry property again, we have that $(X^n \cdot W)_T$ converges to $(X \cdot W)_T$ in $\mathcal{R}_T$.
The uniqueness of the decomposition in the second assertion is established as follows. Suppose that a given $M \in \mathcal{M}^2$ can be decomposed as $N^1 + Z^1 = N^2 + Z^2$. Then $0 = \langle Z^1 - Z^2, N^2 - N^1 \rangle = \langle Z^1 - Z^2 \rangle$, hence the uniqueness follows.
We now prove the existence of the decomposition on an arbitrary interval (but fixed) $[0, T]$. Invoking the already established uniqueness, one can extend the existence to $[0, \infty)$. Since $M_T \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ and R_T is closed, we have the unique decomposition $M_T = N_T + Z_T$, with $N_T \in \mathcal{R}_T$ and $\mathbb{E}\, Z_T N_T' = 0$ for any $N_T' \in \mathcal{R}_T$. Let Z be the right-continuous modification of the martingale defined by $\mathbb{E}\,[Z_T|\mathcal{F}_t]$ and likewise we denote by N the right-continuous modification of

$\mathbb{E}\,[N_T|\mathcal{F}_t]$. Then $M = N + Z$ on $[0,T]$. We now show that $\langle Z, N'\rangle = 0$, where N' is given by $N'_t = \mathbb{E}\,[N'_T|\mathcal{F}_t]$, with $N'_T \in \mathcal{R}_T$ and $t \leq T$. Equivalently, we show that ZN' is a martingale on $[0,T]$. Write $N'_T = (X \cdot W)_T$ for some $X \in \mathcal{P}_T(W)$. Let $F \in \mathcal{F}_t$. Then the process Y given by $Y_u = 1_F 1_{(t,T]}(u) X_u$ is progressive (check!) and in $\mathcal{P}_T(W)$. Moreover we have $(Y \cdot W)_T = 1_F(N'_T - N'_t)$. It follows that

$$
\begin{aligned}
\mathbb{E}\,[1_F Z_T N'_T] &= \mathbb{E}\,[1_F Z_T (N'_T - N'_t)] + \mathbb{E}\,[1_F Z_t N'_t] \\
&= \mathbb{E}\,[Z_T (Y \cdot W)_T] + \mathbb{E}\,[1_F Z_t N'_t].
\end{aligned}
$$

Since Z_T is orthogonal in the L^2-sense to $\mathcal{R}_T$, the expectation $\mathbb{E}\,[Z_T(Y\cdot W)_T] = 0$ and hence ZN' is a martingale on $[0,T]$, since F is arbitrary. □

Remark 8.5. In the proof of the above lemma we have not exploited the fact that we deal with Brownian motion, nor didn't we use the special structure of the filtration (other than it satisfies the usual conditions). The lemma can therefore be extended to other martingales than Brownian motion. However, the next theorem shows that the process Z of lemma 8.4 in the Brownian context is actually zero.

Theorem 8.6 *Let M be a square integrable Brownian martingale with $M_0 = 0$. Then there exists a process $X \in \mathcal{P}(W)$ such that $M = X \cdot W$. The process X is unique in the sense that for any other process X' with the same property one has $||X - X'|| = 0$.*

Proof. As a starting point we take the decomposition $M = N + Z$ of lemma 8.4. The first (and major) step in the process of proving that $Z = 0$ is to show that for any $n \in \mathbb{N}$ and bounded measurable functions f_k $(k = 1, \ldots, n)$ we have for $0 = s_0 \leq \cdots \leq s_n \leq t$ that

$$
\mathbb{E}\,\big(Z_t \prod_{k=1}^n f_k(W_{s_k})\big) = 0. \tag{8.1}
$$

For $n = 0$ this is trivial since $Z_0 = 0$. We use induction to show (8.1). Suppose that (8.1) holds true for a given n and suppose w.l.o.g. that $s_n < t$ and let $s \in [s_n, t]$. By P_n we abbreviate the product $\prod_{k=1}^n f_k(W_{s_k})$. Put for all $\theta \in \mathbb{R}$

$$
\phi(s,\theta) = \mathbb{E}\,\big(Z_t P_n e^{i\theta W_s}\big).
$$

We keep θ fixed for the time being. Observe that by the martingale property of Z we have $\phi(s,\theta) = \mathbb{E}\, Z_s P_n e^{i\theta W_s}$ and by the induction assumption we have $\phi(s_n, \theta) = 0$. Below we will need that

$$
\mathbb{E}\,[Z_s \int_{s_n}^s e^{i\theta W_u}\, dW_u | \mathcal{F}_{s_n}] = 0, \tag{8.2}
$$

which is nothing else but saying that the product of Z and $\int_0^\cdot e^{i\theta W_u}\, dW_u$ is a martingale and this follows from the fact Z is orthogonal to stochastic integrals

that are in $\mathcal{R}$ (lemma 8.4).
As in the proof of proposition 7.14 we use Itô's formula to get

$$e^{i\theta W_s} = e^{i\theta W_{s_n}} + i\theta \int_{s_n}^{s} e^{i\theta W_u}\, dW_u - \frac{1}{2}\theta^2 \int_{s_n}^{s} e^{i\theta W_u}\, du. \tag{8.3}$$

Multiplication of equation (8.3) by $Z_s P_n$, taking conditional expectations w.r.t. $\mathcal{F}_{s_n}$ and using the just fact (8.2), yields

$$\mathbb{E}\,[Z_s P_n e^{i\theta W_s} | \mathcal{F}_{s_n}] = Z_{s_n} P_n e^{i\theta W_{s_n}} - \frac{1}{2}\theta^2 \mathbb{E}\,[Z_s P_n \int_{s_n}^{s} e^{i\theta W_u}\, du | \mathcal{F}_{s_n}].$$

Take then expectations and use reconditioning combined with the fact that we also have $\phi(s,\theta) = \mathbb{E}\,(Z_s P_n e^{i\theta W_s})$ to obtain

$$\phi(s,\theta) = \phi(s_n,\theta) - \frac{1}{2}\theta^2 \mathbb{E}\,\big(P_n \int_{s_n}^{s} e^{i\theta W_u} \mathbb{E}\,[Z_s|\mathcal{F}_u]\, du\big),$$

which then becomes

$$\phi(s,\theta) = \phi(s_n,\theta) - \frac{1}{2}\theta^2 \int_{s_n}^{s} \phi(u)\, du.$$

Since $\phi(s_n,\theta) = 0$, the unique solution to this integral equation is the zero solution. Hence $\phi(s_{n+1},\theta) = 0$ for all θ. Stated otherwise, the Fourier transform of the signed measure μ on $\mathcal{B}(\mathbb{R})$ given by $\mu(B) = \mathbb{E}\,\big(1_B(W_{s_{n+1}}) Z_t P_n\big)$ is identically zero. But then μ must be the zero measure and thus for any bounded measurable function f_{n+1} we have $\int f_{n+1}\, d\mu = 0$. Hence we have (8.1) also for n replaced with $n+1$ and thus for all n, which completes the proof of the first step. By the monotone class theorem (exercise 8.2) we conclude that $\mathbb{E}\, Z_t \xi = 0$ for all bounded $\mathcal{F}_t^W$-measurable functions ξ. But then also for all bounded ξ that are $\mathcal{F}_t$-measurable, because the two σ-algebras differ only by null sets. We conclude that $Z_t = 0$ a.s. Since this holds true for each t separately, we have by right-continuity of Z that Z is indistinguishable from the zero process. The uniqueness of the assertion is trivial. □

Corollary 8.7 *Let M be a right-continuous local martingale adapted to the augmented Brownian filtration $\mathbb{F}$ with $M_0 = 0$. Then there exists a process $X \in \mathcal{P}^*(W)$ such that $M = X \cdot W$. In particular all such local martingales are continuous.*

Proof. Let (T^n) be a localizing sequence for M such that the stopped processes M^{T^n} are bounded martingales. Then we have according to theorem 8.6 the existence of processes X^n such that $M^{T^n} = X^n \cdot W$. By the uniqueness of the representation we have that $X^n 1_{[0,T^{n-1}]} = X^{n-1} 1_{[0,T^{n-1}]}$. Hence we can unambiguously define $X_t = \lim_n X_t^n$ and it follows that $M_t = \lim_n M_{t\wedge T^n} = (X \cdot W)_t$.

8.3 Exercises

8.1 The filtration $\mathbb{F}^W$ is not right-continuous. Let $\Omega = C[0,\infty)$ and $W_t(\omega) = \omega(t)$ for $t \geq 0$. Fix $t > 0$ and let F be the set of functions $\omega \in \Omega$ that have a local maximum at t. Show that $F \in \mathcal{F}_{t+}$. Suppose that $F \in \mathcal{F}_t$. Since any set G in $\mathcal{F}_t$ is determined by the paths of functions ω up to time t, such a set is unchanged if we have continuous continuations of such functions after time t. In particular, if $\omega \in G$, then also $\omega' \in G$, where $\omega'(s) = \omega(s \wedge t) + (s-t)^+$. Notice that for any ω the function ω' doesn't have a local maximum at t. Conclude that $F \notin \mathcal{F}_t$.

8.2 Complete the proof of theorem 8.6 by writing down the Monotone Class argument.

8.3 Provide explicitly the integrands X according to theorem 8.6 for the following martingales. (If $M_0 \neq 0$, you have to adjust this theorem slightly.)
(a) $M_t = W_t^3 - c\int_0^t W_s\,ds$ for a suitable constant c (which one?).
(b) For some fixed time T we have $M_t = \mathbb{E}[e^{W_T}|\mathcal{F}_t]$.
(c) For some fixed time T we take $M_t = \mathbb{E}[\int_0^T W_s\,ds|\mathcal{F}_t]$.
(d) If v is a solution of the partial differential equation

$$\frac{\partial v}{\partial t}(t,x) + \frac{1}{2}\frac{\partial^2 v}{\partial x^2}(t,x) = 0,$$

then $M_t = v(t, W_t)$ is a martingale (show this by the Itô formula, you may assume enough integrability).

8.4 Let M be as in theorem 8.6. Show that $\langle M, W\rangle$ is a.s. absolutely continuous. Express X in terms of $\langle M, W\rangle$.

8.5 Let ξ be a square integrable random variable that is $\mathcal{F}^W_\infty$-measurable, where $\mathcal{F}^W_\infty$ is the σ-algebra generated by a Brownian motion on $[0,\infty)$. Show that there exists a unique process $X \in \mathcal{P}(W)_\infty$ such that

$$\xi = \mathbb{E}\,\xi + \int_0^\infty X\,dW.$$

8.6 The uniqueness result of exercise 8.5 relies on $X \in \mathcal{P}(W)_\infty$. For any $T > 0$ we define $S_T = \inf\{t > T : W_t = 0\}$. Then the S_T are stopping times. Let $\xi = 0$. Show that $\xi = \int_0^\infty 1_{[0,S_T]}\,dW$ for any T.

8.7 The result of theorem 8.6 is not constructive, it is not told how to construct the process X from the given Brownian martingale M. In the following case we can give an explicit expression for X. Suppose that M is of the form $M_t = f(t, W_t)$, where W is a standard Brownian motion and $f : \mathbb{R}^2 \to \mathbb{R}$ is continuously differentiable in the first variable (with partial derivative f_1) and twice continuously differentiable in the second variable (with first and second order partial derivatives denoted by f_2 and f_{22}). Show that the process X is given by $X_t = f_2(t, W_t)$ and also that f satisfies the *backward heat equation*, i.e. $f_1 + \frac{1}{2} f_{22} = 0$. Give two examples of a martingale M that can be written in this form.

9 Absolutely continuous change of measure

In section 7.3 we have seen that the class of semimartingales is closed under smooth transformations; if X is a semimartingale, so is $f(X)$ if f is twice continuously differentiable. In the present section we will see that the semimartingale property is preserved, when we change the underlying probability measure in an absolutely continuous way. This result is absolutely not obvious. Indeed, consider a semimartingale X with decomposition $X = X_0 + A + M$ with all paths of A of bounded variation. Then this property of A evidently still holds, if we change the probability measure $\mathbb{P}$ into any other one. But it is less clear what happens to M. As a matter of fact, M (suppose it is a martingale under $\mathbb{P}$) will in general loose the martingale property. We will see later on that it becomes a semimartingale and, moreover, we will be able to give its semimartingale decomposition under the new probability measure.

9.1 Absolute continuity

Let $(\Omega, \mathcal{F})$ be a measurable space. We consider two measures on this space, μ and ν. One says that ν is absolutely continuous with respect to μ (the notation is $\nu \ll \mu$) if $\nu(F) = 0$ for every $F \in \mathcal{F}$ for which $\mu(F) = 0$. It we have both $\nu \ll \mu$ and $\mu \ll \nu$ we say that μ and ν are equivalent and we write $\mu \sim \nu$. If there exists a set $\Omega_0 \in \mathcal{F}$ for which $\nu(\Omega_0) = 0$ and $\mu(\Omega_0^c) = 0$, then μ and ν are called mutually singular.

If Z is a nonnegative measurable function on Ω, then $\nu(F) = \int_F Z \, d\mu$ defines a measure on $\mathcal{F}$ that is absolutely continuous w.r.t. μ. The content of the Radon-Nikodym theorem (theorem 9.1) is that this is, loosely speaking, the only case of absolute continuity.

Theorem 9.1 *Let μ be σ-finite measure on $(\Omega, \mathcal{F})$ and let ν be a finite measure on $(\Omega, \mathcal{F})$. Then we can uniquely decompose ν as the sum of two measures ν_s and ν_0, where ν_s and μ are mutually singular and $\nu_0 \ll \mu$. Moreover, there exists a unique nonnegative $Z \in L^1(\Omega, \mathcal{F}, \mu)$ such that $\nu_0(F) = \int_F Z \, d\mu$. This function is called the Radon-Nikodym derivative of ν w.r.t. μ and is often written as*

$$Z = \frac{d\nu}{d\mu}.$$

We will be interested in the case where μ and ν are probability measures, as usual called $\mathbb{P}$ and $\mathbb{Q}$, and will try to describe the function Z in certain cases. Notice that for $\mathbb{Q} \ll \mathbb{P}$ we have $\mathbb{E}_{\mathbb{P}} Z = 1$, $\mathbb{Q}(Z = 0) = 0$ and for $\mathbb{Q} \sim \mathbb{P}$ also $\mathbb{P}(Z = 0) = 0$. Expectations w.r.t. $\mathbb{P}$ and $\mathbb{Q}$ are often denoted by $\mathbb{E}_{\mathbb{P}}$ and $\mathbb{E}_{\mathbb{Q}}$ respectively.

Lemma 9.2 *Let $\mathbb{P}$ and $\mathbb{Q}$ be two probability measures on $(\Omega, \mathcal{F})$ and assume that $\mathbb{Q} \ll \mathbb{P}$ with $Z = \frac{d\mathbb{Q}}{d\mathbb{P}}$. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Let X be a random* *variable. Then* $\mathbb{E}_{\mathbb{Q}}|X| < \infty$ *iff* $\mathbb{E}_{\mathbb{P}}|X|Z < \infty$ *and in either case we have*

$$\mathbb{E}_{\mathbb{Q}}[X|\mathcal{G}] = \frac{\mathbb{E}_{\mathbb{P}}[XZ|\mathcal{G}]}{\mathbb{E}_{\mathbb{P}}[Z|\mathcal{G}]} \text{ a.s. w.r.t. } \mathbb{Q}. \tag{9.1}$$

Proof. Let $G \in \mathcal{G}$. We have, using the defining property of conditional expectation both under $\mathbb{Q}$ and $\mathbb{P}$

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left[\mathbb{E}_{\mathbb{Q}}[X|\mathcal{G}]\mathbb{E}_{\mathbb{P}}[Z|\mathcal{G}]1_G\right] &= \mathbb{E}_{\mathbb{P}}\left[\mathbb{E}_{\mathbb{Q}}[X|\mathcal{G}]1_G Z\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\mathbb{E}_{\mathbb{Q}}[X|\mathcal{G}]1_G\right] \\
&= \mathbb{E}_{\mathbb{Q}}[X1_G] \\
&= \mathbb{E}_{\mathbb{P}}[X1_G Z] \\
&= \mathbb{E}_{\mathbb{P}}\left[\mathbb{E}_{\mathbb{P}}[XZ|\mathcal{G}]1_G\right].
\end{aligned}$$

Since this holds for any $G \in \mathcal{G}$, we conclude that $\mathbb{E}_{\mathbb{Q}}[X|\mathcal{G}]\mathbb{E}_{\mathbb{P}}[Z|\mathcal{G}] = \mathbb{E}_{\mathbb{P}}[XZ|\mathcal{G}]$ $\mathbb{P}$- (and thus $\mathbb{Q}$-)a.s. Because

$$\mathbb{Q}(\mathbb{E}_{\mathbb{P}}[Z|\mathcal{G}] = 0) = \mathbb{E}_{\mathbb{P}} Z 1_{\{\mathbb{E}_{\mathbb{P}}[Z|\mathcal{G}]=0\}} = \mathbb{E}_{\mathbb{P}}\mathbb{E}_{\mathbb{P}}[Z|\mathcal{G}]1_{\{\mathbb{E}_{\mathbb{P}}[Z|\mathcal{G}]=0\}} = 0,$$

the division in (9.1) is $\mathbb{Q}$-a.s. justified.

□

9.2 Change of measure on filtered spaces

We consider a measurable space $(\Omega, \mathcal{F})$ together with a right-continuous filtration $\mathbb{F} = \{\mathcal{F}_t, t \geq 0\}$ and two probability measures $\mathbb{P}$ and $\mathbb{Q}$ defined on it. The restrictions of $\mathbb{P}$ and $\mathbb{Q}$ to the $\mathcal{F}_t$ will be denoted by $\mathbb{P}_t$ and $\mathbb{Q}_t$ respectively. Similarly, for a stopping time T, we will denote by $\mathbb{P}_T$ the restriction of $\mathbb{P}$ to $\mathcal{F}_T$. Below we will always assume that $\mathbb{P}_0 = \mathbb{Q}_0$. If $\mathbb{Q} \ll \mathbb{P}$ on $\mathcal{F}$, then necessarily every restriction $\mathbb{Q}_t$ of $\mathbb{Q}$ to $\mathcal{F}_t$ is absolutely continuous w.r.t. the restriction $\mathbb{P}_t$ of $\mathbb{P}$ to $\mathcal{F}_t$ and thus we have a family of *densities* (Radon-Nikodym derivatives) Z_t, defined by

$$Z_t = \frac{d\mathbb{Q}_t}{d\mathbb{P}_t}. \tag{9.2}$$

The process $Z = \{Z_t, t \geq 0\}$ is called the density process (of $\mathbb{Q}$ w.r.t. $\mathbb{P}$). Here is the first property.

Proposition 9.3 *If $\mathbb{Q} \ll \mathbb{P}$ (on $\mathcal{F}$), then the density process Z is a nonnegative uniformly integrable martingale w.r.t. probability measure $\mathbb{P}$ and $Z_0 = 1$.*

Proof. Exercise 9.1. □

However, we will encounter many interesting situations, where we only have $\mathbb{Q}_t \ll \mathbb{P}_t$ for all $t \geq 0$ and where Z is not uniformly integrable. One may also envisage a reverse situation. One is given a nonnegative martingale Z, is it then possible to find probability measures $\mathbb{Q}_t$ (or $\mathbb{Q}$) such that $\mathbb{Q}_t \ll \mathbb{P}_t$ (or $\mathbb{Q} \ll \mathbb{P}$)? The answers are given in the next two propositions.

Proposition 9.4 *Let Z be a nonnegative uniformly integrable martingale under the probability measure $\mathbb{P}$ with $\mathbb{E}_{\mathbb{P}} Z_t = 1$ (for all t). Then there exists a probability measure $\mathbb{Q}$ on $\mathcal{F}_\infty$ such that $\mathbb{Q}$ is absolutely continuous w.r.t. the restriction of $\mathbb{P}$ to $\mathcal{F}_\infty$. If we denote by $\mathbb{P}_t$, respectively $\mathbb{Q}_t$, the restrictions of $\mathbb{P}$, respectively $\mathbb{Q}$, to $\mathcal{F}_t$, then $\frac{d\mathbb{Q}_t}{d\mathbb{P}_t} = Z_t$ for all t.*

Proof. Notice that there exists a $\mathcal{F}_\infty$-measurable random variable Z_∞ such that $\mathbb{E}_{\mathbb{P}}[Z_\infty|\mathcal{F}_t] = Z_t$, for all $t \geq 0$. Simply define $\mathbb{Q}$ on $\mathcal{F}_\infty$ by $\mathbb{Q}(F) = \mathbb{E}_{\mathbb{P}} 1_F Z_\infty$. $\mathbb{Q}$ is a probability measure since $\mathbb{E}_{\mathbb{P}} Z_\infty = 1$. If F is also in $\mathcal{F}_t$, then we have by the martingale property that $\mathbb{Q}_t(F) = \mathbb{Q}(F) = \mathbb{E}_{\mathbb{P}} 1_F Z_t$. □

If we drop the uniform integrability requirement of Z in proposition 9.4, then the conclusion can not be drawn. There will be a family of probability measures $\mathbb{Q}_t$ on the $\mathcal{F}_t$ that is consistent in the sense that the restriction of $\mathbb{Q}_t$ to $\mathcal{F}_s$ coincides with $\mathbb{Q}_s$ for all $s < t$, but there will in general not exist a probability measure $\mathbb{Q}$ on $\mathcal{F}_\infty$ that is absolutely continuous w.r.t. the restriction of $\mathbb{P}$ to the σ-algebra $\mathcal{F}_\infty$. The best possible general result in that direction is the following.

Proposition 9.5 *Let Ω be the space of real continuous functions on $[0,\infty)$ and let X be the coordinate process ($X_t(\omega) = \omega(t)$) on this space. Let $\mathbb{F} = \mathbb{F}^X$ and let $\mathbb{P}$ be a given probability measure on $\mathcal{F}^X_\infty$. Let Z be a nonnegative martingale with $\mathbb{E}_{\mathbb{P}} Z_t = 1$. Then there exists a unique probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}^X_\infty)$ such that the restrictions $\mathbb{Q}_t$ of $\mathbb{Q}$ to $\mathcal{F}_t$ are absolutely continuous w.r.t. the restrictions $\mathbb{P}_t$ of $\mathbb{P}$ to $\mathcal{F}_t$ and $\frac{d\mathbb{Q}_t}{d\mathbb{P}_t} = Z_t$.*

Proof. Let $\mathcal{A}$ be the algebra $\bigcup_{t\geq 0} \mathcal{F}_t$. Observe that one can define a set function $\mathbb{Q}$ on $\mathcal{A}$ unambiguously by $\mathbb{Q}(F) = \mathbb{E}_{\mathbb{P}} 1_F Z_t$ if $A \in \mathcal{F}_t$. The assertion of the proposition follows from Caratheodory's extension theorem as soon as one has shown that $\mathbb{Q}$ is countably additive on $\mathcal{A}$. We omit the proof. □

Remark 9.6. Notice that in proposition 9.5 it is not claimed that $\mathbb{Q} \ll \mathbb{P}$ on $\mathcal{F}^X_\infty$. In general this will not happen, see exercise 9.3.

Proposition 9.7 *Let $\mathbb{P}$ and $\mathbb{Q}$ be probability measures on $(\Omega, \mathcal{F})$ and $\mathbb{Q}_t \ll \mathbb{P}_t$ for all $t \geq 0$. Let Z be their density process and let X be any adapted cadlag process. Then XZ is a martingale under $\mathbb{P}$ iff X is a martingale under $\mathbb{Q}$. If XZ is a local martingale under $\mathbb{P}$ then X is a local martingale under $\mathbb{Q}$. In the latter case equivalence holds under the extra condition that $\mathbb{P}_t \ll \mathbb{Q}_t$ for all $t \geq 0$.*

Proof. We prove the 'martingale version' only. Using lemma 9.2, we have for $t > s$ that

$$\mathbb{E}_{\mathbb{Q}}[X_t|\mathcal{F}_s] = \frac{\mathbb{E}_{\mathbb{P}}[X_t Z_t|\mathcal{F}_s]}{Z_s},$$

from which the assertion immediately follows. The 'local martingale' case is left as exercise 9.4. □

9.3 The Girsanov theorem

In this section we will explicitly describe how the decomposition of semimartingale changes under an absolutely continuous change of measure. This is the content of what is known as Girsanov's theorem, theorem 9.8. The standing assumption in this and the next section is that *the density process Z is continuous.*

Theorem 9.8 *Let X be a continuous semimartingale on $(\Omega, \mathcal{F}, \mathbb{P})$ w.r.t. a filtration $\mathbb{F}$ with semimartingale decomposition $X = X_0 + A + M$. Let $\mathbb{Q}$ be another probability measure on $(\Omega, \mathcal{F})$ such that $\mathbb{Q}_t \ll \mathbb{P}_t$ for all $t \geq 0$ and with density process Z. Then X is also a continuous semimartingale on $(\Omega, \mathcal{F}, \mathbb{Q})$ and the local martingale part $M^{\mathbb{Q}}$ of X under $\mathbb{Q}$ is given by*

$$M^{\mathbb{Q}} = M - Z^{-1} \cdot \langle Z, M \rangle. \tag{9.3}$$

Moreover, if X and Y are semimartingales (under $\mathbb{P}$), then their quadratic covariation process $\langle X, Y \rangle$ is the same under $\mathbb{P}$ and $\mathbb{Q}$.

Proof. To show that $M^{\mathbb{Q}}$ is a local martingale under $\mathbb{Q}$ we use proposition 9.7 and Itô's formula for products. Let $T^n = \inf\{t > 0 : Z_t < 1/n\}$. On each $[0, T^n]$, the process $Z^{-1} \cdot \langle Z, M \rangle$ is well defined. Since $T^n \to \infty$ $\mathbb{Q}$-a.s. (exercise 9.6) $Z^{-1} \cdot \langle Z, M \rangle$ is well defined everywhere $\mathbb{Q}$-a.s.
The product rule gives

$$\begin{aligned} (M^{\mathbb{Q}} Z)_t^{T^n} &= M_0 Z_0 + \int_0^{T^n \wedge t} M^{\mathbb{Q}}\, dZ + \int_0^{T^n \wedge t} Z\, dM^{\mathbb{Q}} + \langle M^{\mathbb{Q}}, Z \rangle_{T^n \wedge t} \\ &= M_0 Z_0 + \int_0^{T^n \wedge t} M^{\mathbb{Q}}\, dZ + \int_0^{T^n \wedge t} Z\, dM. \end{aligned}$$

This shows that the stopped processes $(M^{\mathbb{Q}} Z)^{T^n}$ are local martingales under $\mathbb{P}$. But then also $M^{\mathbb{Q}} Z$ is a local martingale under $\mathbb{P}$ (exercise 9.9), which is what we had to prove. The statement concerning the quadratic variation process one can prove along the same lines, or by invoking proposition 7.4 and by noticing that addition or subtraction of a finite variation process has no influence on the quadratic variation. □

Girsanov's theorem becomes simpler to prove if for all t the measures $\mathbb{P}_t$ and $\mathbb{Q}_t$ are equivalent, in which case the density process is also strictly positive $\mathbb{P}$-a.s. and can be written as a Doléans exponential.

Proposition 9.9 *Let Z be a strictly positive continuous local martingale with $Z_0 = 1$. Then there exists a unique continuous local martingale M such that $Z = \mathcal{E}(M)$.*

Proof. Since $Z > 0$ we can define $Y = \log Z$, a semimartingale whose semimartingale decomposition $Y = M + A$ satisfies $M_0 = A_0 = 0$. We apply the Itô formula to Y and obtain

$$dY_t = \frac{1}{Z_t}\, dZ_t - \frac{1}{2Z_t^2}\, d\langle Z \rangle_t.$$

Hence

$$M_t = \int_0^t \frac{1}{Z_s}\, dZ_s, \tag{9.4}$$

and we observe that $A = -\frac{1}{2}\langle M\rangle$. Hence $Z = \exp(M - \frac{1}{2}\langle M\rangle) = \mathcal{E}(M)$. Showing the uniqueness is the content of exercise 9.10. □

Proposition 9.10 *Let μ be a continuous local martingale and let $Z = \mathcal{E}(\mu)$. Let $T > 0$ and assume that $\mathbb{E}_{\mathbb{P}}\, Z_T = 1$. Then Z is a martingale under $\mathbb{P}$ on $[0,T]$. If M is a continuous local martingale under $\mathbb{P}$, then $M^{\mathbb{Q}} = M - \langle M, \mu\rangle$ is a continuous local martingale under the measure $\mathbb{Q}_T$ defined on $(\Omega, \mathcal{F}_T)$ by $\frac{d\mathbb{Q}_T}{d\mathbb{P}_T} = Z_T$ with time restricted to $[0,T]$.*

Proof. That Z is a martingale follows from proposition 4.7 (ii) and exercise 2.1. Then we apply theorem 9.8 and use that $\langle Z, M\rangle = \int_0^\cdot Z\, d\langle \mu, M\rangle$ to find the representation for $M^{\mathbb{Q}}$. □

Remark 9.11. In the situation of proposition 9.10 we have actually $\mathbb{P}_T \sim \mathbb{Q}_T$ and the density $\frac{d\mathbb{P}_T}{d\mathbb{Q}_T} = Z_T^{-1}$, alternatively given by $\mathcal{E}(-\mu^{\mathbb{Q}})_T$, with $\mu^{\mathbb{Q}} = \mu - \langle\mu\rangle$, in agreement with the notation of proposition 9.10. Moreover, if $M^{\mathbb{Q}}$ is a local martingale under $\mathbb{Q}_T$ over $[0,T]$, then $M^{\mathbb{Q}} + \langle M^{\mathbb{Q}}, \mu\rangle$ is a local martingale under $\mathbb{P}$ over $[0,T]$.

Corollary 9.12 *Let W be a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. Assume that the conditions of proposition 9.10 are in force with $\mu = X \cdot W$, where $X \in \mathcal{P}_T^*(W)$. Then the process $W^{\mathbb{Q}} = W - \int_0^\cdot X_s\, ds$ is a Brownian motion on $[0,T]$ under $\mathbb{Q}_T$ w.r.t. the filtration $\{\mathcal{F}_t, t \in [0,T]\}$.*

Proof. We know from proposition 9.10 that $W^{\mathbb{Q}}$ is a continuous local martingale on $[0,T]$. Since the quadratic variation of this process is the same under $\mathbb{Q}$ as under $\mathbb{P}$ (theorem 9.8), the process $W^{\mathbb{Q}}$ must be a Brownian motion in view of Lévy's characterization (proposition 7.14). □

This corollary only gives us a Brownian motion $W^{\mathbb{Q}}$ under $\mathbb{Q}_T$ on $[0,T]$. Suppose that this would be the case for every T, can we then say that $W^{\mathbb{Q}}$ is a Brownian motion on $[0,\infty)$? For an affirmative answer we would have to extend the family of probability measures $\mathbb{Q}_T$ defined on the $\mathcal{F}_T$ to a probability measure on $\mathcal{F}_\infty$, and as we have mentioned before this is in general impossible. But if we content ourselves with a smaller filtration (much in the spirit of proposition 9.5), something is possible.

Proposition 9.13 *Let W be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathbb{F}^W$ be the filtration generated by W and let X be a process that is progressively measurable w.r.t. $\mathbb{F}^W$ such that $\int_0^T X_s^2\, ds < \infty$ a.s. for all $T > 0$. Then there exists a probability measure $\mathbb{Q}$ on $\mathcal{F}_\infty^W$ such that $W^{\mathbb{Q}} = W - \int_0^\cdot X_s\, ds$ is a Brownian motion on $(\Omega, \mathcal{F}_\infty^W, \mathbb{Q})$.*

Proof. For all T, the probability measures $\mathbb{Q}_T$ induce probability measures $\mathbb{Q}'_T$ on $(\mathbb{R}^{[0,T]}, \mathcal{B}(\mathbb{R}^{[0,T]}))$, the laws of the Brownian motion W up to the times T. Notice that for $T' > T$ the restriction of $\mathbb{Q}'_{T'}$ to $\mathcal{B}(\mathbb{R}^{[0,T]})$ coincides with $\mathbb{Q}'_T$. It follows that the finite dimensional distributions of $W^{\mathbb{Q}}$ form a consistent family. In view of Kolmogorov's theorem there exists a probability measure $\mathbb{Q}'$ on $(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R}^{[0,\infty)}))$ such that the restriction of $\mathbb{Q}'$ to $\mathcal{B}(\mathbb{R}^{[0,T]})$ coincides with $\mathbb{Q}'_T$. A typical set F in $\mathcal{F}^W_\infty$ has the form $F = \{\omega : W_\cdot(\omega) \in F'\}$, for some $F' \in \mathcal{B}(\mathbb{R})$. So we define $\mathbb{Q}(F) = \mathbb{Q}'(F')$ and this results in a probability measure on $\mathcal{F}^W_\infty$. If $F \in \mathcal{F}^W_T$, then we have $\mathbb{Q}(F) = \mathbb{Q}'_T(F') = \mathbb{Q}_T(F)$, so this $\mathbb{Q}$ is the one we are after. Observe that W^Q is adapted to $\mathbb{F}^W$ and let $0 \leq t_1 < \cdots < t_n$ be a given arbitrary n-tuple. Then for $B \in \mathcal{B}(\mathbb{R}^n)$ we have $\mathbb{Q}((W^{\mathbb{Q}}{}_{t_1}, \ldots, W^{\mathbb{Q}}{}_{t_n}) \in B) = \mathbb{Q}_{t_n}((W^{\mathbb{Q}}{}_{t_1}, \ldots, W^{\mathbb{Q}}{}_{t_n}) \in B)$. Since $W^{\mathbb{Q}}$ is Brownian motion on $[0, t_n]$ (corollary 9.12) the result follows. □

Remark 9.14. Consider the probability measure $\mathbb{Q}$ of proposition 9.13. Suppose that all σ-algebras of the filtration $\mathbb{F}$ contains all the $\mathbb{P}$-null sets of $\mathcal{F}^W_\infty$ (which happens if $\mathbb{F}$ is the Brownian filtration). Let $\mathbb{Q}_T$ be the probability measure of corollary 9.12. It is tempting to think that the restriction of $\mathbb{Q}$ to the σ-algebra $\mathcal{F}_T$ coincides with $\mathbb{Q}_T$. This is in general not true. One can find sets F in $\mathcal{F}_T$ for which $\mathbb{P}(F) = 0$ and hence $\mathbb{Q}_T(F) = 0$, although $\mathbb{Q}(F) > 0$ (exercise 9.8).

9.4 The Kazamaki and Novikov conditions

The results in the previous section were based on the fact that the density process Z had the martingale property. In the present section we will see two sufficient conditions in terms of properties of Z that guarantee this. The condition in the first proposition below is called Kazamaki's condition, the one in the second proposition is called Novikov's condition.

Proposition 9.15 *Let μ be a local martingale with $\mu_0 = 0$ and suppose that $\exp(\frac{1}{2}\mu)$ is a uniformly integrable submartingale on an interval $[0, T]$ ($T \leq \infty$), then $\mathcal{E}(\mu)$ is a uniformly integrable martingale on $[0, T]$.*

Proof. We give the proof for $T = \infty$. For other values of T only the notation in the proof changes. Let $a \in (0, 1)$, put $Z(a) = \exp(a\mu/(1 + a))$ and consider $\mathcal{E}(a\mu)$. Note that $\mathcal{E}(a\mu) = \mathcal{E}(\mu)^{a^2} Z(a)^{1-a^2}$. For any set $F \in \mathcal{F}$ (think below of $F = \{\mathcal{E}(a\mu)_t > n\}$ for arbitrary $n \in \mathbb{N}$) we have by Hölder's inequality

$$\mathbb{E}\, 1_F \mathcal{E}(a\mu)_t \leq (\mathbb{E}\, \mathcal{E}(\mu)_t)^{a^2} (\mathbb{E}\, 1_F Z(a)_t)^{1-a^2}.$$

Since $\mathcal{E}(\mu)$ is a nonnegative local martingale, it is a nonnegative supermartingale (proposition 4.7 (ii)). Hence it has an a.s. limit for $t \to \infty$ and by Fatou's lemma $\mathbb{E}\, \mathcal{E}(\mu)_t \leq 1$ and $\mathbb{E}\, \mathcal{E}(\mu)_\infty \leq 1$. Together with the easy to prove fact that $Z(a)$ is uniformly integrable, we obtain that $\mathcal{E}(a\mu)$ is a uniformly integrable martingale (exercise 9.11) and the limits $Z(a)_\infty$ and $\mathcal{E}(a\mu)_\infty$ exist a.s.

Using Hölder's inequality again together with the fact that we now know that $\mathbb{E}\,\mathcal{E}(a\mu)_\infty = 1$ we obtain

$$1 = \mathbb{E}\,\mathcal{E}(a\mu)_\infty \leq \big(\mathbb{E}\,\mathcal{E}(\mu)_\infty\big)^{a^2} \big(\mathbb{E}\,Z(a)_\infty\big)^{1-a^2}. \tag{9.5}$$

Notice that by uniform integrability of $\exp(\frac{1}{2}\mu)$ the limit $\exp(\frac{1}{2}\mu_\infty)$ has finite expectation. The trivial bound on $Z(a)_\infty$

$$Z(a)_\infty \leq 1_{\{\mu_\infty<0\}} + \exp(\frac{1}{2}\mu_\infty)1_{\{\mu_\infty\geq 0\}}.$$

yields

$$\sup_{a<1}(\mathbb{E}\,Z(a)_\infty) < 1 + \mathbb{E}\,\exp(\frac{1}{2}\mu_\infty) < \infty,$$

and thus $\limsup_{a\to 1}(\mathbb{E}\,Z(a)_\infty)^{1-a^2} = 1$. But then we obtain from (9.5), that $\mathbb{E}\,\mathcal{E}(\mu)_\infty \geq 1$. Using the already know inequality $\mathbb{E}\,\mathcal{E}(\mu)_\infty \geq 1$, we conclude that $\mathbb{E}\,\mathcal{E}(\mu)_\infty = 1$, from which the assertion follows. □

In the proof of proposition 9.17 we use the following lemma.

Lemma 9.16 *Let M be a uniformly integrable martingale such that $\mathbb{E}\,\exp(M_\infty) < \infty$. Then $\exp(M)$ is a uniformly integrable submartingale.*

Proof. Exercise 9.12 □

Proposition 9.17 *Let μ be a continuous local martingale such that*

$$\mathbb{E}\,\exp(\frac{1}{2}\langle\mu\rangle_T) < \infty \text{ for } T \in [0,\infty]. \tag{9.6}$$

Then $\mathcal{E}(\mu)$ is a uniformly integrable martingale on $[0,T]$.

Proof. Again we give the proof for $T = \infty$. First we show that μ is uniformly integrable. It follows from the hypothesis and Jensen's inequality that $\mathbb{E}\,\langle\mu\rangle_\infty < \infty$. Let (T^n) be a fundamental sequence for μ such that the stopped processes μ^{T^n} are square integrable. Then $\mathbb{E}\,\mu^2_{T^n\wedge t} \leq \mathbb{E}\,\langle\mu\rangle_\infty < \infty$ for all t and T^n. It follows that the collection of the random variables $\mu_{T^n\wedge}$ is uniformly integrable and therefore also the μ_t. Consequently μ_∞ exists as an a.s. and L^1 limit. We apply the Cauchy-Schwarz inequality to

$$\exp(\frac{1}{2}\mu_\infty) = \big(\mathcal{E}(\mu)_\infty\big)^{1/2}\exp(\frac{1}{4}\langle\mu\rangle_\infty)$$

to get

$$\mathbb{E}\,\exp(\frac{1}{2}\mu_\infty) \leq \big(\mathbb{E}\,\mathcal{E}(\mu)_\infty\big)^{1/2}\big(\mathbb{E}\,\exp(\frac{1}{2}\langle\mu\rangle_\infty)\big)^{1/2} \leq \big(\mathbb{E}\,\exp(\frac{1}{2}\langle\mu\rangle_\infty)\big)^{1/2}.$$

Hence $\mathbb{E}\,\exp(\frac{1}{2}\mu_\infty) < \infty$. It follows from lemma 9.16 that $\mathcal{E}(\frac{1}{2}\mu)$ is a uniformly integrable submartingale and by proposition 9.15 we conclude that $\mathcal{E}(\mu)$ is uniformly integrable on $[0,\infty]$. □

9.5 Exercises

9.1 Prove proposition 9.3.

9.2 Let X be a semimartingale on $(\Omega, \mathcal{F}, \mathbb{P})$. Let Y be a locally bounded process. Show that $Y \cdot X$ is invariant under an absolutely continuous change of measure.

9.3 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a measurable space on which is defined a Brownian motion W. Let $\{\mathbb{P}_\theta : \theta \in \mathbb{R}\}$ be a family of probability measures on $\mathcal{F}^W_\infty$ such that for all θ the process W^θ defined by $W^\theta_t = W_t - \theta t$ is a Brownian motion under $\mathbb{P}_\theta$. Show that this family of measures can be defined, and how to construct it. Show that all the P^θ are mutually singular on $\mathcal{F}^W_\infty$. Suppose that we consider θ as an unknown parameter (value) and that we observe X that under each $\mathbb{P}^\theta$ has the semimartingale decomposition $X_t = \theta t + W^\theta_t$. Take $\theta_t = \frac{X_t}{t}$ as estimator of θ when observations up to time t have been made. Show that θ_t is consistent: $\theta_t \to \theta$, P^θ-a.s. for all θ.

9.4 Finish the proof of proposition 9.7.

9.5 Let $\mathbb{P}$ and $\mathbb{Q}$ be measure on the space $(\Omega, \mathcal{F})$ with a filtration $\mathbb{F}$ and assume that $\mathbb{Q}_t \ll \mathbb{P}_t$ for all $t \geq 0$ with density process Z. Let T be a stopping time. Let $\Omega' = \Omega \cap \{T < \infty\}$ and $\mathcal{F}'_T = \{F \cap \{T < \infty\} : F \in \mathcal{F}_T\}$. Show that with $\mathbb{P}'_T$ the restriction of $\mathbb{P}_T$ to $\mathcal{F}'_T$ (and likewise we have $\mathbb{Q}'_T$) that $\mathbb{Q}'_T \ll \mathbb{P}'_T$ and $\frac{d\mathbb{Q}'_T}{d\mathbb{P}'_T} = Z_T$.

9.6 Let $\mathbb{P}$ and $\mathbb{Q}$ be probability measures on the space $(\Omega, \mathcal{F})$ and assume that $\mathbb{Q}_t \ll \mathbb{P}_t$ for all $t \geq 0$. Let Z be the density process. Let $T^n = \inf\{t : Z_t < \frac{1}{n}\}$. Show that $\mathbb{Q}(T^n < \infty) \leq \frac{1}{n}$ and deduce that $\mathbb{Q}(\inf\{Z_t : t > 0\} = 0) = 0$ and that $\mathbb{Q}(\lim_n T^n = \infty) = 1$.

9.7 Prove the claims made in remark 9.11.

9.8 Consider the situation of remark 9.14. Consider the density process $Z = \mathcal{E}(W)$. Let $F = \{\lim_{t\to\infty} \frac{W_t}{t} = 1\}$. Use this set to show that the probability measures $\mathbb{Q}_T$ and $\mathbb{Q}$ restricted to $\mathcal{F}_T$ are different.

9.9 Let T^n be stopping times such that $T^n \to \infty$ and let M be a process such that M^{T^n} is a local martingale for all n. Show that M is a local martingale.

9.10 Show that uniqueness holds for the local martingale M of proposition 9.9.

9.11 Show that the process $Z(a)$ in the proof of proposition 9.15 is uniformly integrable and that consequently the same holds for $\mathcal{E}(a\mu)$.

9.12 Prove lemma 9.16.

10 Stochastic differential equations

By a *stochastic differential equation* (sde) we mean an equation like

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dW_t \tag{10.1}$$

or

$$dX_t = b(X_t)\, dt + \sigma(X_t)\, dW_t, \tag{10.2}$$

which is sometimes abbreviated by

$$dX = b(X)\, dt + \sigma(X)\, dW.$$

The meaning of (10.1) is nothing else but shorthand notation for the following *stochastic integral equation*

$$X_t = X_0 + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s,\ (t \geq 0). \tag{10.3}$$

Here W is a Brownian motion and b and σ are Borel-measurable functions on $\mathbb{R}^2$ with certain additional requirements to be specified later on. The first integral in (10.3) is a pathwise Lebesgue-Stieltjes integral and the second one a stochastic integral. Of course, both integrals should be well-defined. Examples of stochastic differential equations have already been met in previous sections. For instance, if $X = \mathcal{E}(W)$, then

$$X_t = 1 + \int_0^t X_s\, dW_s,$$

which is of the above type.

We give an infinitesimal interpretation of the functions b and σ. Suppose that a process X can be represented as in (10.3) and that the stochastic integral is a square integrable martingale. Then we have that $\mathbb{E}\,[X_{t+h} - X_t|\mathcal{F}_t] = \mathbb{E}\,[\int_t^{t+h} b(s, X_s)\, ds|\mathcal{F}_t]$. For small h, this should 'approximately' be equal to $b(t, X_t)h$. The conditional variance of $X_{t+h} - X_t$ given $\mathcal{F}_t$ is

$$\mathbb{E}\,[(\int_t^{t+h} \sigma(s, X_s)\, dW_s)^2|\mathcal{F}_t] = \mathbb{E}\,[\int_t^{t+h} \sigma^2(s, X_s)\, ds|\mathcal{F}_t],$$

which is approximated for small h by $\sigma^2(t, X_t)h$. Hence the coefficient b in equation (10.1) tells us something about the direction in which X_t changes and σ something about the variance of the displacement. We call b the *drift* coefficient and σ the *diffusion* coefficient.

A process X should be called a solution with initial condition ξ, if it satisfies equation (10.3) and if $X_0 = \xi$. But this phrase is insufficient for a proper definition of the concept of a solution. In this course we will treat two different concepts, one being *strong solution*, the other one *weak solution*, with the emphasis on the former.

10.1 Strong solutions

In order that the Lebesgue-Stieltjes integral and the stochastic integral in (10.3) exist we have to impose (technical) regularity conditions on the process X that is involved. Suppose that the process X is defined on some probability space with a filtration. These conditions are (of course) that X is progressive (which is the case if X is adapted and continuous) and that

$$\int_0^t (|b(s, X_s)| + \sigma(s, X_s)^2)\, ds < \infty, a.s. \forall t \geq 0. \tag{10.4}$$

In all what follows below, we assume that a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is such that it supports a Brownian motion W and a random variable ξ that is independent of W. The filtration $\mathbb{F} = \{\mathcal{F}_t, t \geq 0\}$ that we will mainly work with in the present section is the filtration generated by W and ξ and then augmented with the null sets. More precisely. We set $\mathcal{F}_t^0 = \mathcal{F}_t^W \vee \sigma(\xi)$, $\mathcal{N}$ the $(\mathbb{P}, \mathcal{F}_\infty^0)$-null sets and $\mathcal{F}_t = \mathcal{F}_t^0 \vee \mathcal{N}$. By previous results we know that this filtration is right-continuous and that W is also Brownian w.r.t. it.

Definition 10.1. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a Brownian motion W defined on it as well as a random variable ξ independent of W, a process X defined on this space is called a *strong solution* of equation (10.1) with *initial condition* ξ if $X_0 = \xi$ as. and X
(i) has continuous paths a.s.
(ii) is $\mathbb{F}$ adapted.
(iii) satisfies condition (10.4)
(iv) satisfies (10.1) a.s.

The main result of this section concerns existence and uniqueness of a strong solution of a stochastic differential equation.

Theorem 10.2 *Assume that the coefficients b and σ are Lipschitz continuous in the second variable, i.e. there exists a constant $K > 0$ such that*

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq K|x - y|,\ \forall t \in [0, \infty), \forall x, y \in \mathbb{R},$$

and that $b(\cdot, 0)$ and $\sigma(\cdot, 0)$ are locally bounded functions. Then, for any initial condition ξ with $\mathbb{E}\,\xi^2 < \infty$ the equation (10.1) has a unique strong solution.

Proof. For given processes X and Y that are such that (10.4) is satisfied for both of them, we define

$$U_t(X) = \xi + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s$$

and $U_t(Y)$ likewise. Note that equation (10.3) with $X_0 = \xi$ can now be written as $X = U(X)$ and a solution of this equation can be considered as a fixed point of U. By U^k we denote the k-fold composition of U. We employ the following notation in the proof. For any process Z we write $||Z||_T = \sup\{|Z_t| : t \leq T\}$. We first prove that for every $T > 0$ there exists a constant C such that

$$\mathbb{E}\,||U^k(X) - U^k(Y)||_T^2 \leq \frac{C^k T^k}{k!} \mathbb{E}\,||X - Y||_T^2. \tag{10.5}$$

Fix $T > 0$. Since for any real numbers p, q it holds that $(p+q)^2 \leq 2(p^2 + q^2)$, we obtain for $t \leq T$

$$|U_t(X) - U_t(Y)|^2 \leq \\ 2\Big(\int_0^t |b(s, X_s) - b(s, Y_s)|\, ds\Big)^2 + 2\Big(\int_0^t (\sigma(s, X_s) - \sigma(s, Y_s))\, dW_s\Big)^2,$$

and thus

$$||U(X) - U(Y)||_t^2 \leq \\ 2\Big(\int_0^t |b(s, X_s) - b(s, Y_s)|\, ds\Big)^2 + 2 \sup_{u \leq t} \Big(\int_0^u (\sigma(s, X_s) - \sigma(s, Y_s))\, dW_s\Big)^2$$

Take expectations and use the Cauchy-Schwarz inequality as well as Doob's L^2-inequality (exercise 10.1) to obtain

$$\mathbb{E}\,||U(X) - U(Y)||_t^2 \leq \\ 2\mathbb{E}\,\Big(t \int_0^t |b(s, X_s) - b(s, Y_s)|^2\, ds\Big) + 8\mathbb{E}\,\Big(\int_0^t (\sigma(s, X_s) - \sigma(s, Y_s))^2\, ds\Big).$$

Now use the Lipschitz condition to get

$$\begin{aligned}
\mathbb{E}\,||U(X) - U(Y)||_t^2 &\leq 2\mathbb{E}\,\Big(tK^2 \int_0^t (X_s - Y_s)^2\, ds\Big) \\
&\quad + 8\mathbb{E}\,\Big(K^2 \int_0^t (X_s - Y_s)^2\, ds\Big) \\
&= 2K^2(t+4) \int_0^t \mathbb{E}\,(X_s - Y_s)^2\, ds \\
&\leq 2K^2(T+4) \int_0^t \mathbb{E}\,||X - Y||_s^2\, ds. \qquad (10.6)
\end{aligned}$$

Let $C = 2K^2(T+4)$. Then we have $\mathbb{E}\,||U(X) - U(Y)||_T^2 \leq CT\mathbb{E}\,||X - Y||_T^2$, which establishes (10.5) for $k = 1$. Suppose that we have proved (10.5) for some k. Then we use induction to get from (10.6)

$$\begin{aligned}
\mathbb{E}\,||U^{k+1}(X) - U^{k+1}(Y)||_T^2 &\leq C \int_0^T \mathbb{E}\,||U^k(X) - U^k(Y)||_s^2\, ds \\
&\leq C \int_0^T \frac{C^k s^k}{k!}\, ds\, \mathbb{E}\,||X - Y||_T^2 \\
&= \frac{C^{k+1} T^{k+1}}{(k+1)!} \mathbb{E}\,||X - Y||_T^2.
\end{aligned}$$

This proves (10.5). We continue the proof by using Picard iteration, i.e. we are going to define recursively a sequence of processes X^n that will have a limit in a suitable sense on any time interval $[0,T]$ and that limit will be the candidate solution. We set $X^0 = \xi$ and define $X^n = U(X^{n-1}) = U^n(X^0)$ for $n \geq 1$. Notice that by construction all the X^n have continuous paths and are $\mathbb{F}$-adapted. It follows from equation (10.5) that we have

$$\mathbb{E}\,||X^{n+1} - X^n||_T^2 \leq \frac{C^n T^n}{n!}\mathbb{E}\,||X^1 - X^0||_T^2 \tag{10.7}$$

From the conditions on b and σ we can conclude that $B := \mathbb{E}\,||X^1 - X^0||_T^2 < \infty$ (exercise 10.8). By Chebychev's inequality we have that

$$\mathbb{P}(||X^{n+1} - X^n||_T > 2^{-n}) \leq B\frac{(4CT)^n}{n!}.$$

It follows from the Borel-Cantelli lemma that the set $\Omega' = \liminf_{n\to\infty}\{||X^{n+1} - X^n||_T \leq 2^{-n}\}$ has probability one. On this set we can find for all ω an integer n big enough such that for all $m \in \mathbb{N}$ one has $||X^{n+m} - X^n||_T^2(\omega) \leq 2^{-n}$. In other words, on Ω' the sample paths of the processes X^n form a Cauchy sequence in $C[0,T]$ w.r.t. the sup-norm, and thus all of them have continuous limits. We call the limit process X (outside Ω' we define it as zero) and show that this process is the solution of (10.1) on $[0,T]$. Since $X_t^{n+1} = (U(X^n)_t - U(X)_t) + U(X)_t$ for each $t \geq 0$ and certainly $X_t^{n+1} \to X_t$ in probability, it is sufficient to show that for each t we have convergence in probability (or stronger) of $U(X^n)_t - U(X)_t$ to zero. We look at

$$\begin{aligned} U(X^n)_t - U(X)_t &= \int_0^t (b(s,X_s^n) - b(s,X_s))\,ds & (10.8)\\ &\quad + \int_0^t (\sigma(s,X_s) - \sigma(s,X_s^n))\,dW_s. & (10.9) \end{aligned}$$

We first consider the integral of (10.8). Notice that on Ω' we have that $||X - X^n||_T^2(\omega) \to 0$. One thus has (again ω-wise on Ω')

$$\begin{aligned} |\int_0^t (b(s,X_s^n) - b(s,X_s))\,ds| &\leq K\int_0^T |X_s^n - X_s|\,ds \\ &\leq KT||X^n - X||_T \to 0. \end{aligned}$$

Hence we have a.s. convergence to zero of the integral in (10.8). Next we look at the stochastic integral of (10.9). One has

$$\begin{aligned} \mathbb{E}\,\Big(\int_0^t (\sigma(s,X_s) - \sigma(s,X_s^n))\,dW_s\Big)^2 &= \mathbb{E}\int_0^t (\sigma(s,X_s) - \sigma(s,X_s^n))^2\,ds \\ &\leq K^2\int_0^t \mathbb{E}\,(X_s - X_s^n)^2\,ds \\ &\leq K^2 T\mathbb{E}\,||X - X^n||_T^2 \end{aligned}$$

and we therefore show that $\mathbb{E}\,||X - X^n||_T^2 \to 0$. From (10.7) we obtain

$$\mathbb{E}\,||X^{n+m} - X^n||_T^2 \leq \Big(\sum_{k=n}^{\infty}\big(\frac{(CT)^k}{k!}\big)^{1/2}\Big)^2 \mathbb{E}\,||X^1 - X^0||_T^2. \tag{10.10}$$

Combined with the already established a.s. convergence of X^n to X in the sup-norm, we get by application of Fatou's lemma from (10.10) that $\mathbb{E}\,||X - X^n||_T^2$ is also bounded from above by the right hand side of this equation and thus converges to zero (it is the tail of a convergent series) for $n \to \infty$.

What is left is the proof of unicity on $[0, T]$. Suppose that we have two solutions X and Y. Then, using the same arguments as those that led us to (10.6), we obtain

$$\begin{aligned}\mathbb{E}\,(X_t - Y_t)^2 &= \mathbb{E}\,(U_t(X) - U_t(Y))^2 \\ &\leq C\int_0^t \mathbb{E}\,(X_s - Y_s)^2\,ds.\end{aligned}$$

It now follows from Gronwall's inequality, exercise 10.7, that $E(X_t - Y_t)^2 = 0$, for all t and hence, by continuity, X and Y are indistinguishable on $[0, T]$. Extension of the solution X to $[0, \infty)$ is then established by the unicity of solutions on any interval $[0, T]$. □

Theorem 10.2 is a classical result obtained by Itô. Many refinements are possible by relaxing the conditions on b and σ. One that is of particular interest concerns the diffusion coefficient. Lipschitz continuity can be weakened to some variation on Hölder continuity.

Proposition 10.3 *Consider equation (10.1) and assume that b satisfies the Lipschitz condition of theorem 10.2, whereas for σ we assume that $|\sigma(t,x) - \sigma(t,y)| \leq h(|x-y|)$, where $h : [0,\infty) \to [0,\infty)$ is a strictly increasing function with $h(0) = 0$ and the property that*

$$\int_0^1 \frac{1}{h(u)^2}\,du = \infty.$$

Then, given an initial condition ξ, equation (10.1) admits at most one strong solution.

Proof. Take a_n $(n \geq 1)$ such that

$$\int_{a_n}^1 \frac{1}{h(u)^2}\,du = \frac{1}{2}n(n+1).$$

Then $\int_{a_n}^{a_{n-1}} \frac{1}{nh(u)^2}\,du = 1$. We can take a smooth disturbance of the integrand in this integral, a nonnegative continuous function ρ_n with support in (a_n, a_{n-1}) and bounded by $\frac{2}{nh(\cdot)^2}$ such that $\int_{a_n}^{a_{n-1}} \rho_n(u)\,du = 1$. We consider even functions ψ_n defined for all $x > 0$ by

$$\psi_n(x) = \psi_n(-x) = \int_0^x \int_0^y \rho_n(u)\,du\,dy.$$

Then the sequence (ψ_n) is increasing, $|\psi'(x)| \leq 1$ and $\lim_n \psi_n(x) = |x|$, by the definition of the ρ_n. Notice also that the ψ_n are in C^2, since the ρ_n are continuous.
Suppose that there are two strong solutions X and Y with $X_0 = Y_0$. Assume for a while the properties $\mathbb{E}\,|X_t| < \infty$, $\mathbb{E}\int_0^t \sigma(s,X_s)^2\,ds < \infty$ for all $t > 0$ and the analogous properties for Y. Consider the difference process $V = X - Y$. Apply Itô's rule to $\psi_n(V)$ and take expectations. Then the martingale term vanishes and we are left with (use the above mentioned properties of the ψ_n)

$$\begin{aligned}
\mathbb{E}\,\psi_n(V_t) &= \mathbb{E}\int_0^t \psi_n'(V_s)(b(s,X_s) - b(s,Y_s))\,ds \\
&\quad + \frac{1}{2}\mathbb{E}\int_0^t \psi_n''(V_s)(\sigma(s,X_s) - \sigma(s,Y_s))^2\,ds \\
&\leq K\int_0^t \mathbb{E}\,|V_s|\,ds + \frac{1}{2}\mathbb{E}\int_0^t \psi_n''(V_s)h^2(V_s)\,ds \\
&\leq K\int_0^t \mathbb{E}\,|V_s|\,ds + \frac{t}{n}.
\end{aligned}$$

Letting $n \to \infty$, we obtain $\mathbb{E}\,|V_t| \leq K\int_0^t \mathbb{E}\,|V_s|\,ds$ and the result follows from the Gronwall inequality and sample path continuity under the temporary integrability assumptions on X and Y. As usual these assumptions can be removed by a localization procedure. In this case one works with the stopping times

$$T^n = \inf\{t > 0 : |X_t| + |Y_t| + \int_0^t (\sigma(s,X_s)^2 + \sigma(s,Y_s)^2)\,ds > n\}.$$

□

A strong solution X by definition satisfies the property that for each t the random variable X_t is $\mathcal{F}_t$-measurable. Sets in $\mathcal{F}_t$ are typically determined by certain realizations of ξ and the paths of W up to time t. This suggests that there is a causal relationship between X and (ξ and) W, that should formally take the form that $X_t = f(t,\xi,W_{[0,t]})$, where $W_{[0,t]}$ stands for the map that takes the ω's to the paths $\{W_s(\omega), s \leq t\}$. One would like the map f to have the appropriate measurability properties. This can be accomplished when one is working with the *canonical set-up*. By the canonical set-up we mean the following. We take $\Omega = \mathbb{R} \times C[0,\infty)$. Let $\omega = (u,f) \in \Omega$. We define $w_t(\omega) = f(t)$ and $\xi(\omega) = u$. A filtration on Ω is obtained by setting $\mathcal{F}_t^0 = \sigma(\xi, w_s, s \leq t)$. Let μ be a probability measure on $\mathbb{R}$, $\mathbb{P}^W$ be the Wiener measure on $C[0,\infty)$ (the unique probability measure that makes the coordinate process a Brownian motion) and $\mathbb{P} = \mathbb{P}^W \times \mu$. Finally, we get the filtration $\mathbb{F}$ by augmenting the $\mathcal{F}_t^0$ with the $\mathbb{P}$-null sets of $\mathcal{F}_\infty^0$. Next to the filtration $\mathbb{F}$ we also consider the filtration $\mathbb{H}$ on $C[0,\infty)$ that consists of the σ-algebras $\mathcal{H}_t = \sigma(h_s, s \leq t)$, where $h_t(f) = f(t)$. We state the next two theorems without proof, but see also exercise 10.11 for a simpler version of these theorems.

Theorem 10.4 *Let the canonical set-up be given. Assume that (10.1) has a strong solution with initial condition ξ (so that in particular condition (10.4) is* *satisfied). Let μ be the law of ξ. Then there exists a functional $F_\mu : \Omega \to C[0,\infty)$ such that for all $t \geq 0$*

$$F_\mu^{-1}(\mathcal{H}_t) \subset \mathcal{F}_t \tag{10.11}$$

and such that $F_\mu(\xi, W)$ and X are $\mathbb{P}$-indistinguishable. Moreover, if we work on another probability space on which all the relevant processes are defined, a strong solution of (10.1) with an initial condition ξ is again given by $F_\mu(\xi, W)$ with the same functional F_μ as above.

If we put $F(x, f) := F_{\delta_x}(x, f)$ (where δ_x is the Dirac measure at x), we would like to have $X = F(\xi, W)$. There is however in general a measurability problem with the map $(x, f) \mapsto F(x, f)$. By putting restrictions on the coefficients this problem disappears.

Theorem 10.5 *Suppose that b and σ are as in theorem 10.2. Then a strong solution X may be represented as $X = F(\xi, W)$, where F satisfies the measurability property of (10.11) and moreover, for each $f \in C[0,\infty)$, the map $x \mapsto F(x, f)$ is continuous. Moreover, on any probability space that supports a Brownian motion W and a random variable ξ, a strong solution X is obtained as $X = F(\xi, W)$, with the same mapping F.*

10.2 Weak solutions

Contrary to strong solutions, that have the interpretation of X as an output process of a machine with inputs W and ξ, weak solutions are basically processes defined on a suitable space that can be represented by a stochastic differential equation. This is formalized in the next definition.

Definition 10.6. A *weak solution* of equation (10.1) by definition consists of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a filtration $\mathbb{F}$ and a pair of continuous adapted processes (X, W) such that W is Brownian motion relative to this filtration and moreover
(i) condition (10.4) is satisfied
(ii) equation (10.1) is satisfied a.s.

The law of X_0, given a certain weak solution, is called the initial distribution. Notice that it follows from this definition that X_0 and W are independent. Related to weak solutions there are different concepts of uniqueness. The most relevant one is *uniqueness in law*, which is defined as follows. Uniqueness in law is said to hold if any two weak solutions $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{F}$, (X, W) and $(\Omega', \mathcal{F}', \mathbb{P}')$, $\mathbb{F}'$, (X', W') with X_0 and X'_0 identically distributed are such that also the processes X and X' have the same distribution.

The proposition below shows that one can guarantee the existence of weak solutions under weaker conditions than those under which existence of strong solutions can be proved, as should be the case. The main relaxation is that we drop the Lipschitz condition on b.

Proposition 10.7 *Consider the equation*

$$dX_t = b(t, X_t)\, dt + dW_t,\ t \in [0, T].$$

Assume that b satisfies the growth condition $|b(t, x)| \leq K(1 + |x|)$, for all $t \geq 0$, $x \in \mathbb{R}$. Then, for any initial law μ, this equation has a weak solution.

Proof. We start out with a measurable space $(\Omega, \mathcal{F})$ and a family of probability measures $\mathbb{P}^x$ on it. Let X be a continuous process on this space such that under each $\mathbb{P}^x$ the process $X - x$ is a standard Brownian motion w.r.t. some filtration $\mathbb{F}$. Let $Z_T = \exp(\int_0^T b(s, X_s)\, dX_s - \frac{1}{2}\int_0^T b(s, X_s)^2\, ds)$. At the end of the proof we show that $\mathbb{E}_{\mathbb{P}^x} Z_T = 1$ for all x. Assuming that this is the case, we can define a probability measure $\mathbb{Q}^x$ on $\mathcal{F}_T$ by $\mathbb{Q}^x(F) = \mathbb{E}_{\mathbb{P}^x} Z_T 1_F$. Put

$$W_t = X_t - X_0 - \int_0^t b(s, X_s)\, ds.$$

It follows from Girsanov's theorem that under each of the measures $\mathbb{Q}^x$ the process W is a standard Brownian motion w.r.t. $\{\mathcal{F}_t\}_{t \leq T}$. Define the probability measure $\mathbb{Q}$ by $\mathbb{Q}(F) = \int \mathbb{Q}^x(F)\, \mu(dx)$. Then W is a Brownian motion under $\mathbb{Q}$ as well (you check why!) and $\mathbb{Q}(X_0 \in B) = \mu(B)$. It follows that the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ together with the filtration $\{\mathcal{F}_t\}_{t \leq T}$ and the processes X and W constitute a weak solution.
We now show that $\mathbb{E}_{\mathbb{P}^x} Z_T = 1$ and write simply $\mathbb{E}\, Z_T$ in the rest of the proof. Our goal is to conclude by application of a variant of Novikov's condition. The ideas is to consider the process Z over a number of small time intervals, to check that Novikov's condition is satisfied on all of them and then to collect the conclusions in the appropriate way.
Let $\delta > 0$ (to be specified later on) and consider the processes b^n given by $b^n_t = b(t, X_t) 1_{[(n-1)\delta, n\delta)}(t)$ as well as the associated processes $Z^n = \mathcal{E}(b^n \cdot X)$. Suppose these are all martingales. Then $\mathbb{E}\, [Z^n_{n\delta} | \mathcal{F}_{(n-1)\delta}] = Z^n_{(n-1)\delta}$, which is equal to one. Since we have

$$Z_{N\delta} = \prod_{n=1}^{N} Z^n_{n\delta},$$

we obtain

$$\mathbb{E}\, Z_{N\delta} = \mathbb{E}\, [\prod_{n=1}^{N-1} Z^n_{n\delta} \mathbb{E}\, [Z^N_{N\delta} | \mathcal{F}_{(N-1)\delta}]] = \mathbb{E} \prod_{n=1}^{N-1} Z^n_{n\delta} = \mathbb{E}\, Z_{(N-1)\delta},$$

which can be seen to be equal to one by an induction argument. To show that the Z^n are martingales we use Novikov's condition. This condition is

$$\mathbb{E} \exp\big(\frac{1}{2} \int_0^T (b^n_t)^2\, dt\big) = \mathbb{E} \exp\big(\frac{1}{2} \int_{(n-1)\delta}^{n\delta} b(t, X_t)^2\, dt\big) < \infty.$$

By the assumption on b this follows as soon as we know that

$$\mathbb{E} \exp\big(\frac{1}{2} \delta K^2 (1 + ||X||_T)^2\big) < \infty,$$

which would follow from

$$\mathbb{E}\exp\left(\delta K^2||X||_T^2\right) < \infty.$$

Consider the positive submartingale V given by $V_t = \exp(\frac{1}{2}\delta K^2 X_t^2)$ and note that $||V||_T^2 = \exp\left(\delta K^2||X||_T^2\right)$. By Doob's L^2-inequality we know that $\mathbb{E}\,||V||_T^2 \leq 4\mathbb{E}\,V_T^2 = 4\mathbb{E}\,\exp(\delta K^2 X_T^2)$. Since X_T has a $N(x,T)$ distribution, the last expectation is finite for $\delta < 1/2K^2T$. Choosing such a delta and an integer N such that $N\delta = T$, we have finished the proof of $\mathbb{E}\,Z_T = 1$. $\square$

We now give an example of a stochastic differential equation that has a weak solution that is unique in law, but that doesn't admit a strong solution. This equation is

$$X_t = \int_0^t \operatorname{sgn}(X_s)\,dW_s \tag{10.12}$$

First we show that a weak solution exists. Take a $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a Brownian motion X and define $W_t = \int_0^t \operatorname{sgn}(X)\,dX$. It follows from Lévy's characterization that also W is a Brownian motion. Moreover, one easily sees that (10.12) is satisfied. However, invoking Lévy's characterization again, any weak solution to this equation must be a Brownian motion. Hence we have uniqueness in law. We proceed by showing that assuming existence of a strong solution leads to an absurdity. The proper argument to be used for this involves *local time*, a process that we don't treat in this course. We hope to convince the reader with a heuristic argumentation. First we write $\operatorname{sgn}(x) = \frac{x}{|x|}1_{x\neq 0}$. The definition of the process W can then be recast as

$$W_t = \frac{1}{2}\int_0^t 1_{X_t\neq 0}\frac{1}{|X_t|}\,d(X_t^2 - t).$$

All the processes involved in the right hand side of this equation are $\mathbb{F}^{|X|}$-adapted and so must be W. But a strong solution satisfies $\mathcal{F}_t^X \subset \mathcal{F}_t$, which would lead to $\mathcal{F}_t^X \subset \mathcal{F}_t^{|X|}$, an inclusion which is absurd.

10.3 Markov solutions

First we define a real transition kernel. It is a function $k : \mathbb{R}\times\mathcal{B}(\mathbb{R}) \to [0,1]$ satisfying the properties

$$x \mapsto k(x,A) \text{ is Borel-measurable for all } A \in \mathcal{B}(\mathbb{R})$$

and

$$A \mapsto k(x,A) \text{ is a probability measure for all } x \in \mathbb{R}\,.$$

Definition 10.8. A process X is Markov w.r.t. a filtration $\mathbb{F}$ if there exists a family of transition kernels $\{P_{t,s} : t \geq s \geq 0\}$ such that for all $t \geq s$ and for all bounded continuous functions f it holds that

$$\mathbb{E}\left[f(X_t)|\mathcal{F}_s\right] = \int f(y)P_{t,s}(X_s, dy). \tag{10.13}$$

A Markov process X is called homogeneous, if there is a family of transition kernels $\{P_u : u \geq 0\}$ such that for all $t \geq s$ one has $P_{t,s} = P_{t-s}$. Such a process is called strong Markov if for every a.s. finite stopping time T one has

$$\mathbb{E}\left[f(X_{T+t})|\mathcal{F}_T\right] = \int f(y)P_t(X_T, dy). \tag{10.14}$$

We are interested in Markov solutions to stochastic differential equations. Since the Markov property (especially if the involved filtration is the one generated by the process under consideration itself) is mainly a property of the distribution of a process, it is natural to consider weak solutions to stochastic differential equations. However, showing (under appropriate conditions) that a solution to a stochastic differential equation enjoys the Markov property is much easier for strong solutions, on which we put the emphasis, and therefore we confine ourselves to this case. The canonical approach to show that weak solutions enjoy the Markov property is via what is known as the *Martingale problem*. For showing that strong solutions have the Markov property, we don't need this concept. The main result of this section is theorem 10.10 below.

First some additional notation and definitions. If X is a (strong) solution to (10.1), then we denote by X^s (not to be confused with similar notation for stopped processes) the process $X_{s+\cdot}$, by b^s the function $b(s+\cdot,\cdot)$ and by σ^s the function $\sigma(s+\cdot,\cdot)$. Likewise we have the process W^s defined by $W^s = W_{s+t} - W_s$ and the σ-algebras $\mathcal{F}^s_t = \mathcal{F}_{t+s}$, constituting the filtration $\mathbb{F}^s$. Notice that W^s is a Brownian motion w.r.t. $\mathbb{F}^s$. It follows that the process X^s satisfies the equation

$$X^s_t = X^s_0 + \int_0^t b^s(u, X^s_u)\, du + \int_0^t \sigma^s(u, X^s_u)\, dW^s_u. \tag{10.15}$$

By $X^{s,x}$ we denote the unique strong solution (assuming that it exists) to (10.15) with $X^s_0 = x$. In what follows we will work with a bounded continuous function $f : \mathbb{R} \to \mathbb{R}$ and for such an f we define the function v^f by $v^f(t, s, x) = \mathbb{E}\, f(X^{s,x}_t)$.

Lemma 10.9 *Assume that the conditions of theorem 10.2 are satisfied. By $X^{s,\xi}$ we denote the strong solution to (10.15) with initial value $X_s = \xi$, where the random variable ξ is independent of W^s and satisfies $\mathbb{E}\,\xi^2 < \infty$. Then v^f is continuous in x and $\mathbb{E}\left[f(X^{s,\xi}_t)|\mathcal{F}_s\right] = v^f(t, s, \xi)$.*

Proof. Let x be a fixed initial condition and consider $X^{s,x}$. Since $X^{s,x}$ is adapted to the augmented filtration generated by $W_u - W_s$ for $u > s$ it is independent of $\mathcal{F}_s$. Hence $\mathbb{E}\left[f(X^{s,x}_t)|\mathcal{F}_s\right] = \mathbb{E}\, f(X^{s,x}_t) = v^f(t, s, x)$. If ξ assumes countably many values ξ_j, we compute

$$\begin{aligned}
\mathbb{E}\left[f(X^{s,\xi}_t)|\mathcal{F}_s\right] &= \sum_j 1_{\{\xi=\xi_j\}} \mathbb{E}\left[f(X^{s,\xi_j}_t)|\mathcal{F}_s\right] \\
&= \sum_j 1_{\{\xi=\xi_j\}} v^f(t, s, \xi_j) \\
&= v^f(t, s, \xi).
\end{aligned}$$

The general case follows by approximation. For arbitrary ξ we define for every n the countably valued random variable $\xi^n = 2^{-n}[2^n\xi]$. Then $\xi^n \leq \xi \leq \xi^n + 2^{-n}$. We compute, using Jensen's inequality for conditional expectations,

$$\begin{aligned}
&\mathbb{E}\left(\mathbb{E}\,[f(X_t^{s,\xi^n})|\mathcal{F}_s] - \mathbb{E}\,[f(X_t^{s,\xi})|\mathcal{F}_s]\right)^2 \\
&= \mathbb{E}\left(\mathbb{E}\,[f(X_t^{s,\xi^n}) - f(X_t^{s,\xi})|\mathcal{F}_s]\right)^2 \\
&\leq \mathbb{E}\,\mathbb{E}\,[(f(X_t^{s,\xi^n}) - f(X_t^{s,\xi}))^2|\mathcal{F}_s] \\
&= \mathbb{E}\,(f(X_t^{s,\xi^n}) - f(X_t^{s,\xi}))^2. \qquad (10.16)
\end{aligned}$$

Now we apply exercise 10.12, that tells us that $\mathbb{E}\,(X_t^{s,\xi^n} - X_t^{s,\xi})^2 \to 0$ for $n \to \infty$. But then we also have the convergence in probability and since f is bounded and continuous also the expression in (10.16) tends to zero. Hence we have L^2-convergence of $\mathbb{E}\,[f(X_t^{s,\xi^n})|\mathcal{F}_s]$ to $\mathbb{E}\,[f(X_t^{s,\xi})|\mathcal{F}_s]$. Applying this result to a deterministic $\xi = x$, we obtain continuity of the function $v^f(t,s,\cdot)$. From the L^2-convergence we obtain a.s. convergence along a suitably chosen subsequence. Recall that we already proved that

$$v^f(t,s,\xi^n) = \mathbb{E}\,[f(X_t^{s,\xi^n})|\mathcal{F}_s].$$

Apply now the continuity result to the left hand side of this equation and the a.s. convergence (along some subsequence) to the right hand side to arrive at the desired conclusion. □

Theorem 10.10 *Let the coefficients b and σ satisfy the conditions of theorem 10.2. Then the solution process is Markov. Under the additional assumption that the coefficients b and σ are functions of the space variable only, the solution process is strong Markov.*

Proof. The previous lemma 10.9 has as a corollary that $\mathbb{E}\,[f(X_{t+s})|\mathcal{F}_s] = v^f(t,s,X_s)$, for all bounded and continuous f. Applying this to functions f of the form $f(x) = \exp(i\lambda x)$, we see that the conditional characteristic function of X_{t+s} given $\mathcal{F}_s$ is a measurable function of X_s, since v^f was continuous in x. It then follows that for $t > s$ the conditional probabilities $\mathbb{P}(X_t \in A|\mathcal{F}_s)$ are of the form $P_{t,s}(X_s, A)$, where the functions $\mathbb{P}_{t,s}(\cdot, A)$ are Borel-measurable and that then $\int f(y)P_{t,s}(X_s, dy) = v^f(t-s,s,X_t) = \mathbb{E}\,[f(X_t)|\mathcal{F}_s]$. Hence X is Markov. To prove the strong Markov property for time homogeneous coefficients we follow a similar procedure. First we observe that the functions b^s and σ^s coincide with b and σ. Hence, if T is an a.s. finite stopping time, we have instead of equation (10.15)

$$X_{T+t} = X_T + \int_0^t b(X_{T+u})\,du + \int_0^t \sigma(X_{T+u})\,dW_u^T,$$

with $W_u^T := W_{T+u} - W_T$. By the strong Markov property of Brownian motion, the process W^T is Brownian w.r.t. the filtration $\{F_{T+t}, t \geq 0\}$. It also follows that the function v^f introduced above doesn't depend on the variable s and we write $v^f(t,x)$ instead of $v^f(t,s,x)$. Hence we can copy the above analysis to arrive at $\mathbb{E}[f(X_{T+t})|\mathcal{F}_T] = v^f(t,X_T)$, which is equivalent to the strong Markov property. $\square$

10.4 Exercises

10.1 Prove that for a right-continuous martingale M it holds that

$$\mathbb{E}\,(\sup_{s\le t}|M_s|)^2 \le 4\mathbb{E}\,M_t^2.$$

Hint: Work on $\lambda^2\mathbb{P}(\sup_{s\le t}|M_s|>\lambda) \le \mathbb{E}\,M_t^2 1_{\{\sup_{s\le t}|M_s|>\lambda\}}$.

10.2 Let $X_0, \varepsilon_1, \varepsilon_2, \dots$ be a sequence of independent random variables. Suppose that we generate a (discrete time) random process by the recursion

$$X_t = f(X_{t-1}, \varepsilon_t, t),\ (t \ge 1),$$

where f is a measurable function. Show that the process X is Markov: $\mathbb{P}(X_{t+1} \in B|\mathcal{F}_t^X) = \mathbb{P}(X_{t+1} \in B|X_t)$. Show also the stronger statement: for any bounded and measurable function h we have

$$\mathbb{E}[h(X_{t+1})|\mathcal{F}_t] = \int h(f(X_t, u, t+1))\,F_{t+1}(du),$$

where F_{t+1} is the distribution function of ε_{t+1}.

10.3 Consider the stochastic differential equation

$$dX_t = 3X_t^{1/3}\,dt + 3X_t^{2/3}\,dW_t,$$

with initial condition $X_0 = 0$. Give two different strong solutions to this equation.

10.4 Consider the one-dimensional equation

$$dX_t = -aX_t\,dt + \sigma\,dW_t,\ X_0,$$

where a and σ are real constants. Show that a strong solution is given by $X_t = e^{-at}X_0 + \sigma e^{-at}\int_0^t e^{as}\,dW_s$. Let X_0 have mean μ_0 and variance $\sigma_0^2 < \infty$. Compute $\mu_t = \mathbb{E}\,X_t$ and $\sigma_t^2 = \mathrm{Var}\,X_t$. If X_0 moreover has a normal distribution, then all X_t have a normal distribution as well. X_0 is said to have the invariant distribution if all X_t have the same distribution as X_0. Find this distribution (for existence you also need a condition on a).

10.5 Let X be defined by $X_t = x_0 \exp\big((b-\frac{1}{2}\sigma^2)t+\sigma W_t\big)$, where W is a Brownian motion and b and σ are constants. Write down a stochastic differential equation that X satisfies. Having found this sde, you make in the sde the coefficients b and σ depending on time. How does the solution of this equation look now?

10.6 Consider the equation

$$dX_t = (\theta - aX_t)\,dt + \sigma\sqrt{X_t \vee 0}\,dW_t,\; X_0 = x_0 \geq 0.$$

Assume that $\theta \geq 0$. Show that X is nonnegative.

10.7 Let g be a nonnegative Borel-measurable function, that is locally integrable on $[0,\infty)$. Assume that g satisfies for all $t \geq 0$ the inequality $g(t) \leq a + b\int_0^t g(s)\,ds$, where $a, b \geq 0$. Show that $g(t) \leq ae^{bt}$. *Hint:* Solve the inhomogeneous integral equation

$$g(t) = a + b\int_0^t g(s)ds - p(t)$$

for a nonnegative function p.

10.8 Show that (cf. the proof of theorem 10.2) $\mathbb{E}\,||X^1 - X^0||_T^2 < \infty$.

10.9 Let $T > 0$. Show that under the assumptions of theorem 10.2 it holds that for all $T > 0$ there is a constant C (depending on T) such that $\mathbb{E}\,X_t^2 \leq C(1 + \mathbb{E}\,\xi^2)\exp(Ct)$, for all $t \leq T$.

10.10 Endow $C[0,\infty)$ with the metric d defined by

$$d(x,y) = \sum_{n=1}^{\infty} 2^{-n}(1 \wedge \max_{t\in[0,n]} |x(t) - y(t)|).$$

Show that the Borel σ-algebra $\mathcal{B}(C[0,\infty))$ coincides with the smallest σ-algebra that makes all finite dimensional projections measurable.

10.11 Let X be strong solution to (10.1) with initial value ξ. Show that for each $t > 0$ there is map $f : \mathbb{R}\times C[0,t] \to \mathbb{R}$ that is $\mathcal{B}(\mathbb{R})\times\mathcal{B}(C[0,t])/\mathcal{B}(\mathbb{R})$-measurable such that $X_t(\omega) = f(\xi(\omega), W_{[0,t]}(\omega))$ for almost all ω.

10.12 Let X and Y be strong solutions with possibly different initial values and assume that the conditions of theorem 10.2 are in force. Show that for all T there is a constant D such that

$$\mathbb{E}\,||X - Y||_t^2 \leq D\mathbb{E}\,|X_0 - Y_0|^2, \forall t \leq T.$$

Hint: Look at the first part of the proof of theorem 10.2. First you use that $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$, mimic the proof and finally you use Gronwall's inequality.

10.13 Show that under the assumptions of theorem 10.2 also equation (10.15) admits a unique strong solution. What can you say about this when we replace s with a finite stopping time T?

10.14 If X is a strong solution of (10.1) and the assumptions of theorem 10.2 are in force. Then the bivariate process $\{(X_t, t), t \geq 0\}$ is strong Markov. Show this.

10.15 If X is a (weak or strong) solution to (10.1), then for all $f \in C^2(\mathbb{R})$, the process $M_t^f = f(X_t) - f(X_0) - \int_0^t \mathcal{L}_s f(X_s)\, ds$ is a local martingale. Here the operators $\mathcal{L}_s$ are defined by $\mathcal{L}_s f(x) = b(s,x)f'(x) + \frac{1}{2}\sigma^2(s,x)f''(x)$ for $f \in C^2(\mathbb{R})$. If we restrict f to be a C_K^∞-function, then the M^f become martingales. Show these statements.

10.16 Consider the equation

$$X_t = 1 + \int_0^t X_s\, dWs.$$

Apply the Picard-Lindelöf iteration scheme of the proof of theorem 10.2 (so $X_t^0 \equiv 1$, etc.). Show that

$$X_t^n = \sum_{k=0}^n \frac{1}{k!} H_k(W_t, t),$$

where the functions H_k are those of exercise 7.1. Conclude that $X_t^n \to \mathcal{E}(W)_t$ a.s. for $n \to \infty$, for every $t \geq 0$. Do we also have a.s. convergence, uniform over compact time intervals?

11 Partial differential equations

The simplest example of a partial differential equation whose solutions can be expressed in terms of a diffusion process is the *heat equation*

$$u_t - \frac{1}{2}u_{xx} = 0. \tag{11.1}$$

The fundamental solution of this equation for $t > 0$ and $x \in \mathbb{R}$ is the density of the $N(0,t)$ distribution,

$$p(t,x) = \frac{1}{\sqrt{2\pi t}}\exp(-\frac{x^2}{2t}).$$

This solution is clearly not unique, $u(t,x) = t + x^2$ is another one. For uniqueness one needs e.g. an initial value, $u(0,x) = f(x)$ say, with some prespecified function f. Leaving technicalities aside for the moment, one may check by scrupulously interchanging differentiation and integration that

$$u(t,x) = \int_{\mathbb{R}} f(y)p(t,x-y)\,dy$$

satisfies equation (11.1) for $t > 0$. Moreover, we can write this function u as $u(t,x) = \mathbb{E}\,f(x + W_t)$, where we extend the domain of the t-variable to $t \geq 0$. Notice that for $t = 0$ we get $u(0,x) = f(x)$. Under appropriate conditions on f one can show that this gives the unique solution to (11.1) with the initial condition $u(0,x) = f(x)$.

If we fix a terminal time T, we can define $v(t,x) = u(T-t,x)$ for $t \in [0,T]$, with u a solution of the heat equation. Then v satisfies the *backward heat equation* equation

$$v_t + \frac{1}{2}v_{xx} = 0,$$

with terminal condition $v(T,x) = f(x)$ if $u(0,x) = f(x)$. It follows that we have the representation $v(t,x) = \mathbb{E}\,f(x + W_{T-t}) = \mathbb{E}\,f(x + W_T - W_t)$. Denote by $W^{t,x}$ a process defined on $[t,\infty)$ that starts at t in x and whose increments have the same distribution as those of Brownian motion. Then we can identify this process as $W^{t,x}_s = x + W_s - W_t$ for $s \geq t$. Hence we have $v(t,x) = \mathbb{E}\,f(W^{t,x}_T)$.

In this section we will look at partial differential equations that are more general than the heat equation, with initial conditions replaced by a terminal condition. The main result is that solutions to such equations can be represented as functionals of solutions to stochastic differential equations.

11.1 Feynman-Kac formula

Our starting point is the stochastic differential equation (10.1). Throughout this section we assume that the coefficients b and σ are continuous and that the linear growth condition

$$|b(t,x)| + |\sigma(t,x)| \leq K(1 + |x|) \tag{11.2}$$

is satisfied. Moreover we assume that (10.1) allows for each pair (t,x) a weak solution involving a process $(X^{t,x})$ that is unique in law and satisfies $X^{t,x}_t = x$ a.s. We also need the family of operators $\{\mathcal{L}_t, t > 0\}$ (called *generator* in Markov process language) acting on functions $f \in C^{1,2}([0,\infty)\times\mathbb{R})$ defined by

$$\mathcal{L}_t f(t,x) = b(t,x)f_x(t,x) + \frac{1}{2}\sigma^2(t,x)f_{xx}(t,x). \tag{11.3}$$

Below, in the proof of lemma 11.2, we use inequality (11.4), known as one of the Burkholder-Davis-Gundy inequalities.

Lemma 11.1 *Let $p \geq 2$. There exists a constant C_p such that for all continuous local martingales M with $M_0 = 0$ and all finite stopping times T one has*

$$\mathbb{E}\sup_{t\leq T}|M_t|^p \leq C_p \mathbb{E}\,\langle M\rangle_T^{p/2}. \tag{11.4}$$

Proof. Assume first that M is bounded. The function $x \mapsto |x|^p$ is in $C^2(\mathbb{R})$ so we can apply Itô's formula to get

$$|M_T|^p = p\int_0^T \operatorname{sgn}(M_t)|M_t|^{p-1}\,dM_t + \frac{1}{2}p(p-1)\int_0^T |M_t|^{p-2}\,d\langle M\rangle_t.$$

Since M is bounded, it is a honest martingale and so is the first term in the above equation. Taking expectations, we thus get with $||M||_T = \sup_{t\leq T}|M_t|$ and Hölder's inequality at the last step below

$$\begin{aligned}
\mathbb{E}\,|M_T|^p &= \frac{1}{2}p(p-1)\mathbb{E}\int_0^T |M_t|^{p-2}\,d\langle M\rangle_t \\
&\leq \frac{1}{2}p(p-1)\mathbb{E}\big(\sup_{t\leq T}|M_t|^{p-2}\langle M\rangle_T\big) \\
&\leq \frac{1}{2}p(p-1)(\mathbb{E}\,||M||_T^p)^{1-2/p}(\mathbb{E}\,\langle M\rangle_T^{p/2})^{2/p}.
\end{aligned}$$

Doob's inequality $\mathbb{E}\,||M||_T^p \leq \big(\frac{p}{p-1}\big)^p\mathbb{E}\,|M_T|^p$ (see (2.3)) then gives

$$\mathbb{E}\,||M_T||^p \leq (\frac{p}{p-1})^p\frac{1}{2}p(p-1)(\mathbb{E}\,||M||_T^p)^{1-2/p}(\mathbb{E}\,\langle M\rangle_T^{p/2})^{2/p},$$

from which we obtain

$$\mathbb{E}\,||M_T||^p \leq (\frac{p}{p-1})^{p^2/2}(\frac{1}{2}p(p-1))^{p/2}\mathbb{E}\,\langle M\rangle_T^{p/2},$$

which proves the assertion for bounded M. If M is not bounded we apply the above result to M^{T^n}, where the stopping times T^n are such that M^{T^n} are martingales bounded by n:

$$\mathbb{E}\,||M^{T^n}||_T^p \leq C_p\mathbb{E}\,\langle M^{T^n}\rangle_T^{p/2} = C_p\mathbb{E}\,\langle M\rangle_{T\wedge T^n}^{p/2}.$$

In this inequality the right hand side is less than or equal to $C_p\mathbb{E}\,\langle M\rangle_T^{p/2}$. The result is then obtained by applying Fatou's lemma to the left hand side. $\square$

Lemma 11.2 *Let X be (part of) a weak solution of (10.1). Then for any finite time T and $p \geq 2$ there is a constant C such that*

$$\mathbb{E}\sup_{t\leq T}|X_t|^p \leq Ce^{CT}(1+\mathbb{E}\,|X_0|^p).$$

Proof. Exercise 11.3. □

We now consider the *Cauchy problem.* Let $T>0$ and let functions $f:\mathbb{R}\to\mathbb{R}$, $g,k:[0,T]\times\mathbb{R}\to\mathbb{R}$ be given. Find a (unique) solution $v:[0,T]\times\mathbb{R}$ that belongs to $C^{1,2}([0,T)\times\mathbb{R})$ and that is continuous on $[0,T]\times\mathbb{R}$ such that

$$v_t + \mathcal{L}_t v = kv - g, \tag{11.5}$$

and

$$v(T,\cdot) = f. \tag{11.6}$$

The following *assumptions* are imposed on f, g and k. They are all continuous on their domain, k is nonnegative and f, g satisfy the following growth condition. There exist constants $L>0$ and $\lambda \geq 1$ such that

$$|f(x)| + \sup_{0\leq t\leq T}|g(t,x)| \leq L(1+|x|^{2\lambda}) \tag{11.7}$$

Theorem 11.3 *Let under the stated assumptions the equation (11.5) with terminal condition (11.6) have a solution v satisfying the growth condition*

$$\sup_{0\leq t\leq T}|v(t,x)| \leq M(1+|x|^{2\mu}), \tag{11.8}$$

for some $M>0$ and $\mu \geq 1$. Let $X^{t,x}$ be the weak solution to (10.1) starting at t in x that is unique in law. Then v admits the stochastic representation

$$\begin{aligned} v(t,x) &= \mathbb{E}\,[f(X_T^{x,t})\exp(-\int_t^T k(u,X_u^{t,x})\,du)] \\ &\quad + \mathbb{E}\,[\int_t^T g(r,X_r^{t,x})\exp(-\int_t^r k(u,X_u^{t,x})\,du)\,dr]. \end{aligned} \tag{11.9}$$

on $[0,T]\times\mathbb{R}$ and is thus unique.

Proof. In the proof we simply write X instead of $X^{t,x}$. Let

$$Y_s = v(s,X_s)\exp(-\int_t^s k(u,X_u)\,du) \text{ for } s \geq t.$$

An application of Itô's formula combined with the fact that v solves (11.5) yields

$$\begin{aligned} Y_s - Y_t &= \int_t^s v_x(r,X_r)\sigma(r,X_r)\exp(-\int_t^r k(u,X_u)\,du)\,dW_r \\ &\quad - \int_t^s g(r,X_r)\exp(-\int_t^r k(u,X_u)\,du)\,dr. \end{aligned} \tag{11.10}$$

Notice that $Y_t = v(t,x)$ and that $Y_T = f(X_T)\exp(-\int_t^T k(u,X_u)\,du)$. If the stochastic integral in (11.10) would be a martingale, then taking expectations for $s = T$ would yield the desired result. Since this property is not directly guaranteed under the prevailing assumptions, we will reach our goal by stopping. Let $T^n = \inf\{s \geq t : |X_s| \geq n\}$. Consider the stochastic integral in (11.10) at $s = T \wedge T^n$. It can be written as

$$\int_t^T 1_{\{r \leq T^n\}} v_x(r, X_r^{T^n})\sigma(r, X^{T^n})\exp(-\int_t^r k(u, X_u^{T^n})\,du)\,dW_r.$$

Since v_x and σ are bounded on compact sets, $|X^{T^n}|$ is bounded by n and $k \geq 0$, the integrand in the above stochastic integral is bounded and therefore the stochastic integral has zero expectation. Therefore, if we evaluate (11.10) and take expectations we obtain

$$\mathbb{E}\,Y_{T\wedge T^n} - v(t,x) = -\mathbb{E}\int_t^{T\wedge T^n} g(r,X_r)\exp(-\int_t^r k(u,X_u)\,du)\,dr. \tag{11.11}$$

Consider first the left hand side of (11.11). It can be written as the sum of

$$\mathbb{E}\left(f(X_T)\exp(-\int_t^T k(u,X_u)\,du)1_{\{T\leq T^n\}}\right) \tag{11.12}$$

and

$$\mathbb{E}\left(v(T^n, X_{T^n})\exp(-\int_t^{T^n} k(u,X_u)\,du)1_{\{T>T^n\}}\right). \tag{11.13}$$

The expression (11.12) is bounded in absolute value by $L\mathbb{E}\,(1+|X_T|^{2\lambda})$in view of (11.7). Since $\mathbb{E}\,|X_T|^{2\lambda} \leq Ce^{C(T-t)}(1+|x|^{2\lambda}) < \infty$ in view of lemma 11.2, we can apply the dominated convergence theorem to show that the limit of (11.12) for $n \to \infty$ is equal to $\mathbb{E}\,f(X_T)\exp(-\int_t^T k(u,X_u)\,du)$. The absolute value of (11.13) is bounded from above by

$$M(1+n^{2\mu})\mathbb{P}(T^n \leq T). \tag{11.14}$$

Now,

$$\mathbb{P}(T^n \leq T) = \mathbb{P}(\sup_{t\leq T}|X_t| \geq n) \leq n^{-2p}\mathbb{E}\sup_{t\leq T}|X_t|^{2p}.$$

The expectation here is in view of lemma 11.2 less than or equal to $C(1 + |x|^{2p})e^{C(T-t)}$. Hence we can bound (11.14) from above by $M(1+n^{2\mu})n^{-2p}C(1+|x|^{2p})e^{C(T-t)}$. By choosing $p > \mu$, we see that the contribution of (11.13) vanishes for $n \to \infty$. Summing up,

$$\mathbb{E}\,Y_{T\wedge T^n} \to \mathbb{E}\,f(X_T)\exp(-\int_t^T k(u,X_u)\,du).$$

Next we turn to the right hand side of (11.11). Write it as

$$-\mathbb{E}\int_t^T 1_{\{r\leq T^n\}}g(r,X_r)\exp(-\int_t^r k(u,X_u)\,du)\,dr.$$

The absolute value of the integrand we can bound by $L(1+|X_r|^{2\lambda})$, whose expectation is finite by another application of lemma 11.2. Hence the dominated convergence theorem yields the result. □

Remark 11.4. The nonnegative function k that appears in the Cauchy problem is connected with *exponential killing*. Suppose that we have a process $X^{t,x}$ starting at time t in x and that there is an independent random variable Y with a standard exponential distribution. Let $\partial \notin \mathbb{R}$ be a so-called *coffin* or *cemetery state*. Then we define the process $X^{\partial,t,x}$ by

$$X_s^{\partial,t,x} = \begin{cases} X_s^{t,x} & \text{if } \int_t^s k(u, X_u^{t,x})\,du < Y \\ \partial & \text{if } \int_t^s k(u, X_u^{t,x})\,du \geq Y \end{cases}$$

Functions f defined on $\mathbb{R}$ will be extended to $\mathbb{R} \cup \{\partial\}$ by setting $f(\partial) = 0$. If $X^{t,x}$ is a Markov process (solving equation (10.1) for instance), then $X^{\partial,t,x}$ is a Markov process as well. If, in the terminology of the theory of Markov processes, $X^{t,x}$ has generator $\mathcal{L}$, then $X^{\partial,t,x}$ has generator $\mathcal{L}^k$ defined by $\mathcal{L}_t^k f(t,x) = \mathcal{L}_t f(t,x) - k(t,x)f(t,x)$. Furthermore we have

$$\mathbb{E}\, f(X_s^{\partial,t,x}) = \mathbb{E}\, f(X_s^{t,x}) \exp(-\int_t^u k(u, X_u^{t,x})\,du. \tag{11.15}$$

The above considerations enable one to connect the theory of solving the Cauchy problem with $k = 0$ to solving the problem with arbitrary k by jumping in the representation from the process $X^{t,x}$ to $X^{\partial,t,x}$.

11.2 Exercises

11.1 Show that the growth conditions on f and g are not needed in order to prove theorem 11.3, if we assume instead that next to k also f and g are nonnegative.

11.2 Consider equation (11.5) with $k = 0$ and $g = 0$. The equation is then called *Kolmogorov's backward equation*. Let f be continuous with compact support. Show that $v^{r,f}(t,x) = \mathbb{E}\, f(X_r^{t,x})$ satisfies Kolmogorov's backward equation for all $r > t$. Suppose that there exists a function p of four variables t, x, r, y such that for all f that are continuous with compact support one has $v^{r,f}(t,x) = \int_{\mathbb{R}} f(y)p(t,x,r,y)\,dy$ and $\lim_{t\uparrow r} v^{r,f}(t,x) = f(x)$. Show that for fixed r, y the function $(t,x) \to p(t,x,r,y)$ satisfies Kolmogorov's backward equation. What is the interpretation of the function p?

11.3 Prove lemma 11.2. *Hint:* Proceed as in exercise 10.12 and use the Doob and Burkholder-Davis-Gundy inequalities.

11.4 The Black-Scholes partial differential equation is

$$v_t(t,x) + \frac{1}{2}\sigma^2 x^2 v_{xx}(t,x) + rxv_x(t,x) = rv(t,x),$$

with constants $r, \sigma > 0$. Let X be the price process of some financial asset and $T > 0$ a finite time horizon. A simple contingent claim is a measurable function of X_T, $f(X_T)$ say, representing the value at time T of some *derived* financial product. The pricing problem is to find the right price at any time t prior to T that should be charged to trade this claim. Clearly, at time $t = T$, this price should be equal to $f(X_T)$. The theory of Mathematical Finance dictates that (under the appropriate assumptions) this price is equal to $v(t, X_t)$, with v a solution to the Black-Scholes equation. Give an explicit solution for the case of a *European call option*, i.e. $f(x) = \max\{x - K, 0\}$, where $K > 0$ is some positive constant.

11.5 Here we consider the Cauchy problem with an initial condition. We have the partial differential equation

$$u_t + ku = \mathcal{L}_t u + g,$$

with the initial condition $u(0, \cdot) = f$. Formulate sufficient conditions such that this problem has a unique solution which is given by

$$\begin{aligned} u(t,x) &= \mathbb{E}\, f(X_t^x) \exp(-\int_0^t k(u, X_u^x)\, du) \\ &\quad + \mathbb{E} \int_0^t g(s, X_s^x) \exp(-\int_0^s k(u, X_u^x)\, du)\, ds, \end{aligned}$$

where X^x is a solution to (10.1) with $X_0 = x$.

11.6 Consider equation (11.5) with $g = 0$,

$$v_t + \mathcal{L}_t v = kv.$$

A fundamental solution to this equation is a function $(t, x, s, y) \mapsto p(t, x; s, y)$ such that for all $s > t$ and continuous f with compact support the function $(t, x) \mapsto v(t, x; s) = \int_{\mathbb{R}} f(y) p(t, x; s, y)\, dy$ satisfies this equation and such that $\lim_{t \uparrow s} v(t, x; s) = f(x)$. Assume that a fundamental solution exists. Show that the solution to the Cauchy problem (11.5), (11.6) takes the form

$$v(t,x) = \int_{\mathbb{R}} p(t, x; T, y) f(y)\, dy + \int_t^T \int_{\mathbb{R}} p(t, x; s, y) g(s, y)\, dy\, ds.$$

Index