LECTURE NOTES

Graduate Course on

Stochastic Differential Equations

With 1 figure

By Patrik Albin

18th July 2009

Figure 1: We.

Contents

1 Lecture 1, Wednesday March 18

1.1 Stochastic processes

Definition 1.1. *A* stochastic process $X = \{X(t)\}_{t \in T}$ *with* parameter set T *is a function* $X : \Omega \times T \to \mathbb{R}$ *(or possibly to some other measurable space than* $\mathbb{R}$*) such that* $X(\cdot, t) : \Omega \to \mathbb{R}$ *(or some other space ...) is a random variable for each* $t \in T$*.*

The dependence of $\omega \in \Omega$ for a stochastic process X is often suppressed in the notation (as we did already in the definition), so that we write $X(t)$ or $\{X(t)\}_{t \in T}$ instead of $X(\omega, t)$ or $\{X(\omega, t)\}_{(\omega,t) \in \Omega \times T}$.

Definition 1.2. *The* finite dimensional distributions *(fidi's)* $\{F_{X(t_1),\ldots,X(t_n)} : t_1, \ldots, t_n \in T, n \in \mathbb{N}\}$ *of a stochastic process* $\{X(t)\}_{t \in T}$ *are given by*

$$F_{X(t_1),\ldots,X(t_n)}(x_1, \ldots, x_n) = \mathbf{P}\{X(t_1) \leq x_1, \ldots, X(t_n) \leq x_n\} \quad \textit{for } x_1, \ldots, x_n \in \mathbb{R}.$$

It is natural to belive that a stochastic process is more or less "determined" by its univariate marginal distributions $F_{X(t)}(x) = \mathbf{P}\{X(t) \leq x\}$ for $x \in \mathbb{R}$, for each $t \in T$. But this is not true at all, see Exercise 1 below. In fact, in general, not even all the fidi's of a process are sufficient for that purpose, see Exercise 2 below. However, for "decent" processes, such as the ones we will encounter, the latter does not happen.

Exercise 1. Consider the stochastic process $X(t) = \xi$ for $t \in \mathbb{R}$, where ξ is a single standard normal $\mathrm{N}(0,1)$ distributed random variable. Let $Y(t)$ be a stochastic process that is $\mathrm{N}(0,1)$ distributed at each $t \in \mathbb{R}$, but with all random values of the process at different times independent of each other. Find the univariate marginal distributions $F_{X(t)}$ and $F_{Y(t)}$. Pick a "typical" $\omega \in \Omega$ and, for that choice of ω, plot a likely appearance of the graphs, the so called *realisations*,

$$\mathbb{R} \ni t \curvearrowright X(t) = X(\omega, t) \in \mathbb{R} \quad \text{and} \quad \mathbb{R} \ni t \curvearrowright Y(t) = Y(\omega, t) \in \mathbb{R}.$$

Definition 1.3. *Two stochastic processes* $\{X(t)\}_{t \in T}$ *and* $\{Y(t)\}_{t \in T}$ *defined on a common probability space are said to be* versions *of each other if* $\mathbf{P}\{X(t) = Y(t)\} = 1$ *for each* $t \in T$*.*

Some authors make the stronger requirement that $\mathbf{P}\{X(t) = Y(t) \text{ for all } t \in T\} = 1$ to call X and Y versions of each other. Albeit this is a stronger requirement in general, these two definitions will coincide for the processes we will encounter.

Probabilities of events for processes X and Y that are versions of each other need not be equal, see Exercise 2 below. However, usually there is no need to regard processes that are versions of each other, but not equal, as really different, but rather as different expressions one single process can take.

Exercise 2. Find two stochastic processes $\{X(t)\}_{t\in[0,1]}$ and $\{Y(t)\}_{t\in[0,1]}$ that have common fidi's, and that are versions of each other, but that satisfy

$$\mathbf{P}\{X(t)\neq Y(t) \ \text{ for some } \ t\in[0,1]\}=1.$$

There exists a rather difficult theory of *separability* for stochastic processes, that adresses the issue of what probailities for a process are determined by the fidi's. This theory is very important in stochastic process theory in general, but we will not need it as our processes will be sufficiently nice to be understood by more direct methods.

1.2 Gaussian processes

Definition 1.4. *A stochastic process* $\{X(t)\}_{t\in T}$ *is* Gaussian (normal) *if all fidi's are multivariate Gaussian (normal) distributed, that is, if for each choice of constants* $a_1,\ldots,a_n\in\mathbb{R}$ *parameters* $t_1,\ldots,t_n\in T$ *and* $n\in\mathbb{N}$ *the linear combination*

$$\sum_{i=1}^{n} a_i X(t_i) \quad \textit{is univariate Gaussian (normal) distributed}$$

Exercise 3. Prove that an $\mathbb{R}^n$-valued random variable $(X_1,\ldots,X_n)$ is multivariate Gaussian distributed if and only if $\{X_i\}_{i\in\{1,\ldots,n\}}$ is a Gaussian process.

Definition 1.5. *The* mean function $m:T\to\mathbb{R}$ *and* covariance function $r:T\times T\to\mathbb{R}$ *for a stochastic process* $\{X(t)\}_{t\in T}$ *are defined as* $m(t)=\mathbf{E}\{X(t)\}$ *and* $r(s,t)=\mathbf{Cov}\{X(s),X(t)\}$, *respectively (whenever the right-hand sides make sense).*

Exercise 4. Show that the fidi's of a Gaussian process $\{X(t)\}_{t\in T}$ with mean function m and covariance function r have characteristic functions given by

$$\mathbf{E}\big\{\mathrm{e}^{i(s_1X(t_1)+\ldots+s_nX(t_n))}\big\}$$
$$=\exp\left\{i\,(s_1\ \ldots\ s_n)\begin{pmatrix}m(t_1)\\ \vdots\\ m(t_n)\end{pmatrix}-\frac{1}{2}(s_1\ \ldots\ s_n)\,\big(r(t_i,t_j)\big)_{i,j}\begin{pmatrix}s_1\\ \vdots\\ s_n\end{pmatrix}\right\}.$$

Theorem 1.6. *The fidi's of a Gaussian process are determined by the mean function m and the covariance function r of the process.*

Proof. By Exercise 4 m and r determine the characteristic functions of the fidi's. □

Corollary 1.7. *If $\{X(t)\}_{t\in T}$ is a Gaussian process and R and S are subsets of T, then the processes $\{X(t)\}_{t\in R}$ and $\{X(t)\}_{t\in S}$ are independent if and only if* $\mathbf{Cov}\{X(r),X(s)\}=0$ *for all $r\in R$ and $s\in S$.*

Proof. The implication to the right is immediate. For the implication to the left, assume that $\mathbf{Cov}\{X(r),X(s)\}=0$ for $r\in R$ and $s\in S$. Note that X takes non-random values on $R\cap S$, as variances there are zero. Hence it is enough to prove that $\{X(t)\}_{t\in R\setminus S}$ and $\{X(t)\}_{t\in S\setminus R}$ are independent. For this, in turn (by the very definition of what that independence amounts to), it is enough to prove that $\{X(r_i)\}_{i\in\{1,\ldots,m\}}$ and $\{X(s_j)\}_{j\in\{1,\ldots,n\}}$ are independent for every choice of $r_1,\ldots,r_m\in R\setminus S$ and $s_1,\ldots,s_n\in S\setminus R$. Using Theorem 1.6, this in turn follows from the fact that the process $\{X(t)\}_{t\in\{r_1,\ldots,r_m\}\cup\{s_1,\ldots,s_n\}}$ has the same fidi's as when composed by two components $\{X(r_i)\}_{i\in\{1,\ldots,m\}}$ and $\{X(s_j)\}_{j\in\{1,\ldots,n\}}$ with the requested independence properties, as the latter two then also have zero covariances between them. □

Example 1.8. Let $\{X(t)\}_{t\geq 0}$ be a Gaussian process that is continuous for each ω. Then the process $Y(t)=\int_0^t X(s)\,ds$, $s\geq 0$, is also a Gaussian process, since

$$\sum_{i=1}^n a_i Y(t_i) \leftarrow \sum_{i=1}^n a_i \sum_{s_j\leq t_i} X(s_{j-1})\,(s_j-s_{j-1})$$

as we consider finer and finer grids $0=s_0<s_1<\ldots<s_k=\max_{1\leq i\leq n} t_i$ such that $\max_{1\leq j\leq k} s_j-s_{j-1}\downarrow 0$ (Riemann sum). Now, the sum on the right-hand side can be rewritten as a linear combination of process values of X, and is thus Gaussian distributed. As Gaussian random variables can only converge to Gaussian limits, see Exercise 5 below, it follows that every linear combination $\sum_{i=1}^n a_i Y(t_i)$ of process values of Y is Gaussian distributed. Hence Y is a Gaussian process.

Exercise 5. Show that if a sequence of Gaussian random variables converges to a limit random variable (in distribution, in probability, in mean-square, or almost surely), then the limit random variable is Gaussian.

Exercise 6. Find an example of a stochastic process that has Gaussian univariate marginal distributions, but that is not Gaussian.

Exercise 7. Let the Gaussian process X in Example 1.8 have mean function m and covariance function r. Find the mean function and the covariance function of the process Y.

1.3 Brownian motion (BM)

Definition 1.9. *A stochastic process* $\{B(t)\}_{t\geq 0}$ *with* $B(0)=0$ *is a* Brownian motion (BM)*, which is the same thing as a* Wiener process*, if it has the following properties:*

∘ (Continuity) $[0,\infty) \ni t \curvearrowright B(\omega,t) \in \mathbb{R}$ *is continuous for all (or almost all)* $\omega \in \Omega$*;*

∘ (Independent increments) $B(t)-B(s)$ *is independent of* $\{B(r)\}_{r\in[0,s]}$ *for* $0 \leq s < t$*;*

∘ (Stationary normal increments) $B(t)-B(s)$ *is* $\mathrm{N}(0,t-s)$*-distributed for* $0 \leq s \leq t$*.*

Exercise 8. Show that B is a zero-mean Gaussian process with covariance function $\mathbf{Cov}\{B(s),B(t)\} = s\wedge t = \min\{s,t\}$.

Exercise 9. It is customary to use B^x as notation for a process which has the properties of BM, with the only exception that $B^x(0) = x$ for a (non-random) constant $x \in \mathbb{R}$: Show that $\{B^x(t)\}_{t\geq 0}$ has the same fidi's as $\{B(t)+x\}_{t\geq 0}$.

Exercise 10. Prove the following elementary formulas $\mathbf{E}\{B(t)-B(s)\} = 0$, $\mathbf{E}\big\{(B(t)-B(s))^2\big\} = t-s$ and $\mathbf{Var}\big\{(B(t)-B(s))^2\big\} = 2\,(t-s)^2$.

Exercise 11. Plot a few sample paths of BM in a computer.

Theorem 1.10. *In the sense of convergence in mean-square, BM has quadratic variation over the interval* $[s,t] \subseteq [0,\infty)$ *given by*

$$\lim_{\max_{1\leq i\leq n} t_i - t_{i-1}\downarrow 0} \sum_{i=1}^{n} (B(t_i)-B(t_{i-1}))^2 = t-s,$$

where $s = t_0 < t_1 < \ldots < t_n = t$ *are finer and finer partitions of* $[s,t]$*.*

Exercise 12. Prove Theorem 1.10 by means of considering expectations and variances.

Note that the quadratic variation of any "nice" function over a finite interval is zero. For example, any continuously differentiable function has zero quadratic variation, see Exercise 95 below.

Theorem 1.11. *BM has infinite variation over any interval* $[s,t] \subseteq [0,\infty)$ *with length* $t-s>0$*, that is,*

$$\lim_{\max_{1\le i\le n} t_i - t_{i-1}\downarrow 0} \sum_{i=1}^{n} |B(t_i)-B(t_{i-1})| = \infty \quad \text{with probability 1,}$$

where $s=t_0<t_1<\ldots<t_n=t$ *are finer and finer partitions of* $[s,t]$.

Exercise 13. Prove Theorem 1.11 by means of considering expectations and variances.

Note that the variation of any "nice" function over a finite interval is finite. For example, any continuously differentiable function has finite variation, see Exercise 95 below.

2 Lecture 2, Monday March 23

2.1 Non-differentiability of BM

Theorem 2.1. *Let $\{X(t)\}_{t\geq 0}$ be a stochastic process with independent increments such that, for some constant $\varepsilon > 0$, it holds that*

$$\lim_{n\to\infty} n^{\varepsilon} \sup_{t\geq 0} \mathbf{P}\big\{|X(t+1/n) - X(t)| \leq K/n\big\} = 0 \quad \text{for each } K > 0. \tag{2.1}$$

Then the process X is not differentiable anywhere with probability 1.

Proof. Pick an $N \in \mathbb{N}$ such that $N\varepsilon \geq 2$. Notice that if X is differentiable at some $s \geq 0$, then we have

$$|X(t) - X(s)| \leq \ell(t-s) \leq \ell(N+2)/n \quad \text{for } t \in (s, s+(N+2)/n),$$

for all sufficiently large $\ell, n \in \mathbb{N}$. Choosing $k \in \mathbb{N}$ such that $k/n, \ldots, (k+N)/n \in (s, s+(N+2)/n)$, this gives

$$|X((i+1)/n) - X(i/n)| \leq |X((i+1)/n) - X(s)| + |X(s) - X(i/n)| \leq 2\ell(N+2)/n$$

for $i = k, \ldots, k+N-1$. Hence the event that X is differentiable somewhere is contained in the event

$$\bigcup_{\ell=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \bigcup_{k=1}^{n^2} \bigcap_{i=k}^{k+N-1} \bigg\{|X((i+1)/n) - X(i/n)| \leq \frac{2\ell(N+2)}{n}\bigg\}.$$

Therefore it is enough to prove that

$$\mathbf{P}\bigg\{\bigcap_{n=m}^{\infty} \bigcup_{k=1}^{n^2} \bigcap_{i=k}^{k+N-1} \bigg\{|X((i+1)/n) - X(i/n)| \leq \frac{2\ell(N+2)}{n}\bigg\}\bigg\} = 0, \tag{2.2}$$

which in turn holds if

$$\mathbf{P}\bigg\{\bigcup_{k=1}^{n^2} \bigcap_{i=k}^{k+N-1} \bigg\{|X((i+1)/n) - X(i/n)| \leq \frac{2\ell(N+2)}{n}\bigg\}\bigg\} \to 0 \quad \text{as } n\to\infty, \tag{2.3}$$

as the probability on the left-hand side in (2.2) is bounded by that on the left-hand side in (2.3) for $n \geq m$. However, by Boole's inequality together with independence of increments and (2.1), the probability on the left-hand side in (2.3) is at most

$$\begin{aligned}
&\sum_{k=1}^{n^2} \prod_{i=k}^{k+N-1} \mathbf{P}\bigg\{|X((i+1)/n) - X(i/n)| \leq \frac{2\ell(N+2)}{n}\bigg\} \\
&\qquad \leq n^{2-N\varepsilon} \bigg(n^{\varepsilon} \sup_{t\geq 0} \mathbf{P}\bigg\{|X(t+1/n) - X(t)| \leq \frac{2\ell(N+2)}{n}\bigg\}\bigg)^{N} \\
&\qquad \to 0 \quad \text{as } n\to\infty.
\end{aligned}$$

□

Corollary 2.2. *BM is not differentiable anywhere with probability* 1.

Proof. The definition of BM gives the hypothesis of the previous theorem except (2.1). However, we get (2.1) for $\varepsilon \in (0,1/2)$ from observing that

$$\mathbf{P}\big\{|X(t+1/n)-X(t)| \leq K/n\big\} = \mathbf{P}\big\{\{|\mathrm{N}(0,1/n)| \leq K/n\big\} \leq \frac{2\,K}{\sqrt{2\pi\,n}}. \quad \square$$

Exercise 14. Plot the approximative derivative process $\{(B(t+h)-B(t))/h\}_{t\geq 0}$ in a computer for a small $h>0$ to illustrate the non-differentiability of BM.

Exercise 15. The (non-existing) derivative process $\{B'(t)\}_{t\geq 0}$ of BM is what electrical engineering people use as "white noise" through out their science, more or less: Can you say anything about why that is so?

⋆2.2 Existence of BM[1]

⋆Theorem 2.3. *BM exists.*

⋆*Proof.* We prove the existence of BM $\{B(t)\}_{t\in[0,1]}$. To that end it is sufficient show that there exists a zero-mean Gaussian process B with covariance function $\mathbf{Cov}\{B(s),B(t)\} = s\wedge t$ that is continuous with probability 1, because then we have $\mathbf{Cov}\{B(r),B(t)-B(s)\} = 0$ for $0\leq r\leq s\leq t$, so that increments are independent by Corollary 1.7. Further, this gives $\mathbf{Var}\{B(0)\} = 0$, so that $B(0) = 0$, as well as $\mathbf{Var}\{B(t)-B(s)\} = t-s$, so that $B(t)-B(s)$ is $\mathrm{N}(0,t-s)$-distributed.

Let $\xi_0, \xi_1, \ldots$ be independent $\mathrm{N}(0,1)$-distributed random variables, and set

$$B(t) = \sum_{k=0}^{\infty} \frac{\sqrt{2}}{\pi}\,\frac{2}{2k+1}\,\sin\Big(\frac{(2k+1)\pi\,t}{2}\Big)\xi_k \quad \text{for } t\in[0,1], \tag{2.4}$$

where the convergence is in the mean-square sense. Note that, by the Cauchy criterion, this limit is well-defined if and only if

$$\mathbf{E}\bigg\{\bigg[\sum_{k=0}^{m} \frac{\sqrt{2}}{\pi}\,\frac{2}{2k+1}\,\sin\Big(\frac{(2k+1)\pi\,t}{2}\Big)\xi_k - \sum_{\ell=0}^{n} \frac{\sqrt{2}}{\pi}\,\frac{2}{2\ell+1}\,\sin\Big(\frac{(2\ell+1)\pi\,t}{2}\Big)\xi_\ell\bigg]^2\bigg\} \to 0$$

as $m,n\to\infty$. This in turn holds as the expression on the left-hand side is equal to

$$\mathbf{E}\bigg\{\bigg[\sum_{k=(m\wedge n)+1}^{(m\vee n)} \frac{\sqrt{2}}{\pi}\,\frac{2}{2k+1}\,\sin\Big(\frac{(2k+1)\pi\,t}{2}\Big)\xi_k\bigg]^2\bigg\} = \sum_{k=(m\wedge n)+1}^{(m\vee n)} \frac{8\,\sin\big((2k+1)\pi\,t/2\big)^2}{\pi^2(2k+1)^2}.$$

[1] Material that is marked ⋆ is non-mandatory. In particular, examination procedures of the course will not assume that you have read ⋆-marked material.

By symmetry in (2.4) we have $\mathbf{E}\{B(t)\} = 0$. To find the covariance function of B we use the fact that covariances commute with mean-square limits to obtain

$$\mathbf{Cov}\{B(s),B(t)\} = \sum_{k=0}^{\infty} \frac{8\,\sin\big((2k+1)\pi\,s/2\big)\,\sin\big((2k+1)\pi\,t/2\big)}{\pi^2\,(2k+1)^2}.$$

By the elementary identity $2\sin(x)\sin(y) = \cos(x-y) - \cos(x+y)$ together with *Exercise 16 below, the right-hand side of this identity in turn is equal to

$$\begin{aligned}\sum_{k=0}^{\infty} \frac{4}{\pi^2\,(2k+1)^2}&\left[\cos\bigg(\frac{(2k+1)\pi\,(s-t)}{2}\bigg) - \cos\bigg(\frac{(2k+1)\pi\,(s+t)}{2}\bigg)\right]\\ &= \frac{1-|t-s|}{2} - \frac{1-|t+s|}{2}\\ &= s\wedge t \quad \text{for } s,t\in[0,1],\end{aligned}$$

which is the covariance function desired.

As the process B quite obviously is Gaussian, recall Exercise 5, it only remains to prove that B is continuous with probability 1. To that end we notice that

$$\begin{aligned}&\mathbf{P}\bigg\{\sum_{k=0}^{\infty}\frac{\sqrt{2}}{\pi}\,\frac{2}{2k+1}\,\sin\bigg(\frac{(2k+1)\pi\,t}{2}\bigg)\xi_k \text{ is continuous for } t\in[0,1]\bigg\}\\ &\geq \mathbf{P}\bigg\{\sum_{n=0}^{\infty}\sum_{k=2^n-1}^{2^{n+1}-2}\frac{\sqrt{2}}{\pi}\,\frac{2}{2k+1}\,\sin\bigg(\frac{(2k+1)\pi\,t}{2}\bigg)\xi_k \text{ converges uniformly for } t\in[0,1]\bigg\}\\ &\geq 1-\mathbf{P}\bigg\{\sup_{t\in[0,1]}|X_n(t)|>2^{-n/8} \text{ for infinitely many } n\bigg\}\end{aligned} \tag{2.5}$$

(see also Exercise 47 below), where X_n is the zero-mean Gaussian process given by

$$X_n(t) = \sum_{k=2^n-1}^{2^{n+1}-2}\frac{\sqrt{2}}{\pi}\,\frac{2}{2k+1}\,\sin\bigg(\frac{(2k+1)\pi\,t}{2}\bigg)\xi_k \quad \text{for } t\in[0,1].$$

Now, by the elementary identity $\sin(x)-\sin(y) = 2\cos((x+y)/2)\sin((x-y)/2)$ together with the elementary inequality $|\sin(x)|\leq|x|^{1/4}$, we readily get

$$\begin{aligned}\mathbf{E}\big\{(X_n(t)-X_n(s))^2\big\} &= \sum_{k=2^n-1}^{2^{n+1}-2}\frac{32\,\cos\big((2k+1)\pi\,(t+s)/4\big)^2\,\sin\big((2k+1)\pi\,(t-s)/4\big)^2}{\pi^2\,(2k+1)^2}\\ &\leq \sum_{k=2^n-1}^{2^{n+1}-2}\frac{16\,|t-s|^{1/2}}{\pi^{3/2}\,(2k+1)^{3/2}}\\ &\leq \frac{16\,2^n\,|t-s|^{1/2}}{\pi^{3/2}\,(2^{n+1}-1)^{3/2}}.\end{aligned} \tag{2.6}$$

As X_n is continuous and symmetrically distributed with $X_n(0)=0$, we therefore get

$$s_n \equiv \mathbf{P}\big\{\sup_{t\in[0,1]}|X_n(t)|>2^{-n/8}\big\}$$

$$
\begin{aligned}
&\le 2\,\mathbf{P}\Big\{\bigcup_{k=0}^{\infty}\bigcup_{\ell=1}^{2^k}\{X_n(2^{-k}\ell)>2^{-n/8}\}\Big\}\\
&\le 2\,\mathbf{P}\Big\{\bigcup_{k=0}^{\infty}\bigcup_{\ell=1}^{2^k}\Big\{X_n(2^{-k}\ell)>2^{-n/8-1}\Big(1+(1-2^{-1/8})\textstyle\sum_{j=0}^{k}2^{-j/8}\Big)\Big\}\Big\}\\
&= 2\,\mathbf{P}\Big\{X_n(0)>2^{-n/8-1}\Big(1+(1-2^{-1/8})\Big)\Big\}\\
&\qquad +2\sum_{k=1}^{\infty}\mathbf{P}\Big\{\bigcup_{\ell=1}^{2^k}\Big\{X_n(2^{-k}\ell)>2^{-n/8-1}\Big(1+(1-2^{-1/8})\textstyle\sum_{j=0}^{k}2^{-j/8}\Big)\Big\},\\
&\qquad\qquad \bigcap_{m=0}^{k-1}\bigcap_{\ell=1}^{2^m}\Big\{X_n(2^{-m}\ell)\le 2^{-n/8-1}\Big(1+(1-2^{-1/8})\textstyle\sum_{j=0}^{m}2^{-j/8}\Big)\Big\}\Big\}\\
&\le 0+2\sum_{k=1}^{\infty}\sum_{\ell=0}^{2^{k-1}-1}\mathbf{P}\Big\{X_n(2^{-k}(2\ell+1))>2^{-n/8-1}\Big(1+(1-2^{-1/8})\textstyle\sum_{j=0}^{k}2^{-j/8}\Big),\\
&\qquad\qquad X_n(2^{-k+1}\ell)\le 2^{-n/8-1}\Big(1+(1-2^{-1/8})\textstyle\sum_{j=0}^{k-1}2^{-j/8}\Big)\Big\}\\
&\le 2\sum_{k=1}^{\infty}\sum_{\ell=0}^{2^{k-1}-1}\mathbf{P}\Big\{X_n(2^{-k}(2\ell+1))-X_n(2^{-k}2\ell)>2^{-n/8-1}(1-2^{-1/8})2^{-k/8}\Big\}\\
&= 2\sum_{k=1}^{\infty}\sum_{\ell=0}^{2^{k-1}-1}\mathbf{P}\left\{\mathrm{N}(0,1)>\frac{2^{-n/8-1}(1-2^{-1/8})2^{-k/8}}{\sqrt{\mathbf{E}\{(X_n(2^{-k}(2\ell+1))-X_n(2^{-k}2\ell))^2\}}}\right\}\\
&\le \sum_{k=1}^{\infty}2^k\,\mathbf{P}\left\{\mathrm{N}(0,1)>\frac{2^{-n/8-1}(1-2^{-1/8})2^{-k/8}\,\pi^{3/4}\,(2^{n+1}-1)^{3/4}}{4\,2^{n/2}\,2^{-k/4}}\right\},
\end{aligned}
$$

where we used (2.6) for the last inequality. Hence $\sum_{n=0}^{\infty}s_n<\infty$ (see *Exercise 17 below), so that the right-hand side of (2.5) is 1 by the Borel-Cantelli lemma. □

***Exercise 16.** Show the indentity

$$
\sum_{k=0}^{\infty}\frac{4}{\pi^2\,(2k+1)^2}\cos\bigg(\frac{(2k+1)\pi t}{2}\bigg)=\frac{1-t}{2}\quad\text{for } t\in[0,2],
$$

for example, by means of calculating the one-sided Laplace tranform on both sides.

***Exercise 17.** Why is $\sum_{n=0}^{\infty}s_n<\infty$ at the end of the proof of Theorem 2.3?

2.3 Introduction to martingales

We will learn a lot of martingale theory in this course. Here are the first few steps:

Definition 2.4. *A family $\{\mathcal{F}_t\}_{t\in T}$, $T\subseteq[0,\infty)$, of σ-algebras on a probability space $(\Omega,\mathcal{F},\mathbf{P})$, $\mathcal{F}_t\subseteq\mathcal{F}$, that is increasing, $\mathcal{F}_s\subseteq\mathcal{F}_t$ for $T\ni s<t\in T$, is called a* filtration.

Definition 2.5. *A stochastic process* $\{X(t)\}_{t\in T}$ *is* adapted *to a filtration* $\{\mathcal{F}_t\}_{t\in T}$ *if* $X(t)$ *is* $\mathcal{F}_t$*-measurable for each* $t\in T$.

Definition 2.6. *Let* $\{X(t)\}_{t\in T}$ *be a stochastic process that is adapted to a filtration* $\{\mathcal{F}_t\}_{t\in T}$ *and that is integrable,* $\mathbf{E}\{|X(t)|\}<\infty$ *for* $t\in T$. *We say that* $\{(X(t),\mathcal{F}_t)\}_{t\in T}$ *is a* martingale *if*

$$\mathbf{E}\{X(t)\,|\,\mathcal{F}_s\} = X(s) \quad \textit{for } T\ni s<t\in T.$$

We say that $\{(X(t),\mathcal{F}_t)\}_{t\in T}$ *is a* submartingale *if*

$$\mathbf{E}\{X(t)\,|\,\mathcal{F}_s\} \geq X(s) \quad \textit{for } T\ni s<t\in T.$$

We say that $\{(X(t),\mathcal{F}_t)\}_{t\in T}$ *is a* supermartingale *if*

$$\mathbf{E}\{X(t)\,|\,\mathcal{F}_s\} \leq X(s) \quad \textit{for } T\ni s<t\in T.$$

Exercise 18. If $\{\xi_i\}_{i=1}^{\infty}$ are independent random variables with finite and zero/positive/negative expected values, then the process $\{\sum_{i=1}^{n}\xi_i\}_{n\in\mathbb{N}}$ is a martingale/submartingale/supermartingale with respect to itself.

Exercise 19. Show that BM is a martingal with respect to itself.

Exercise 20. Show that the process $\{B(t)^2-t\}_{t\geq 0}$ is a martingale with respect to the filtration $\{\sigma(B(s):0\leq s\leq t)\}_{t\geq 0}$.

Exercise 21. Show that the process $\{\mathrm{e}^{cB(t)-c^2t/2}\}_{t\geq 0}$ is a martingale with respect to the filtration $\{\sigma(B(s):0\leq s\leq t)\}_{t\geq 0}$ for any constant $c\in\mathbb{R}$.

Exercise 22. Show that if X is a martingale and $g:\mathbb{R}\to\mathbb{R}$ a convex function such that the process $g(X)$ is integrable, then $g(X)$ is a submartingale.

Exercise 23. Show that if X is a submartingale and $g:\mathbb{R}\to\mathbb{R}$ a convex non-decreasing function such that the process $g(X)$ is integrable, then $g(X)$ is a submartingale.

Definition 2.7. *A martingale/submartingale/supermartingale* $\{X(t)\}_{t\in T}$ *that is a martingale/submartingale/supermartingale with respect to the filtration* $\{\sigma(X(s)):T\ni s\leq t)\}_{t\in T}$ *is called a* martingale/submartingale/supermartingale with respect to itself.

Exercise 24. Prove that every martingale/submartingale/supermartingale is a martingale/submartingale/supermartingale with respect to itself.

3 Lecture 3, Wednesday March 25

3.1 Doob-Kolmogorov inequality

The following powerful inequality will be used to derive two more inequalities, that in turn will be crucial for the build up of the theory.

Theorem 3.1. (Doob inequality) *For a non-negative right-continuous submartingale $\{Y(t)\}_{t\in[0,T]}$, we have*

$$\mathbf{P}\Big\{\sup_{0\le t\le T} Y(t)\ge\lambda\Big\}\le\frac{\mathbf{E}\big\{Y(T)\,I_{\{\sup_{0\le t\le T}Y(t)\ge\lambda\}}\big\}}{\lambda}\quad\textit{for }\lambda>0. \tag{3.1}$$

Proof. By Exercise 25 below, it is sufficient to prove that

$$\mathbf{P}\Big\{\sup_{0\le t\le T} Y(t)>\lambda\Big\}\le\frac{\mathbf{E}\big\{Y(T)\,I_{\{\sup_{0\le t\le T}Y(t)>\lambda\}}\big\}}{\lambda}\quad\text{for }\lambda>0. \tag{3.2}$$

By Exercise 26 below and the right-contiuity of Y, (3.2) holds if

$$\mathbf{P}\Big\{\max_{0\le i\le n} Y(t_i)>\lambda\Big\}\le\frac{\mathbf{E}\big\{Y(T)\,I_{\{\max_{0\le i\le n}Y(t_i)>\lambda\}}\big\}}{\lambda}\quad\text{for }\lambda>0, \tag{3.3}$$

for every partition $0=t_0<t_1<\ldots<t_n=T$ of $[0,T]$. However, writing $\tau=\min\{t_i : Y(t_i)>\lambda\}$ and noting that $\{\tau=t_i\}\in\mathcal{F}_{t_i}$, (3.3) in turn follows as

$$\begin{aligned}
\mathbf{E}\big\{Y(T)\,I_{\{\max_{0\le i\le n}Y(t_i)>\lambda\}}\big\} &= \sum_{i=0}^{n}\int_{\{\tau=t_i\}} Y(T)\,d\mathbf{P}\\
&= \sum_{i=0}^{n}\int_{\{\tau=t_i\}} \mathbf{E}\{Y(T)\,|\,\mathcal{F}_{t_i}\}\,d\mathbf{P}\\
&\ge \sum_{i=0}^{n}\int_{\{\tau=t_i\}} Y(t_i)\,d\mathbf{P}\\
&= \mathbf{E}\big\{Y(\tau)\,I_{\{\max_{0\le i\le n}Y(t_i)>\lambda\}}\big\}\\
&\ge \lambda\,\mathbf{P}\Big\{\max_{0\le i\le n}Y(t_i)>\lambda\Big\}.
\end{aligned}$$

□

Corollary 3.2. (Doob-Kolmogorov inequality) *Pick a constant $p\ge1$. For a right-continuous martingale $\{X(t)\}_{t\in[0,T]}$ such that the process $\{|X(t)|^p\}_{t\in[0,T]}$ is integrable, we have*

$$\mathbf{P}\Big\{\sup_{0\le t\le T}|X(t)|\ge\lambda\Big\}\le\frac{\mathbf{E}\{|X(T)|^p\}}{\lambda^p}\quad\textit{for }\lambda>0.$$

Proof. As the function $|\cdot|^p$ is convex the corollary follows from applying Doob's inequality to the submartingale $Y=|X|^p$ (recall Exercise 22), see Exercise 28 below. □

Exercise 25. Show that the inequality (3.1) follows from the inequality (3.2).

Exercise 26. Explain in detail how (3.2) follows from (3.3). Also, explain how we can conclude that $\sup_{0\leq t\leq T} Y(t)$ really is a well-defined random variable.

Exercise 27. Let $\{Y(t)\}_{t\in[0,T]}$ be a right-continuous process. Explain why we cannot in general deduce that

$$\mathbf{P}\left\{\sup_{0\leq t\leq T} Y(t)\geq\lambda\right\}\leq\frac{\mathbf{E}\{Y(T)\}}{\lambda}\quad\text{for }\lambda>0,$$

if

$$\mathbf{P}\left\{\max_{0\leq i\leq n} Y(t_i)\geq\lambda\right\}\leq\frac{\mathbf{E}\{Y(T)\}}{\lambda}\quad\text{for }\lambda>0$$

for every partition $0=t_0<t_1<\ldots<t_n=T$ of $[0,T]$.

Exercise 28. Show that Corollary 3.2 follows from Theorem 3.1.

Exercise 29. Find a (non-rightcontinuous) martingale that does not obey the Doob-Kolmogorov inequality.

3.2 Augmented filtrations

Definition 3.3. *A filtration $\{\mathcal{F}_t\}_{t\in T}$ on a complete probability space $(\Omega,\mathcal{F},\mathbf{P})$ is called* augmented *if each member $\mathcal{F}_t$ of the filtration contains all $\mathbf{P}$-null-sets of $\mathcal{F}$.*

Exercise 30. Show that every filtration can be enlarged to a unique smallest possible augmented filtration. In what sense is the latter filtration the "smallest possible"?

Exercise 31. Show that a martingale/submartingale/supermartingale with respect to a certain filtration is a martingale/submartingale/supermartingale also with respect to the smallest possible enlarged augmented version of the filtration.

Definition 3.4. *Let $\{B(t)\}_{t\geq0}$ be a BM that is adapted to an augmented filtration $\{\mathcal{F}_t\}_{t\geq0}$ on a complete probability space, and that is such that $B(t)-B(s)$ is independent of $\mathcal{F}_s$ for $0\leq s<t$. This we call* the usual setup *or* the usual conditions.

Exercise 32. Let B be BM and X a random variable that is independent of B defined on a common complete probability space $(\Omega,\mathcal{F},\mathbf{P})$ with $\mathbf{P}$-null sets $\mathcal{N}$. Show that the ususal conditions holds for $\mathcal{F}_t=\sigma(X)\vee\mathcal{N}\vee\sigma(B(s):s\leq t)$.

Exercise 33. Prove that $\mathcal{F}_0$ is independent of B under the usual conditions.

3.3 Measurable processes

Definition 3.5. *A stochastic process $\{X(t)\}_{t\in T}$, $T\subseteq[0,\infty)$, is* measurable *if the map $X:\Omega\times T\to\mathbb{R}$ is measurable.*

Exercise 34. What does the second "measurable" mean in Definition 3.5?

Theorem 3.6. *A right-continuous/left-continuous process $\{X(t)\}_{t\geq 0}/\{X(t)\}_{t\in[0,T]}$ is measurable.*

Proof. For $\{X(t)\}_{t\geq 0}$ right-continuous, we have

$$X(\omega;t)\leftarrow\sum_{k=1}^{n^2}X\Big(\omega;\frac{k-1}{n}\Big)\,I_{[(k-1)/n,k/n)}(t)\quad\text{as } n\to\infty.$$

Here each term in the sum on the right-hand side measurable, since

$$\Big\{(\omega,t)\in\Omega\times[0,\infty):X\Big(\omega;\frac{k-1}{n}\Big)\,I_{[(k-1)/n,k/n)}(t)\leq x\Big\}$$
$$=\bigg(\Big\{\omega\in\Omega:X\Big(\omega;\frac{k-1}{n}\Big)\leq x\Big\}\bigcap\Big[\frac{k-1}{n},\frac{k}{n}\Big)\bigg)\bigcup\bigg(\{\omega\in\Omega:0\leq x\}\bigcap\Big[\frac{k-1}{n},\frac{k}{n}\Big)^c\bigg).$$

Hence the sum is measurable for each n, as is then the limit as $n\to\infty$. $\square$

Exercise 35. Can you give an example of a non-measurable stochastic process?

Theorem 3.7. *If $\{X(t)\}_{t\in T}$, $T\subseteq[0,\infty)$, is a measurable stochastic process and τ a T-valued random variable, then $X(\tau)$ is a random variable.*

Proof. Recall that compositions of measurable functions are measurable, and note that $X(\tau(\omega))=X(\omega,\tau(\omega))=(X\circ h)(\omega)$, where

$$\Omega\ni\omega\curvearrowright h(\omega)=(\omega,\tau(\omega))\in\Omega\times T.$$

Hence it is sufficient to show that h is measurable. This in turn is so because

$$\{\omega\in\Omega:h(\omega)\in A\times B\}=A\cap\{\omega\in\Omega:\tau(\omega)\in B\}\in\mathcal{F}\quad\text{for } A\in\mathcal{F}\text{ and } B\in\mathcal{B}(T),$$

where $\mathcal{B}(T)$ denotes the Borel-sets of T. As the family $\{C\in\mathcal{F}\times\mathcal{B}(T):h^{-1}(C)\in\mathcal{F}\}$ is a σ-algebra, we conclude that this σ-algebra must be $\mathcal{F}\times\mathcal{B}(T)$, see Exercise 36 below. And so h is measurable, as required. $\square$

Exercise 36. Explain the details of the last part of the proof of Theorem 3.7.

Theorem 3.8. *A measurable stochastic process $\{X(t)\}_{t\geq 0}$ has measurable sample paths, that is, the map $[0,\infty)\ni t\curvearrowright X(\omega,t)\in\mathbb{R}$ is measurable for each $\omega\in\Omega$.*

Lemma 3.9. *Let $(\mathfrak{G},\mathcal{G})$ and $(\mathfrak{H},\mathcal{H})$ be measurable spaces. For the so called* sections *$E^{(\cdot)}$ and $E_{(\cdot)}$ of a set E in the product σ-algebra $\mathcal{G}\times\mathcal{H}$, we have*

$$\begin{cases} E^y=\{x\in\mathfrak{G}:(x,y)\in E\}\in\mathcal{G} & \textit{for } y\in\mathfrak{H} \\ E_x=\{y\in\mathfrak{H}:(x,y)\in E\}\in\mathcal{H} & \textit{for } x\in\mathfrak{G} \end{cases}.$$

**Proof.* It is enough to prove the statement for E^y, see *Execise 37 below. Note that

$$\big\{E\in\mathcal{G}\times\mathcal{H}:E^y\in\mathcal{G}\ \text{ for all }\ y\in\mathfrak{H}\big\} \tag{3.4}$$

is a σ-algebra, see *Execise 37 below. Moreover, we have

$$(G\times H)^y=\begin{cases} G & \text{if } y\in H \\ \emptyset & \text{if } y\notin H \end{cases}\in\mathcal{G}\quad\text{for } y\in\mathfrak{H},\quad \text{for } G\times H\in\mathcal{G}\times\mathcal{H}. \tag{3.5}$$

This proves the lemma, see *Execise 37 below. $\square$

***Exercise 37.** Why is it enough to prove Lemma 3.9 for E^y? Why is the family (3.4) a σ-algebra? Why does Lemma 3.9 follow from (3.4) together with (3.5).

Proof of Theorem 3.8. Pick an $\omega\in\Omega$ and a $C\in\mathcal{B}(\mathbb{R})$. By Lemma 3.9, we have

$$\big\{t\in[0,\infty):X(\omega,t)\in C\big\}=\big\{(\omega,t)\in\Omega\times[0,\infty):X(\omega,t)\in C\big\}_\omega\in\mathcal{B}([0,\infty)),$$

since $\{(\omega,t):X(\omega,t)\in C\}\in\mathcal{F}\times\mathcal{B}([0,\infty))$ by the measurability of X. $\square$

*3.4 Progressively measurable processes

We have selected not to use the concept of progressive measurability (see Definition 3.10 below) in these notes as it introduces additional unproven building elements, and as it is only one proof – that of Theorem 4.4 below – that could have been significantly simplified using progressive measurability. However, as many authors use progressive measurability, let us say just a little about it.

***Definition 3.10.** *A stochastic process $\{X(t)\}_{t\in[0,T]}$ is called* progressively measurable *with respect to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ if the map $X:(\Omega,\mathcal{F}_t)\times([0,t],\mathcal{B}([0,t]))\to(\mathbb{R},\mathcal{B}(\mathbb{R}))$ is measurable for each $t\in[0,T]$.*

It is not too hard to see that a progressively measurable process is measurable and adapted, see ⋆Exercise 38 below. It might seem clear that the converse that a measurable and adapted process is progressively measurable should hold as well, but this turns out to not necessarily be the case. However, the weaker statement that a measurable and adapted process has a progressively measurable version is an important theorem famed for its difficult proof.

To see what progressive measurability can be used for, let us assume the usual conditions and that $\{X(t)\}_{t\in[0,T]}$ is a measurable and adapted process such that $\int_0^T |X(s)|\,ds < \infty$ (almost surely), so that the process $\{\int_0^t X(s)\,ds\}_{t\in[0,T]}$ is well-defined (with probability 1). It might seem immediate that this process is adapted, but the proof of this fact is difficult without using progressive measurability. However, if we pick a progressively measurable version $\{Y(t)\}_{t\in[0,T]}$ of X (supported by the above mentioned difficult to prove theorem), then Fubini's theorem shows that

$$\begin{aligned}
&\mathbf{E}\Big\{\int_0^t |X(s)-Y(s)|\,ds\Big\} \\
&= \int_0^t \mathbf{E}\big\{I_{\{X(s)>Y(s)\}}\,(X(s)-Y(s))\big\}\,ds + \int_0^t \mathbf{E}\big\{I_{\{Y(s)>X(s)\}}\,(Y(s)-X(s))\big\}\,ds \\
&= 0.
\end{aligned}$$

Hence we have $\int_0^t X(s)\,ds = \int_0^t Y(s)\,ds$ with probability 1, and as $\{\int_0^t Y(s)\,ds\}_{t\in[0,T]}$ is adapted (as an immediate consequence of progressive measurability and Fubini's theorem), it follows that $\{\int_0^t X(s)\,ds\}_{t\in[0,T]}$ is adapted (as the filtration is augmented).

⋆Exercise 38. Show that a progressively measurable process $\{X(t)\}_{t\in[0,T]}$ is measurable and adapted.

3.5 Introduction to stochastic integration

A substantial part of our work in this course will be devoted to construct stochastic integrals. Here comes the first few steps on that journey:

Definition 3.11. *A stochastic process $\{X(t)\}_{t\in[0,T]}$ belongs to the class S_T of* simple processes *on $[0,T]$, if for some constants $0 = t_0 < t_1 < \ldots < t_n = T$ and for some random variables $X(0), X_{t_0}, \ldots, X_{t_{n-1}}$ that are adapted to $\mathcal{F}_0, \mathcal{F}_{t_0}, \ldots, \mathcal{F}_{t_{n-1}}$, respectively, and that satisfy $|X(0)|, |X_{t_0}|, \ldots, |X_{t_{n-1}}| \leq C$ for some (non-random) constant $C > 0$, it holds that*

$$X(t) = X(0)I_{\{0\}}(t) + \sum_{i=1}^{n} X_{t_{i-1}}\, I_{(t_{i-1},t_i]}(t) \quad \textit{for } t\in[0,T]. \tag{3.6}$$

Definition 3.12. *A stochastic process* $\{X(t)\}_{t\in[0,T]}$ *belongs to the class* E_T*, if* X *is measurable and adapted to* $\{\mathcal{F}_t\}_{t\geq 0}$ *with*

$$\mathbf{E}\bigg\{\int_0^T X(r)^2\,dr\bigg\} < \infty.$$

Definition 3.13. *A stochastic process* $\{X(t)\}_{t\in[0,T]}$ *belongs to the class* P_T *of* predictable processes on $[0,T]$*, if* X *is measurable and adapted to* $\{\mathcal{F}_t\}_{t\geq 0}$ *with*

$$\mathbf{P}\bigg\{\int_0^T X(r)^2\,dr < \infty\bigg\} = 1.$$

We will define the so called *Itô integral* $\int_0^t X\,dB$, $t\in[0,T]$, of processes $X\in P_T$ with respect to BM B. This will be done by means of first defining the integral for $X\in S_T$, and then extend that integral, first to $X\in E_T$, and then finally to $X\in P_T$.

Exercise 39. Prove that $S_T\subseteq E_T\subseteq P_T$. In particular, explain why processes in S_T are measurable and adapted, and why the Lebesgue integral $\int_0^T X(r)^2\,dr$ is a well-defined random variable for $X\in P_T$.

Example 3.14. As BM B does not have finite variation, we cannot hope to use the usual Stieltjes integral theory to define $\int_0^T X\,dB$. Indeed, as B is continuous, if $\int_0^T B\,dB$ were well-defined as a Stieltjes integral, then we would have (see also ⋆Example 3.15 below)

$$\begin{aligned}\int_0^T B\,dB &= \lim_{\max_{1\leq i\leq n} t_i-t_{i-1}\downarrow 0}\sum_{i=1}^n B(t_{i-1})\,(B(t_i)-B(t_{i-1}))\\ &= \lim_{\max_{1\leq i\leq n} t_i-t_{i-1}\downarrow 0}\sum_{i=1}^n B(t_i)\,(B(t_i)-B(t_{i-1})),\end{aligned}$$

where $0=t_0<t_1<\ldots<t_n=T$ are finer and finer partitions of $[0,T]$. However, by Theorem 1.10 the two limits above are not equal, as their difference

$$\lim_{\max_{1\leq i\leq n} t_i-t_{i-1}\downarrow 0}\bigg(\sum_{i=1}^n B(t_i)\,(B(t_i)-B(t_{i-1})-\sum_{i=1}^n B(t_{i-1})\,(B(t_i)-B(t_{i-1}))\bigg) = T.$$

Exercise 40. Prove that a function has finite variation over an interval if and only if it can be written as the difference between two non-decreasing functions. (It is thus integrals with respect to finite variation functions that have a well-defined signed Stieltjes' signed integration theory, and BM does not fit into this context by Theorem 1.10. This explains the things that happen in Example 3.14.)

***Example 3.15.** (Riemann-Stieltjes integral) If $f:[0,T]\to\mathbb{R}$ is continuous and $g:[0,T]\to\mathbb{R}$ has finite variation, then the Riemann-Stieltjes integral

$$\int_0^T f\,dg = \lim_{\max_{1\le i\le n} t_i - t_{i-1}\downarrow 0} \sum_{i=1}^n f(s_i)\,(g(t_i)-g(t_{i-1})) \tag{3.7}$$

is well-defined, where $0=t_0<t_1<\ldots<t_n=T$ are finer and finer partitions of $[0,T]$ and $s_i\in[t_{i-1},t_i]$. To see this we note that if we pick two different such grids in $[0,T]$, $0=t_0\le s_1\le t_1\le\ldots\le s_m\le t_m=T$ and $0=t_0'\le s_1'\le t_1'\le\ldots\le s_n'\le t_n'=T$, and if $0=t_0<t_1''<\ldots<t_{m+n-1}''=T$ is a refinement of the first two grids that contains all members of them, then we readily get

$$\begin{aligned}
&\left|\sum_{i=1}^m f(s_i)\,(g(t_i)-g(t_{i-1})) - \sum_{i=1}^n f(s_i')\,(g(t_i')-g(t_{i-1}'))\right| \\
&\le \Bigg(\sup_{s,t\in[0,T]\,:\,|s-t|\le \max_{1\le i\le m} t_i-t_{i-1}} |f(s)-f(t)| + \sup_{s,t\in[0,T]\,:\,|s-t|\le \max_{1\le i\le n} t_i'-t_{i-1}'} |f(s)-f(t)|\Bigg) \\
&\qquad\times \sum_{i=1}^{m+n-1} |g(t_i'')-g(t_{i-1}'')|
\end{aligned} \tag{3.7}$$

$\to 0$ as $\max_{1\le i\le m} t_i-t_{i-1}\downarrow 0$ and $\max_{1\le i\le n} t_i'-t_{i-1}'\downarrow 0$,

by uniform continuity of f and finite variation of g. From this we see that

$$\left\{\sum_{i=1}^n f(t_{i-1})\,(g(t_i)-g(t_{i-1}))\right\}$$

is a Cauchy sequence when $0=t_0<t_1<\ldots<t_n=T$ satisfy $\max_{1\le i\le m} t_i-t_{i-1}\downarrow 0$. This Cauchy sequence converges to some limit $\int_0^T f\,dg$ as $\max_{1\le i\le m} t_i-t_{i-1}\downarrow 0$. Moreover, by another application of (3.7), it follows that any sum of the type (3.7) must converge to that same limit $\int_0^T f\,dg$ as $\max_{1\le i\le m} t_i-t_{i-1}\downarrow 0$.

4 Lecture 4, Monday March 30

Henceforth we always assume the usual conditions!

4.1 Itô integrals for the space S_T

Note that a simple process X in S_T

$$X(t) = X(0)I_{\{0\}}(t) + \sum_{i=1}^{n} X_{t_{i-1}} I_{(t_{i-1},t_i]}(t) \quad \text{for } t \in [0,T],$$

$0 = t_0 < t_1 < \ldots < t_n = T$, is not the same thing as a simple function in Lebesgue integration, but rather a step functions of the kind employed in Riemann integration.

Definition 4.1. *The* Itô integral process $\{\int_0^t X\,dB\}_{t\in[0,T]}$ *of a simple process* $X \in S_T$ *is defined by* $\int_0^0 X\,dB = 0$ *and*

$$\int_0^t X\,dB = \sum_{i=1}^{m} X_{t_{i-1}} (B(t_i) - B(t_{i-1})) + X_{t_m} (B(t) - B(t_m)) \quad \textit{for } t \in (t_m, t_{m+1}],$$

for $m = 0, \ldots, n-1$. *Further, we define*

$$\int_s^t X\,dB = \int_0^t X\,dB - \int_0^s X\,dB \quad \textit{for } s,t \in [0,T].$$

Exercise 41. (Consistency) Let an $X \in S_T$ have two representations

$$X(t) = X(0)I_{\{0\}}(t) + \sum_{i=1}^{m} X_{t_{i-1}} I_{(t_{i-1},t_i]}(t) = X(0)I_{\{0\}}(t) + \sum_{i=1}^{n} X'_{t'_{i-1}} I_{(t'_{i-1},t'_i]}(t)$$

for $t \in [0,T]$. By means of introducing a third grid that contains all times of the grids $0 = t_0 < t_1 < \ldots < t_m = T$ and $0 = t'_0 < t'_1 < \ldots < t'_n = T$, show that the Itô integral process $\{\int_0^t X\,dB\}_{t\in[0,T]}$ coincides for the two representations of X.

Exercise 42. (Continuity) Show that $\{\int_0^t X\,dB\}_{t\in[0,T]}$ is a continuous stochastic process for $X \in S_T$.

Exercise 43. (Nada) Show that for an $X \in S_T$ we have

$$\int_r^t X\,dB = -\int_t^r X\,dB \quad \text{and} \quad \int_0^t I_{[r,s]} X\,dB = \int_r^s X\,dB \quad \text{for } 0 \le r \le s \le t \le T.$$

Exercise 44. (Linearity) Show that for $X, Y \in S_T$ (not necessarily with the same grids) and constants $a, b \in \mathbb{R}$, we have

$$\int_0^t (aX + bY)\,dB = a\int_0^t X\,dB + b\int_0^t Y\,dB \quad \text{for } t \in [0,T].$$

Exercise 45. (ADAPTEDNESS) Show that $\{\int_0^t X\,dB\}_{t\in[0,T]}$ is an adapted stochastic process for $X \in S_T$.

Theorem 4.2. (MARTINGALE) *For $X \in S_T$, $\{\int_0^t X\,dB\}_{t\in[0,T]}$ is a martingale.*

Proof. By Exercise 45, $\int_0^t X\,dB$ is adapted. Further, we have $\mathbf{E}\{|\int_0^t X\,dB|\} < \infty$ by inspection of the definition of $\int_0^t X\,dB$ together with the boundedness of the X process. To show that $\mathbf{E}\{\int_0^t X\,dB\,|\,\mathcal{F}_s\} = \int_0^s X\,dB$ for $0 \le s < t \le T$ we can assume that $s = t_j < t_{m+1} = t$ for some members t_j and t_{m+1} of the grid $0 = t_0 < t_1 < \ldots < t_n = T$ (cf. Definition 3.11), as otherwise the grid can be enriched to that end without affecting the values of the process X or its Itô integral process, see Exercise 41. Now, as $B(t_i) - B(t_{i-1})$ is independent of $\mathcal{F}_{t_{i-1}}$, and all BM-values $\{B(t)\}_{t\le t_{i-1}}$ as well as $X_{t_{i-1}}$ are $\mathcal{F}_{t_{i-1}}$-measurable, we have (by the basic rules for conditional expectations)

$$
\begin{aligned}
&\mathbf{E}\bigg\{\int_0^t X\,dB\,\bigg|\,\mathcal{F}_s\bigg\}\\
&= \mathbf{E}\bigg\{\sum_{i=1}^{j} X_{t_{i-1}}(B(t_i)-B(t_{i-1}))\,\bigg|\,\mathcal{F}_s\bigg\} + \mathbf{E}\bigg\{\sum_{i=j+1}^{m+1} X_{t_{i-1}}(B(t_i)-B(t_{i-1}))\,\bigg|\,\mathcal{F}_s\bigg\}\\
&= \sum_{i=1}^{j} X_{t_{i-1}}(B(t_i))-B(t_{i-1})) + \sum_{i=j+1}^{m+1}\mathbf{E}\Big\{\mathbf{E}\big\{X_{t_{i-1}}(B(t_i)-B(t_{i-1}))\big|\mathcal{F}_{t_{i-1}}\big\}\,\Big|\,\mathcal{F}_s\Big\} \qquad (4.1)\\
&= \int_0^s X\,dB + \sum_{i=j+1}^{m+1}\mathbf{E}\Big\{X_{t_{i-1}}\,\mathbf{E}\big\{B(t_i)-B(t_{i-1})\big|\mathcal{F}_{t_{i-1}}\big\}\,\Big|\,\mathcal{F}_s\Big\}\\
&= \int_0^s X\,dB + 0.
\end{aligned}
$$

□

Exercise 46. (ZERO-MEAN) Show that $\mathbf{E}\{\int_0^t X\,dB\} = 0$ for $X \in S_T$ and $t \in [0,T]$.

Theorem 4.3. (ISOMETRY) *For $X, Y \in S_T$ and $t \in [0,T]$, we have*

$$\mathbf{E}\bigg\{\bigg(\int_0^t X\,dB\bigg)\bigg(\int_0^t Y\,dB\bigg)\bigg\} = \int_0^t \mathbf{E}\{X(r)Y(r)\}\,dr = \mathbf{E}\bigg\{\int_0^t X(r)Y(r)\,dr\bigg\}.$$

Proof. We can without loss assume that X and Y have a common grid $0 = t_0 < t_1 < \ldots < t_n = T$ and that $t = t_{m+1}$ belongs to that grid (see Exercise 41 and the proof of Theorem 4.2). Now, since $X_{t_{i-1}}Y_{t_{j-1}}(B(t_i)-B(t_{i-1}))$ is $\mathcal{F}_{t_{j-1}}$-measurable and independent of $B(t_j) - B(t_{j-1})$ for $i < j$, we have

$$\mathbf{E}\bigg\{\bigg(\int_0^t X\,dB\bigg)\bigg(\int_0^t Y\,dB\bigg)\bigg\}$$

$$
\begin{aligned}
&= \mathbf{E}\Big\{\Big(\sum_{i=1}^{m+1} X_{t_{i-1}}(B(t_i)-B(t_{i-1}))\Big)\Big(\sum_{j=1}^{m+1} Y_{t_{j-1}}(B(t_j)-B(t_{j-1}))\Big)\Big\} \\
&= \mathbf{E}\Big\{\sum_{i=1}^{m+1} X_{t_{i-1}}Y_{t_{i-1}}(B(t_i)-B(t_{i-1}))^2\Big\} \\
&\qquad + \mathbf{E}\Big\{\sum_{1\le i,j\le m+1,\, i\neq j} X_{t_{i-1}}Y_{t_{j-1}}(B(t_i)-B(t_{i-1}))\,(B(t_j)-B(t_{j-1}))\Big\} \\
&= \sum_{i=1}^{m+1} \mathbf{E}\{X_{t_{i-1}}Y_{t_{i-1}}\}\,\mathbf{E}\{(B(t_i)-B(t_{i-1}))^2\} \\
&\qquad + \sum_{1\le i<j\le m+1} \mathbf{E}\{B(t_j)-B(t_{j-1})\}\,\mathbf{E}\{X_{t_{i-1}}Y_{t_{j-1}}(B(t_i)-B(t_{i-1}))\} \\
&\qquad + \sum_{1\le j<i\le m+1} \mathbf{E}\{B(t_i)-B(t_{i-1})\}\,\mathbf{E}\{X_{t_{i-1}}Y_{t_{j-1}}(B(t_j)-B(t_{j-1}))\} \\
&= \sum_{i=1}^{m+1} \mathbf{E}\{X_{t_{i-1}}Y_{t_{i-1}}\}\,(t_i-t_{i-1}) + 0 + 0 \\
&= \int_0^t \mathbf{E}\{X(r)Y(r)\}\,dr \\
&= \mathbf{E}\Big\{\int_0^t X(r)Y(r)\,dr\Big\},
\end{aligned}
$$

using Fubini's theorem in the last step (recall that $X, Y \in S_T$ are measurable by Exercise 39). Note that all expectations are finite by boundedness of $X, Y \in S_T$. □

4.2 Itô integrals for the space E_T

The next Theorem 4.4 is crucial to construct the Itô integrals for the space E_T. We will return to the rather difficult proof of this theorem later.

Theorem 4.4. *For $X \in E_T$ there exists a sequence $\{X_n\}_{n=1}^{\infty} \subseteq S_T$ such that*

$$\lim_{n\to\infty} \mathbf{E}\Big\{\int_0^T \big(X_n(r)-X(r)\big)^2\,dr\Big\} = 0. \tag{4.2}$$

Definition and Theorem 4.5. *The* Itô integral process $\{\int_0^t X\,dB\}_{t\in[0,T]}$ *for $X \in E_T$ is well-defined as the unique up to version and continuous with probability 1 stochastic process that is given as the limit in the sense of convergence in mean-square by*

$$\int_0^t X\,dB \leftarrow \int_0^t X_n\,dB \quad \text{as } n\to\infty \text{ for } t\in[0,T], \tag{4.3}$$

where $\{X_n\}_{n=1}^{\infty} \subseteq S_T$ satisfies (4.2). Further, we define

$$\int_s^t X\,dB = \int_0^t X\,dB - \int_0^s X\,dB \quad \text{for } s,t\in[0,T].$$

Proof. The mean-square limit (4.3) exists as $\{\int_0^t X_n\,dB\}_{n=1}^\infty$ is a mean-square Cauchy sequence. This is so since by linearity and isometry for the integral on S_T together with the elementary inequality $(x+y)^2 \le 2\,x^2 + 2\,y^2$ for $x,y\in\mathbb{R}$ and (4.2), we have

$$\begin{aligned}
&\mathbf{E}\bigg\{\bigg(\int_0^t X_m\,dB - \int_0^t X_n\,dB\bigg)^2\bigg\}\\
&\quad= \mathbf{E}\bigg\{\int_0^t (X_m(r)-X_n(r))^2\,dr\bigg\}\\
&\quad\le 2\,\mathbf{E}\bigg\{\int_0^T \big(X_m(r)-X(r)\big)^2\,dr\bigg\} + 2\,\mathbf{E}\bigg\{\int_0^T \big(X(r)-X_n(r)\big)^2\,dr\bigg\}\\
&\quad\to 0 \quad \text{as } m,n\to\infty.
\end{aligned} \tag{4.4}$$

Further, in the sense of uniqueness up to version, the limit in (4.3) does not depend on which sequence $\{X_n\}_{n=1}^\infty$ we choose that satisfies (4.2), because if $\{X_n'\}_{n=1}^\infty \subseteq S_T$ also satisfies (4.2) and converges to the limit $\oint_0^t X\,dB$ in (4.3), then the fact that mean-square limits and second moments commute gives by inspection of (4.4)

$$\begin{aligned}
&\mathbf{E}\bigg\{\bigg(\int_0^t X\,dB - \oint_0^t X\,dB\bigg)^2\bigg\}\\
&= \lim_{n\to\infty} \mathbf{E}\bigg\{\bigg(\int_0^t X_n\,dB - \int_0^t X_n'\,dB\bigg)^2\bigg\}\\
&\le 2\limsup_{n\to\infty}\mathbf{E}\bigg\{\int_0^T \big(X_n(r)-X(r)\big)^2\,dr\bigg\} + 2\limsup_{n\to\infty}\mathbf{E}\bigg\{\int_0^T \big(X(r)-X_n'(r)\big)^2\,dr\bigg\}\\
&= 0.
\end{aligned}$$

As for the existence of a continuous version of the limit process in (4.3), use Theorem 4.4 to pick a sequence $\{X_n\}_{n=1}^\infty \subseteq S_T$ such that

$$\mathbf{E}\bigg\{\int_0^T \big(X(r)-X_n(r)\big)^2\,dr\bigg\} \le \frac{2^{-n}}{3} \quad \text{for } n\in\mathbb{N}.$$

From this together with the inequality $(x+y)^2 \le 2\,x^2 + 2\,y^2$ we get

$$\begin{aligned}
&\mathbf{E}\bigg\{\int_0^T \big(X_{n+1}(r)-X_n(r)\big)^2\,dr\bigg\}\\
&\le 2\,\mathbf{E}\bigg\{\int_0^T \big(X_{n+1}(r)-X(r)\big)^2\,dr\bigg\} + 2\,\mathbf{E}\bigg\{\int_0^T \big(X(r)-X_n(r)\big)^2\,dr\bigg\}\\
&\le 2^{-n} \quad \text{for } n\in\mathbb{N}.
\end{aligned} \tag{4.5}$$

By Definition and Theorem 4.5 together with a telescope sum argument, we have

$$\int_0^t X_1\,dB + \sum_{n=1}^N \int_0^t (X_{n+1}-X_n)\,dB \to \int_0^t X\,dB \quad \text{in mean-square as } N\to\infty.$$

Since each term in the sum on the left-hand side is continuous (cf. Exercise 42), we get the claimed existence of a continuous version of $\int_0^t X\,dB$ if the sum converges uniformly for $t \in [0,T]$ with probability one. This in turn holds if

$$\mathbf{P}\bigg\{\sum_{n=1}^{\infty} \sup_{t\in[0,T]} \bigg|\int_0^t (X_{n+1}-X_n)\,dB\bigg| \ \text{converges}\bigg\} = 1, \tag{4.6}$$

see Exercise 47 below. To prove (4.6) it is sufficient to show that

$$\mathbf{P}\bigg\{\sup_{t\in[0,T]} \bigg|\int_0^t (X_{n+1}-X_n)\,dB\bigg| > \frac{1}{n^2} \ \text{for infinitely many} \ n\in\mathbb{N}\bigg\} = 0. \tag{4.7}$$

By the Borel-Cantelli Lemma, (4.7) in turn will follow if

$$\sum_{n=1}^{\infty} \mathbf{P}\bigg\{\sup_{t\in[0,T]} \bigg|\int_0^t (X_{n+1}-X_n)\,dB\bigg| > \frac{1}{n^2}\bigg\} < \infty. \tag{4.8}$$

However, since $\int_0^t (X_{n+1}-X_n)\,dB$ is a continuous martingale (by Exercise 42 and Theorem 4.2), the Doob-Kolmogorov inequality (Corollary 3.2) together with isometry for the space S_T and (4.5) show that the sum in (4.8) is at most

$$\begin{aligned}\sum_{n=1}^{\infty} \mathbf{E}\bigg\{\bigg(\int_0^T (X_{n+1}-X_n)\,dB\bigg)^2\bigg\}\bigg/\bigg(\frac{1}{n^2}\bigg)^2 &= \sum_{n=1}^{\infty} \mathbf{E}\bigg\{\int_0^T \big(X_{n+1}(r)-X_n(r)\big)^2\,dr\bigg\}\, n^4 \\ &\leq \sum_{n=1}^{\infty} \frac{n^4}{2^n} \\ &< \infty. \end{aligned}$$

□

Exercise 47. Prove the following fact used in the proofs of *Theorem 2.3 and Definition and Theorem 4.5, that if $\{a_n : [0,T]\to\mathbb{R}\}_{n=1}^{\infty}$ are continuous functions such that $\sum_{n=1}^{\infty} \sup_{t\in[0,T]} |a_n(t)| < \infty$, then $\sum_{n=1}^{\infty} a_n : [0,T]\to\mathbb{R}$ is continuous.

We will verify several properties for Itô integral processes below. Many of those verifications consist of checking that two continuous with probability 1 stochastic processes $\{Y_1(t)\}_{t\in[0,T]}$ and $\{Y_2(t)\}_{t\in[0,T]}$ are versions of each other, that is, they agree with probability 1 for each $t\in[0,T]$. (Itô integral process values will be defined in the sense of convergence in mean-square or convegence in probability, so that two Itô integral process values agree means that they are equal with probability 1.) However, as the processes Y_1 and Y_2 are continuous with probability 1, they are versions of each other if and only if $\mathbf{P}\{Y_1(t) = Y_2(t) \ \text{for all} \ t\in[0,T]\} = 1$, see Exercise 48 below. So when we verify several properties for Itô integral processes in the sense that the property holds with probability 1 for each $t\in[0,T]$ below, it does in fact follow that the property holds simultaneuously for all $t\in[0,T]$.

Exercise 48. Show that two continuous with probability 1 processes $\{Y_1(t)\}_{t\in[0,T]}$ and $\{Y_2(t)\}_{t\in[0,T]}$ defined on a common complete probability space are versions of each other if and only if it holds that $\mathbf{P}\{Y_1(t) = Y_2(t) \ \text{for all} \ t\in[0,T]\} = 1$.

Exercise 49. (Nada) Show that for an $X\in E_T$ we have
$$\int_r^t X\,dB=-\int_t^r X\,dB \quad\text{and}\quad \int_0^t I_{[r,s]}\,X\,dB=\int_r^s X\,dB \quad\text{for } 0\le r\le s\le t\le T.$$

Exercise 50. (Linearity) Show that for $X,Y\in E_T$ and constants $a,b\in\mathbb{R}$, we have
$$\int_0^t (aX+bY)\,dB=a\int_0^t X\,dB+b\int_0^t Y\,dB \quad\text{for } t\in[0,T].$$

Exercise 51. (Adaptedness) Show that $\{\int_0^t X\,dB\}_{t\in[0,T]}$ is an adapted stochastic process for $X\in E_T$.

Exercise 52. (Martingale) Using Exercise 53 below, show that $\{\int_0^t X\,dB\}_{t\in[0,T]}$ is a square-integrable martingale for $X\in E_T$.

Exercise 53. Let $\mathcal{G}$ be a σ-algebra (contained in $\mathcal{F}$) and let $Z,Z_1,Z_2,\ldots$ be random variables. Show that as $n\to\infty$ we have
$$Z_n\to Z \text{ in } \mathbb{L}^1 \text{ (/mean-square)} \;\Rightarrow\; \mathbf{E}\{Z_n|\mathcal{G}\}\to\mathbf{E}\{Z|\mathcal{G}\} \text{ in } \mathbb{L}^1 \text{ (/mean-square)}.$$

Exercise 54. (Zero-mean) Show that $\mathbf{E}\{\int_0^t X\,dB\}=0$ for $X\in E_T$ and $t\in[0,T]$.

Exercise 55. (Isometry) Show that for $X,Y\in E_T$ and $t\in[0,T]$, we have
$$\mathbf{E}\bigg\{\Big(\int_0^t X\,dB\Big)\Big(\int_0^t Y\,dB\Big)\bigg\}=\int_0^t \mathbf{E}\{X(r)Y(r)\}\,dr=\mathbf{E}\bigg\{\int_0^t X(r)Y(r)\,dr\bigg\}.$$

Exercise 56. (Convergence) Show that if $X\in E_T$ and $\{X_n\}_{n=1}^\infty\subseteq E_T$ satisfy
$$\lim_{n\to\infty}\mathbf{E}\bigg\{\int_0^T \big(X_n(r)-X(r)\big)^2\,dr\bigg\}=0,$$
then we have in the sense of convergence in mean-square
$$\int_0^t X_n\,dB\to\int_0^t X\,dB \quad\text{as } n\to\infty \text{ for } t\in[0,T].$$

Exercise 57. (Zipp) Show that $\int_0^t I_{[s,t]}(r)\,Y X(r)\,dB(r)=Y\int_s^t X(r)\,dB(r)$ for $0\le s\le t\le T$ and $X\in E_T$, when Y is an $\mathcal{F}_s$-measurable random variable that is bounded by a (non-random) constant.

Exercise 58. Let $f:[0,T]\to\mathbb{R}$ be a non-random function belonging to E_T, which is to say that $f\in\mathbb{L}^2([0,T])$. Show that $\{\int_0^t f\,dB\}_{t\in[0,T]}$ is a zero mean Gaussian process with covariance function $\mathbf{Cov}\{\int_0^s f\,dB,\int_0^t f\,dB\}=\int_0^{\min\{s,t\}} f(r)^2\,dr$.

5 Lecture 5, Wednesday April 1

5.1 Proof of Theorem 4.4.

Given an $X \in E_T$ and a constant $\varepsilon > 0$, we need to prove that

$$\mathbf{E}\bigg\{\int_0^T (Y(r)-X(r))^2\,dr\bigg\} \leq \varepsilon \quad \text{for some } Y \in S_T.$$

To that end truncate X as

$$X^{(N)}(r) = \begin{cases} -N & \text{if} \quad X(r) < -N \\ X(t) & \text{if} \quad |X(r)| \leq N \\ N & \text{if} \quad X(r) > N \end{cases}. \tag{5.1}$$

Since $X^{(N)}(r) \to X(r)$ as $N \to \infty$ with $(X^{(N)}(r)-X(r))^2 \leq X(r)^2$, we then have

$$\mathbf{E}\bigg\{\int_0^T \big(X^{(N)}(r)-X(r)\big)^2\,dr\bigg\} \to 0 \quad \text{as } N \to \infty \tag{5.2}$$

(by dominated convergence as $X \in E_T$). Using the inequality $(x+y)^2 \leq 2\,x^2 + 2\,y^2$ it follows that it is enough to prove that, given $X \in E_T$, $\varepsilon > 0$ and $N \in \mathbb{N}$, we have

$$\mathbf{E}\bigg\{\int_0^T \big(Y(r)-X^{(N)}(r)\big)^2\,dr\bigg\} \leq \varepsilon \quad \text{for some } Y \in S_T. \tag{5.3}$$

That this is so in the particular case when X is continuous is Exercise 59 below.

*In the general case of an $X \in E_T$ that is not necessarily continuous, the proof of (5.3) is much more difficult than when X is continuous. It goes like this: Given constants $\tau \in [0,1]$ and $n \in \mathbb{N}$, define a discrete approximation $\{X_n^{(N,\tau)}(t)\}_{t\in[0,T]}$ of $X^{(N)}$ as

$$X_n^{(N,\tau)}(t) = \sum_{k=0}^{\infty} I_{(2^{-n}(k+\tau),2^{-n}(k+\tau+1)]\,\cap\,[0,T]}(t)X^{(N)}(2^{-n}(k+\tau)) \quad \text{for } t \in [0,T].$$

As it is immediate that $X_n^{(N,\tau)} \in S_T$, the sufficiency criterion (5.3) holds if

$$\liminf_{n\to\infty}\mathbf{E}\bigg\{\int_0^T \big(X_n^{(N,\tau)}(r)-X^{(N)}(r)\big)^2\,dr\bigg\} = 0 \quad \text{for some } \tau \in [0,1]. \tag{5.4}$$

Now, if (5.4) does not hold, then Fubini's theorem and Fatou's lemma show that

$$\begin{aligned} &\liminf_{n\to\infty}\mathbf{E}\bigg\{\int_{\tau=0}^{\tau=1}\int_{r=0}^{r=T} \big(X_n^{(N,\tau)}(r)-X^{(N)}(r)\big)^2\,dr\,d\tau\bigg\} \\ &\qquad \geq \int_{\tau=0}^{\tau=1} \liminf_{n\to\infty}\mathbf{E}\bigg\{\int_{r=0}^{r=T} \big(X_n^{(N,\tau)}(r)-X^{(N)}(r)\big)^2\,dr\bigg\}\,d\tau \\ &\qquad > 0. \end{aligned}$$

Hence it is sufficient to prove that

$$\lim_{n\to\infty}\mathbf{E}\bigg\{\int_{\tau=0}^{\tau=1}\int_{r=0}^{r=T} \big(X_n^{(N,\tau)}(r)-X^{(N)}(r)\big)^2\,dr\,d\tau\bigg\} = 0. \tag{5.5}$$

The next step of the proof is to obeserve that it is sufficient to prove

$$\lim_{h\downarrow 0}\mathbf{E}\bigg\{\int_h^T \big(X^{(N)}(r)-X^{(N)}(r-h)\big)^2\,dr\bigg\}=0. \tag{5.6}$$

This is so because if (5.6) holds, then we get (5.5) by elementary calculations as

$$\begin{aligned}
&\mathbf{E}\bigg\{\int_{\tau=0}^{\tau=1}\int_{r=0}^{r=T}\big(X_n^{(N,\tau)}(r)-X^{(N)}(r)\big)^2\,dr d\tau\bigg\}\\
&=\mathbf{E}\bigg\{\int_{\tau=0}^{\tau=1}\sum_{k=0}^{\infty}\int_{r\in(2^{-n}(k+\tau),2^{-n}(k+\tau+1)]\cap[0,T]}\big(X^{(N)}(2^{-n}(k+\tau))-X^{(N)}(r)\big)^2\,dr d\tau\bigg\}\\
&=\sum_{k=0}^{\infty}\mathbf{E}\bigg\{\int_{r=0}^{r=T}\int_{\tau\in[2^n r-(k+1),2^n r-k)\cap[0,1]}\big(X^{(N)}(2^{-n}(k+\tau))-X^{(N)}(r)\big)^2\,d\tau dr\bigg\}\\
&=\sum_{k=0}^{\infty}\mathbf{E}\bigg\{2^n\int_{r=0}^{r=T}\int_{\tilde\tau=-2^{-n}}^{\tilde\tau=0} I_{[2^{-n}k-r,2^{-n}(k+1)-r]}(\tilde\tau)\big(X^{(N)}(r+\tilde\tau)-X^{(N)}(r)\big)^2\,d\tilde\tau dr\bigg\}\\
&=\mathbf{E}\bigg\{2^n\int_{\tilde\tau=-2^{-n}}^{\tilde\tau=0}\int_{r=0}^{r=T}\sum_{k=0}^{\infty} I_{[2^{-n}k-\tilde\tau,2^{-n}(k+1)-\tilde\tau]}(r)\big(X^{(N)}(r+\tilde\tau)-X^{(N)}(r)\big)^2\,dr d\tilde\tau\bigg\}\\
&\le(2N)^2\int_0^{2^{-n}}dr+\mathbf{E}\bigg\{2^n\int_{\tilde\tau=-2^{-n}}^{\tilde\tau=0}\int_{r=2^{-n}}^{r=T}\big(X^{(N)}(r+\tilde\tau)-X^{(N)}(r)\big)^2\,dr d\tilde\tau\bigg\}\\
&\le(2N)^2\int_0^{2^{-n}}dr+\sup_{h\in[0,2^{-n}]}\mathbf{E}\bigg\{\int_h^T\big(X^{(N)}(r-h)-X^{(N)}(r)\big)^2\,dr\bigg\}\\
&\to 0\quad\text{as } n\to\infty.
\end{aligned}$$

To prove (5.6) in turn it is sufficient to prove that, given any $\varepsilon>0$, we have

$$\mathbf{E}\bigg\{\int_0^T\big(X^{(N)}(r)-Z(r)\big)^2\,dr\bigg\}\le\varepsilon \tag{5.7}$$

for some continuous process $\{Z(r)\}_{r\in[0,T]}$ with $\sup_{r\in[0,T]}|Z(r)|\le N$. This is so because then we may use the inequality $(x+y+z)^2\le 3\,x^2+3\,y^2+3\,z^2$ to deduce that

$$\begin{aligned}
\mathbf{E}\bigg\{\int_h^T(X^{(N)}(r)-X^{(N)}(r-h))^2\,dr\bigg\}&\le 3\,\mathbf{E}\bigg\{\int_h^T(X^{(N)}(r)-Z(r))^2\,dr\bigg\}\\
&\quad+3\,\mathbf{E}\bigg\{\int_h^T(Z(r)-Z(r-h))^2\,dr\bigg\}\\
&\quad+3\,\mathbf{E}\bigg\{\int_h^T(Z(r-h)-X^{(N)}(r-h))^2\,dr\bigg\}\\
&\le 3\varepsilon+3\,\mathbf{E}\bigg\{\sup_{r\in[h,T]}(Z(r)-Z(r-h))^2\int_0^T dr\bigg\}+3\varepsilon\\
&\to 6\varepsilon\quad\text{as } h\downarrow 0
\end{aligned}$$

by uniform continuity of Z together with the bounded convergence theorem.

Define $Y(t)=\int_0^t X^{(N)}(s)\,ds$ for $t\in[0,T]$ and

$$Z_n(t)=n\big(Y(t^+\wedge T)-Y((t-1/n)^+\wedge T)\big)=n\int_{(t-1/n)^+\wedge T}^{t^+\wedge T}X^{(N)}(s)\,ds \tag{5.8}$$

for $t \in [0,T]$. Notice that $|Z_n(t)| \le N$ since the integral is over an interval of length at most $1/n$ and $|X^{(N)}(r)| \le N$. Further, Z_n is continuous because Y is absolutely continuous. As absolutely continuous functions are differentiable almost everywhere, it follows that $Y'(t) = \lim_{n\to\infty} Z_n(t) = X^{(N)}(t)$ for almost all $t \in [0,T]$. Since $(X^{(N)}(t) - Z_n(t))^2 \le (2N)^2$, the dominated convergence theorem now gives

$$\lim_{n\to\infty} \mathbf{E}\bigg\{\int_0^T \big(X^{(N)}(r) - Z_n(r)\big)^2\, dr\bigg\} = 0. \tag{5.9}$$

Picking n sufficiently large we conclude that (5.7) holds. □

Exercise 59. Give a direct proof of (5.3) in the particular case when X is continuous [so that $X^{(N)}$ is continuous and bounded].

***Exercise 60.** It is tempting to conclude (5.3) directly from (5.9) together with Exercise 59. Explain why this is not possible (unless we use the fact that every measurable and adapted process has a progressively measurable version).

5.2 Itô integrals for the space P_T

The construction of Itô integrals for the space P_T will require the following two help Theorems 5.1 and 5.2, the proofs of which are given later:

Theorem 5.1. *For $X \in P_T$, we have in the sense of convergence in probability*

$$\int_0^T (X_n(r) - X(r))^2\, dr \to 0 \quad \text{as } n\to\infty \quad \text{for some sequence } \{X_n\}_{n=1}^\infty \subseteq E_T. \tag{5.10}$$

Theorem 5.2. *For $X \in E_T$ and a constant $C > 0$, we have*

$$\mathbf{P}\bigg\{\sup_{t\in[0,T]}\bigg|\int_0^t X\, dB\bigg| > \lambda\bigg\} \le \frac{C}{\lambda^2} + \mathbf{P}\bigg\{\int_0^T X(r)^2\, dr \ge C\bigg\} \quad \text{for } \lambda > 0. \tag{5.11}$$

Definition and Theorem 5.3. *The* Itô integral process $\{\int_0^t X\, dB\}_{t\in[0,T]}$ *for $X \in P_T$ is well-defined as the unique up to version and continuous with probability 1 stochastic process that is given as the limit in the sense of convergence in probability by*

$$\int_0^t X\, dB \leftarrow \int_0^t X_n\, dB \quad \text{as } n\to\infty \text{ for } t\in[0,T], \tag{5.12}$$

where $\{X_n\}_{n=1}^\infty \subseteq E_T$ satisfies (5.10). Further, we define

$$\int_s^t X\, dB = \int_0^t X\, dB - \int_0^s X\, dB \quad \text{for } s,t \in [0,T].$$

Proof. To prove that the limit in probability exists it is sufficient to check the Cauchy criterion for convergence in probability, which is to say that

$$\lim_{m,n\to\infty} \mathbf{P}\bigg\{\bigg|\int_0^t X_m\,dB - \int_0^t X_n\,dB\bigg| > \varepsilon\bigg\} = 0 \quad \text{for each } \varepsilon > 0. \tag{5.13}$$

However, by (5.10) and (5.11), given any $\delta > 0$, the left-hand side of (5.13) is at most

$$\begin{aligned}
&\limsup_{m,n\to\infty} \mathbf{P}\bigg\{\sup_{t\in[0,T]}\bigg|\int_0^t X_m\,dB - \int_0^t X_n\,dB\bigg| > \varepsilon\bigg\} \\
&\leq \frac{\delta\,\varepsilon^2}{\varepsilon^2} + \limsup_{m,n\to\infty} \mathbf{P}\bigg\{\int_0^T (X_m(r)-X_n(r))^2\,dr \geq \delta\,\varepsilon^2\bigg\} \\
&\leq \delta + \limsup_{m,n\to\infty} \mathbf{P}\bigg\{2\int_0^T (X_m(r)-X(r))^2\,dr + 2\int_0^T (X(r)-X_n(r))^2\,dr \geq \delta\,\varepsilon^2\bigg\} \\
&\leq \delta + 2\limsup_{n\to\infty} \mathbf{P}\bigg\{\int_0^T (X_n(r)-X(r))^2\,dr \geq \frac{\delta\,\varepsilon^2}{4}\bigg\} \\
&= \delta
\end{aligned} \tag{5.14}$$

(using Boole's inequality for the last inequality). This shows that (5.13) holds.

In the sense of uniqueness up to version, the limit in (5.12) does not depend on which sequence $\{X_n\}_{n=1}^\infty$ we choose that satisfies (5.10), because if $\{X_n'\}_{n=1}^\infty \subseteq E_T$ also satisfies (5.10) and converges to the limit $\oint_0^t X\,dB$ in (5.12), then the limits $\int_0^t X\,dB$ and $\oint_0^t X\,dB$ must agree. This is so since (5.11) gives [cf. (5.14)]

$$\begin{aligned}
&\mathbf{P}\bigg\{\bigg|\int_0^t X\,dB - \oint_0^t X\,dB\bigg| > \varepsilon\bigg\} \\
&\leq \limsup_{n\to\infty} \mathbf{P}\bigg\{\bigg|\int_0^t X_n\,dB - \int_0^t X_n'\,dB\bigg| > \varepsilon\bigg\} \\
&\leq \delta + \limsup_{n\to\infty} \mathbf{P}\bigg\{\int_0^T (X_n(r)-X_n'(r))^2\,dr \geq \delta\varepsilon^2\bigg\} \\
&\leq \delta + \limsup_{n\to\infty} \bigg[\mathbf{P}\bigg\{\int_0^T (X_n(r)-X(r))^2\,dr \geq \frac{\delta\varepsilon^2}{4}\bigg\} + \mathbf{P}\bigg\{\int_0^T (X_n'(r)-X(r))^2\,dr \geq \frac{\delta\varepsilon^2}{4}\bigg\}\bigg] \\
&= \delta \quad \text{for any choice of } \varepsilon, \delta > 0.
\end{aligned}$$

As for the existence of a continuous version of the limit process in (5.12), use (5.10) to pick a sequence $\{X_n\}_{n=1}^\infty \subseteq E_T$ such that

$$\mathbf{P}\bigg\{\int_0^T \big(X(r)-X_n(r)\big)^2\,dr > \frac{1}{4\,n^6}\bigg\} \leq \frac{2^{-n}}{3} \quad \text{for } n \in \mathbb{N}.$$

From this together with the inequality $(x+y)^2 \leq 2\,x^2 + 2\,y^2$ we get [cf. (4.5)]

$$\begin{aligned}
&\mathbf{P}\bigg\{\int_0^T \big(X_{n+1}(r)-X_n(r)\big)^2\,dr > \frac{1}{n^6}\bigg\} \\
&\leq \mathbf{P}\bigg\{2\int_0^T \big(X_{n+1}(r)-X(r)\big)^2\,dr > \frac{1}{2\,n^6}\bigg\} + \mathbf{P}\bigg\{2\int_0^T \big(X(r)-X_n(r)\big)^2\,dr > \frac{1}{2\,n^6}\bigg\}
\end{aligned}$$

$$\leq 2^{-n} \quad \text{for } n \in \mathbb{N}. \tag{5.15}$$

By Definition and Theorem 5.3 together with a telescope sum argument, we have

$$\int_0^t X_1\, dB + \sum_{n=1}^{N} \int_0^t (X_{n+1} - X_n)\, dB \to \int_0^t X\, dB \quad \text{in probability as } N \to \infty.$$

Since each term in the sum on the left-hand side is continuous (by Definition and Theorem 4.5), we get the claimed existence of a continuous version of $\int_0^t X\, dB$ if the sum converges uniformly for $t \in [0, T]$ with probability 1. Recall from the proof of Definition and Theorem 4.5 that this in turn will hold if

$$\sum_{n=1}^{\infty} \mathbf{P}\bigg\{ \sup_{t \in [0,T]} \bigg| \int_0^t (X_{n+1} - X_n)\, dB \bigg| > \frac{1}{n^2} \bigg\} < \infty. \tag{5.16}$$

However, (5.16) holds as it follows from Theorem 5.2 together with (5.15) that

$$\mathbf{P}\bigg\{ \sup_{t \in [0,T]} \bigg| \int_0^t (X_{n+1} - X_n)\, dB \bigg| > \frac{1}{n^2} \bigg\} \leq \frac{1}{n^2} + \mathbf{P}\bigg\{ \int_0^T \big(X_{n+1}(r) - X_n(r)\big)^2 dr > \frac{1}{n^6} \bigg\}$$
$$\leq \frac{1}{n^2} + 2^{-n} \quad \text{for } n \in \mathbb{N}. \qquad \square$$

The convergence property in Exercise 56 for the space E_T has the following natural analogue for the space P_T:

Theorem 5.4. *If $X \in P_T$ and $\{X_n\}_{n=1}^{\infty} \subseteq P_T$ satisfy*

$$\int_0^T \big(X_n(r) - X(r)\big)^2 dr \to 0 \quad \textit{in probability as } n \to \infty,$$

then we have in the sense of convergence in probability

$$\int_0^t X_n\, dB \to \int_0^t X\, dB \quad \textit{as } n \to \infty \textit{ for } t \in [0, T]. \tag{5.17}$$

Proof. Use Theorem 5.1 to find sequences $\{X_i^{(n)}\}_{i=1}^{\infty} \subseteq E_T$ such that

$$\int_0^T \big(X_i^{(n)}(r) - X_n(r)\big)^2 dr \to 0 \quad \text{in probability as } i \to \infty \text{ for each } n \in \mathbb{N}. \tag{5.18}$$

By Boole's inequality and the inequality $(x+y)^2 \leq 2\,x^2 + 2\,y^2$ together with (5.17) and (5.18) we have

$$\limsup_{n\to\infty} \limsup_{i\to\infty} \mathbf{P}\bigg\{ \int_0^T \big(X_i^{(n)}(r) - X(r)\big)^2 dr > \delta \bigg\}$$
$$\leq \limsup_{n\to\infty} \limsup_{i\to\infty} \mathbf{P}\bigg\{ \int_0^T \big(X_i^{(n)}(r) - X_n(r)\big)^2 dr > \frac{\delta}{4} \bigg\}$$
$$+ \limsup_{n\to\infty} \mathbf{P}\bigg\{ \int_0^T \big(X_n(r) - X(r)\big)^2 dr > \frac{\delta}{4} \bigg\}$$
$$= 0 \quad \text{for each choice of } \delta > 0.$$

It follows readily that we can find sequences $\{n(j)\}_{j=1}^{\infty}, \{i(j)\}_{j=1}^{\infty} \subseteq \mathbb{N}$ such that

$$\mathbf{P}\bigg\{\int_0^T \big(X_{i(j)}^{(n(j))}(r) - X(r)\big)^2\, dr > \frac{1}{j}\bigg\} \le \frac{1}{j} \quad \text{for } j \in \mathbb{N}.$$

This implies that in the sense of convergence in probability

$$\int_0^T \big(X_{i(j)}^{(n(j))}(r) - X(r)\big)^2\, dr \to 0 \quad \text{as } j \to \infty. \tag{5.19}$$

Now, given any constant $\varepsilon > 0$, in the presence of (5.18) and (5.19), it follows from Boole's inequality and the inequality $(x+y+z)^2 \le 3\,x^2 + 3\,y^2 + 3\,z^2$ together with Theorem 5.2 and Definition and Theorem 5.3 that

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbf{P}\bigg\{\bigg|\int_0^t X\, dB - \int_0^t X_n\, dB\bigg| > \varepsilon\bigg\} \\
&\quad \le \limsup_{j\to\infty} \mathbf{P}\bigg\{\bigg|\int_0^t X\, dB - \int_0^t X_{i(j)}^{(n(j))}\, dB\bigg| > \frac{\varepsilon}{3}\bigg\} \\
&\qquad + \limsup_{n\to\infty} \limsup_{i\to\infty} \limsup_{j\to\infty} \mathbf{P}\bigg\{\bigg|\int_0^t X_{i(j)}^{(n(j))}\, dB - \int_0^t X_i^{(n)}\, dB\bigg| > \frac{\varepsilon}{3}\bigg\} \\
&\qquad + \limsup_{n\to\infty} \limsup_{i\to\infty} \mathbf{P}\bigg\{\bigg|\int_0^t X_i^{(n)}\, dB - \int_0^t X_n\, dB\bigg| > \frac{\varepsilon}{3}\bigg\} \\
&\quad \le 0 + \delta + \limsup_{n\to\infty} \limsup_{i\to\infty} \limsup_{j\to\infty} \mathbf{P}\bigg\{\int_0^T \big(X_{i(j)}^{(n(j))}(r) - X_i^{(n)}(r)\big)^2\, dr \ge \frac{\delta\varepsilon^2}{9}\bigg\} + 0 \\
&\quad \le \delta + \limsup_{j\to\infty} \mathbf{P}\bigg\{\int_0^T \big(X_{i(j)}^{(n(j))}(r) - X(r)\big)^2\, dr \ge \frac{\delta\varepsilon^2}{81}\bigg\} \\
&\qquad + \limsup_{n\to\infty} \mathbf{P}\bigg\{\int_0^T \big(X(r) - X_n(r)\big)^2\, dr \ge \frac{\delta\varepsilon^2}{81}\bigg\} \\
&\qquad + \limsup_{n\to\infty} \limsup_{i\to\infty} \mathbf{P}\bigg\{\int_0^T \big(X_n(r) - X_i^{(n)}(r)\big)^2\, dr \ge \frac{\delta\varepsilon^2}{81}\bigg\} \\
&\quad = \delta \quad \text{for } t \in [0,T], \quad \text{for each choice of } \delta > 0. \qquad \square
\end{aligned}$$

6 Lecture 6, Monday April 27

6.1 Stopping times

Definition 6.1. *A random variable* τ *with values in* $[0,\infty]$ *is called a* stopping time *with respect to a filtration* $\{\mathcal{F}_t\}_{t\geq 0}$ *if it holds that* $\{\tau\leq t\}\in\mathcal{F}_t$ *for* $t\geq 0$.

Exercise 61. Discuss the topic of random variables that can take infinite values.

Exercise 62. Let τ be a stopping time. Show that the events $\{\tau\geq t\}$, $\{\tau<t\}$, $\{\tau>t\}$ and $\{\tau=t\}$ all are $\mathcal{F}_t$-measurable.

Exercise 63. Show that every non-negative real number is a stopping time.

Exercise 64. Show that if τ_1 and τ_2 are stopping times, then $\tau_1\wedge\tau_2$, $\tau_1\vee\tau_2$ and $\tau_1+\tau_2$ are also stopping times.

Example 6.2. If $\{X(t)\}_{t\geq 0}$ is a continuous adapted stochastic process, then the hitting time $\tau_n=\inf\{s\geq 0: X(s)\geq n\}$ of any level $n\in\mathbb{R}$ is a stopping time (recall that $\inf\{\emptyset\}=\infty$), because (see Exercise 65 below)

$$\begin{aligned}\{\tau_n>t\} &= \big\{\inf\{s\geq 0: X(s)\geq n\}>t\big\}\\ &= \big\{\sup\nolimits_{s\in[0,t]}X(s)<n\big\}\\ &= \bigcup_{m=1}^{\infty}\big\{\sup\nolimits_{s\in[0,t]}X(s)\leq n-1/m\big\}\\ &= \bigcup_{m=1}^{\infty}\bigcap_{s\in[0,t]\cap\mathbb{Q}}\{X(s)\leq n-1/m\},\end{aligned}\tag{6.1}$$

where the event on the right-hand side belongs to $\mathcal{F}_t$ as X is adapted.

Exercise 65. Prove the second equality in (6.1).

Exercise 66. Show that $\inf\{s\in[0,T]: X(s)\geq n\}$ is a stopping time for a continuous adapted process $\{X(t)\}_{t\in[0,T]}$.

Exercise 67. If $\{X(t)\}_{t\geq 0}$ is a continuous adapted stochastic process, show that $\inf\{s\geq 0: X(s)\geq n \text{ or } X(s)\leq m\}$ is a stopping time for $\mathbb{R}\ni m<n\in\mathbb{R}$.

Exercise 68. On page 52 in his book[2] Klebaner claims that $\bigcap_{s\in[0,t]}\{X(s)\in D\}=\bigcap_{s\in[0,t]\cap\mathbb{Q}}\{X(s)\in D\}$ for an open $D\subseteq\mathbb{R}$ and a continuous X: Is this true?

Here is one very important result about stopped Itô integral processes:

[2] Fima C. Klebaner: "Introduction to Stochastic Calculus with Applications".

Theorem 6.3. (STOPPING) *For $X \in E_T$ and a stopping time τ we have $\{I_{[0,\tau]}(t) X(t)\}_{t\in[0,T]} \in E_T$ and $\int_0^{t\wedge\tau} X\,dB = \int_0^t I_{[0,\tau]} X\,dB$ for $t \in [0,T]$.*

Proof. By Exercise 62 the process $I_{[0,\tau]}X$ is measurable and adapted, see Exercise 69 below. Consider the following discrete approximation of τ:

$$\tau_n = 2^{-n}\lfloor 2^n\tau+1\rfloor = 2^{-n}(k+1) \quad \text{for } \tau \in [2^{-n}k, 2^{-n}(k+1)), \quad \text{for } k \in \mathbb{N}. \tag{6.2}$$

Then τ_n is a stopping time for $n \in \mathbb{N}$ and $\tau_n \downarrow \tau$ as $n \to \infty$, see Exercise 70 below. Hence the continuity of the Itô integral for E_T gives $\int_0^{t\wedge\tau_n} X\,dB \to \int_0^{t\wedge\tau} X\,dB$ almost surely as $n \to \infty$, while dominated convergence gives

$$\mathbf{E}\bigg\{\int_0^t \big(I_{[0,\tau_n]}(r)X(r) - I_{[0,\tau]}(r)X(r)\big)^2\,dr\bigg\} = \mathbf{E}\bigg\{\int_0^t I_{[\tau,\tau_n)}(r)X(r)^2\,dr\bigg\} \to 0,$$

so that $\int_0^t I_{[0,\tau_n]}X\,dB \to \int_0^t I_{[0,\tau]}X\,dB$ in mean-square as $n \to \infty$ by Exercise 56. Therefore it is sufficient to prove the theorem for each one of the processes $I_{[0,\tau_n]}X$. However, the proof for $I_{[0,\tau_n]}X$ is done as follows using Exercises 49 and 57:

$$\begin{aligned}
\int_0^{t\wedge\tau_n} X\,dB &= \int_0^t X\,dB - \int_{t\wedge\tau_n}^t X\,dB \\
&= \int_0^t X\,dB - \sum_{k=1}^{\infty} I_{\{\tau_n=2^{-n}k\}} \int_{t\wedge 2^{-n}k}^t X\,dB \\
&= \int_0^t X\,dB - \sum_{k=1}^{\infty} \int_0^t I_{[t\wedge 2^{-n}k,t]}\, I_{\{\tau_n=2^{-n}k\}}\, X\,dB \\
&= \int_0^t I_{[0,t]} X\,dB - \int_0^t I_{[t\wedge\tau_n,t]} X\,dB \\
&= \int_0^t I_{[0,t\wedge\tau_n]} X\,dB \\
&= \int_0^t I_{[0,t]}\, I_{[0,\tau_n]} X\,dB \\
&= \int_0^t I_{[0,\tau_n]} X\,dB.
\end{aligned}$$

□

Exercise 69. Show that the process $I_{[0,\tau]}X$ in Theorem 6.3 is measurable and adapted.

Exercise 70. Show that the discrete approximation τ_n of the stopping time τ in (6.2) is a stopping time for each $n \in \mathbb{N}$ and that $\tau_n \downarrow \tau$ as $n \to \infty$.

6.2 Proof of Theorem 5.1

Given an integer $n > 0$, let $X_n = I_{[0,\tau_n]}X$, where

$$\tau_n = T \wedge \inf\left\{t \in [0,T] : \int_0^t X(r)^2\,dr \geq n\right\}. \tag{6.3}$$

As the process $\int_0^t X(r)^2\,dr$ is adapted by Lemma 6.4 below, τ_n is a stopping time by Exercise 64 together with Example 6.2. Hence Exercise 69 shows that X_n is measurable and adapted. Further, we have

$$\int_0^T X_n(r)^2\,dr = \int_0^{\tau_n} X(r)^2\,dr \leq n \tag{6.4}$$

as an immediate consequence of the fact that $\{\int_0^t X(r)^2\,dr\}_{t\in[0,T]}$ is a continuous stochastic process (see Exercise 71 below). Hence we have $X_n \in E_T$. Now, given any constants $\delta, \varepsilon > 0$, since $X \in P_T$, we have $\mathbf{P}\{\int_0^T X(r)^2\,dr \geq n\} \leq \varepsilon$ for some $n = n(\varepsilon) \in \mathbb{N}$. For $n = n(\varepsilon)$ we therefore have

$$\begin{aligned}
\mathbf{P}\left\{\int_0^T (X_n(r) - X(r))^2\,dr > \delta\right\} &\leq \mathbf{P}\{X_n(r) \neq X(r) \text{ for some } r \in [0,T]\} \\
&\leq \mathbf{P}\left\{\int_0^T X(r)^2\,dr \geq n\right\} \\
&\leq \varepsilon.
\end{aligned}$$

□

Exercise 71. Show that the process $\{\int_0^t X(s)^2\,ds\}_{t\in[0,T]}$ is continuous for $X \in P_T$.

Lemma 6.4. *For $X \in P_T$ the process $\{\int_0^t X(r)^2\,dr\}_{t\in[0,T]}$ is adapted.*

Proof. By dominated convergence, with the notation (5.1), we have that $\int_0^t X^{(N)}(r)^2\,dr \to \int_0^t X(r)^2\,dr$ almost surely as $N \to \infty$ (since $X \in P_T$). Hence it is enough to prove the lemma for $X \in E_T$ (since $X^{(N)} \in E_T$). However, for $X \in E_T$ the lemma is a consequence of Theorem 4.4, see Exercise 72 below. □

Exercise 72. Show that Lemma 6.4 for $X \in E_T$ follows from Theorem 4.4.

6.3 Proof of Theorem 5.2

With the notation (6.3), on the event $\{\int_0^T X(r)^2\,dr < C\}$ we have $\tau_C = T$ while Theorem 6.3 gives $\int_0^t X\,dB = \int_0^{t\wedge\tau_C} X\,dB = \int_0^t X_C\,dB$ for $t \in [0,T]$. Hence the Doob-Kolmogorov inequality together with isometry for the Itô integral on E_T and (6.4) give

$$\begin{aligned}
\mathbf{P}\left\{\sup_{t\in[0,T]}\left|\int_0^t X\,dB\right| > \lambda\right\} &\leq \mathbf{P}\left\{\sup_{t\in[0,T]}\left|\int_0^t X_C\,dB\right| > \lambda\right\} + \mathbf{P}\left\{\int_0^T X(r)^2\,dr \geq C\right\} \\
&= \mathbf{E}\left\{\int_0^T X_C(r)^2\,dr\right\}\Big/\lambda^2 + \mathbf{P}\left\{\int_0^T X(r)^2\,dr \geq C\right\} \\
&\leq \frac{C}{\lambda^2} + \mathbf{P}\left\{\int_0^T X(r)^2\,dr \geq C\right\}.
\end{aligned}$$

□

Exercise 73. Show that if (5.11) holds for $X \in E_T$ and $C = 1$, then (5.11) holds for $X \in E_T$ and all $C > 0$.

6.4 Properties of Itô integrals for the space P_T

Exercise 74. (NADA) Show that for an $X \in P_T$ we have

$$\int_r^t X\,dB = -\int_t^r X\,dB \quad \text{and} \quad \int_0^t I_{[r,s]} X\,dB = \int_r^s X\,dB \quad \text{for } 0 \le r \le s \le t \le T.$$

Exercise 75. (LINEARITY) Show that for $X, Y \in P_T$ and $a, b \in \mathbb{R}$, we have

$$\int_0^t (aX + bY)\,dB = a\int_0^t X\,dB + b\int_0^t Y\,dB \quad \text{for } t \in [0,T].$$

Exercise 76. (ADAPTEDNESS) Show that for $X \in P_T$ the Itô integral process $\{\int_0^t X\,dB\}_{t\in[0,T]}$ is adapted.

Exercise 77. (ZIPP) Show that $\int_0^t I_{[s,t]}(r) Y X(r)\,dB(r) = Y \int_s^t X(r)\,dB(r)$ for $0 \le s \le t \le T$ and $X \in P_T$, when Y is an $\mathcal{F}_s$-measurable random variable that is bounded by a (non-random) constant.

As Exercises 74 and 77 extend the results of Exercises 49 and 57 from E_T to P_T, and as Theorem 5.4 gives a version for P_T of the convegence result in Exercise 56 for E_T, it is straightforward to modify the proof of Theorem 6.3 to work for P_T:

Theorem 6.5. (STOPPING) *For $X \in P_T$ and a stopping time τ we have $\{I_{[0,\tau]}(t)\, X(t)\}_{t\in[0,T]} \in P_T$ and $\int_0^{t\wedge\tau} X\,dB = \int_0^t I_{[0,\tau]} X\,dB$ for $t \in [0,T]$.*

Exercise 78. Prove Theorem 6.5.

Exercise 79. Discuss the stopping property for the space S_T.

6.5 Approximation of Itô integrals

Exercise 80. Show that for an $X \in P_T$ we have

$$\int_0^T (X_n(r) - X(r))^2\,dr \to 0 \quad \text{as } n \to \infty \quad \text{for some sequence } \{X_n\}_{n=1}^{\infty} \subseteq S_T$$

in the sense of convergence in probability. Conclusions?

With the stopping property available for the space P_T in Theorem 6.5, the proof of Theorem 5.2 for the space E_T carries over to work also for the space P_T:

Theorem 6.6. *For $X \in P_T$ and a constant $C > 0$, we have*

$$\mathbf{P}\left\{\sup_{t\in[0,T]}\left|\int_0^t X\,dB\right| > \lambda\right\} \leq \frac{C}{\lambda^2} + \mathbf{P}\left\{\int_0^T X(r)^2\,dr \geq C\right\} \quad \textit{for } \lambda > 0.$$

Exercise 81. Prove Theorem 6.6.

The next approximation result is a key ingredient to prove so called Itô formulas:

Theorem 6.7. *A continuous adapted process $\{X(t)\}_{t\in[0,T]}$ belongs to P_T and satisfies*

$$\sup_{t\in[0,T]}\left|\int_0^t X\,dB - \int_0^t \sum_{i=1}^n X(t_{i-1}) I_{(t_{i-1},t_i]}\,dB\right| \to 0 \quad \textit{in probability}$$

for partitions $0 = t_0 < t_1 < \ldots < t_n = T$ of $[0,T]$ such that $\max_{1\leq i\leq n} t_i - t_{i-1} \downarrow 0$.

Proof. Clearly, we have $X \in P_T$. Further, Theorem 6.6 shows that

$$\begin{aligned}
&\mathbf{P}\left\{\sup_{t\in[0,T]}\left|\int_0^t X\,dB - \int_0^t \sum_{i=1}^n X(t_{i-1}) I_{(t_{i-1},t_i]}\,dB\right| > \varepsilon\right\} \\
&\qquad \leq \delta + \mathbf{P}\left\{\sum_{i=1}^n \int_{t_{i-1}}^{t_i} (X(r) - X(t_{i-1}))^2\,dr \geq \delta\varepsilon^2\right\} \\
&\qquad \leq \delta + \mathbf{P}\left\{\sup_{s,t\in[0,T],\,|s-t|\leq \max_{1\leq i\leq n} t_i - t_{i-1}} (X(s)-X(t))^2 \int_0^T dr \geq \delta\varepsilon^2\right\} \\
&\qquad \to \delta \quad \text{as} \quad \max_{1\leq i\leq n} t_i - t_{i-1} \downarrow 0 \quad \text{for each choice of } \delta, \varepsilon > 0,
\end{aligned}$$

by uniform continutity of X. $\square$

Exercise 82. Explain why the conclusion of Theorem 6.7 cannot hold for a general (not necessarily continuous) $X \in P_T$.

7 Lecture 7, Wednesday April 29

7.1 Itô processes and stochastic differentials

Definition 7.1. *If $\{\mu(t)\}_{t\in[0,T]}$ is an adapted measurable stochastic process such that*

$$\mathbf{P}\bigg\{\int_0^T |\mu(r)|\,dr < \infty\bigg\} = 1,$$

if $\sigma \in P_T$, and if $X(0)$ is an $\mathcal{F}_0$-measurable random variable, then we call

$$X(t) = X(0) + \int_0^t \mu(r)\,dr + \int_0^t \sigma\,dB, \quad t \in [0,T],$$

an Itô process *and*

$$dX(t) = \mu(t)\,dt + \sigma(t)\,dB(t)$$

the corresponding stochastic differential.

Observe that the the stochastic differential only is a convenient short-hand notation for the corresponding Itô process, from which it earns its mathematical meaning.

Note the difference between the concepts of Itô integral process and Itô process!

Example 7.2. The two most basic examples of stochastic differentials are $dX(t) = dt$ and $dX(t) = dB(t)$.

Exercise 83. Explain why Itô processes are continuous with probability 1.

Exercise 84. Show that the Lebesgue integral part $\{\int_0^t \mu(r)\,dr\}_{t\in[0,T]}$ of an Itô process is adapted, e.g., by applying Lemma 6.4 to $\sqrt{\mu^+}, \sqrt{\mu^+} \in P_T$.

Definition 7.3. *If $dX(t) = \mu(t)\,dt + \sigma(t)\,dB(t)$ is a stochastic differential and $\{Y(t)\}_{t\in[0,T]}$ an adapted measurable process such that*

$$\mathbf{P}\bigg\{\int_0^T |Y(r)|\,|\mu(r)|\,dr < \infty\bigg\} = 1$$

and $Y\sigma \in P_T$, then we define the Itô process $\{\int_0^t Y\,dX\}_{t\in[0,T]}$ by

$$\int_0^t Y\,dX = \int_0^t Y(r)\mu(r)\,dr + \int_0^t Y\sigma\,dB \quad \textit{for } t \in [0,T].$$

In order to make sure that the above definition of $\int_0^t Y\,dX$ is consistent, in the sense of not being multi-valued, we must check that if two Itô processes agree

$$\mathbf{P}\bigg\{\int_0^t \mu_1(r)\,dr + \int_0^t \sigma_1\,dB = \int_0^t \mu_2(r)\,dr + \int_0^t \sigma_2\,dB \text{ for } t \in [0,T]\bigg\} = 1, \qquad (7.1)$$

then

$$\mathbf{P}\bigg\{\int_0^t Y(r)\mu_1(r)\,dr+\int_0^t Y\sigma_1\,dB=\int_0^t Y(r)\mu_2(r)\,dr+\int_0^t Y\sigma_2\,dB \text{ for } t\in[0,T]\bigg\}=1. \tag{7.2}$$

However, this follows from Corollary 8.7 below, according to which (7.1) implies

$$\mathbf{P}\{\sigma_1(t)=\sigma_2(t) \text{ a.e. for } t\in[0,T]\}=\mathbf{P}\{\mu_1(t)=\mu_2(t) \text{ a.e. for } t\in[0,T]\}=1,$$

which in turn implies the required (7.2) (recall Definition and Theorem 5.3).

Exercise 85. Explain why the Itô process $\int_0^t Y\,dX$ is well-defined when X is an Itô process and Y is a continuous adapted process. Conclude that the Itô process $\int_0^t Y\,dX$ is well-defined when X and Y are both Itô processes.

The following generalization of Theorem 6.7 will be a crucial tool in the sequel:

Theorem 7.4. *For an Itô process $\{X(t)\}_{t\in[0,T]}$ and a continuous adapted process $\{Y(t)\}_{t\in[0,T]}$, we have*

$$\sup_{t\in[0,T]}\bigg|\int_0^t Y\,dX-\int_0^t\sum_{i=1}^n Y(t_{i-1})I_{(t_{i-1},t_i]}\,dX\bigg|\to 0 \quad \textit{in probability}$$

for partitions $0=t_0<t_1<\ldots<t_n=T$ *of* $[0,T]$ *such that* $\max_{1\leq i\leq n} t_i-t_{i-1}\downarrow 0$.

Proof. First note the obvious fact that

$$\begin{aligned}&\sup_{t\in[0,T]}\bigg|\int_0^t Y\,dX-\int_0^t\sum_{i=1}^n Y(t_{i-1})I_{(t_{i-1},t_i]}\,dX\bigg|\\ &\quad\leq \sup_{t\in[0,T]}\bigg|\int_0^t Y(r)\mu(r)\,dr-\int_0^t\sum_{i=1}^n Y(t_{i-1})I_{(t_{i-1},t_i]}(r)\mu(r)\,dr\bigg|\\ &\qquad+\sup_{t\in[0,T]}\bigg|\int_0^t Y\sigma\,dB-\int_0^t\sum_{i=1}^n Y(t_{i-1})I_{(t_{i-1},t_i]}\sigma\,dB\bigg|.\end{aligned}$$

Here the first term on the righ-hand side is at most

$$\sup_{r,s\in[0,T],\,|r-s|\leq\max_{1\leq i\leq n} t_i-t_{i-1}}|Y(r)-Y(s)|\int_0^T|\mu(r)|\,dr\to 0 \quad \text{almost surely as } n\to\infty$$

by uniform continuity of Y (see Exercise 83). To prove that the second term on the right hand goes to 0 in probability we just need the following slight adjustment of the argument employed in the proof of Theorem 6.7 based on Theorem 6.6:

$$\mathbf{P}\bigg\{\sup_{t\in[0,T]}\bigg|\int_0^t Y\sigma\,dB-\int_0^t\sum_{i=1}^n Y(t_{i-1})I_{(t_{i-1},t_i]}\sigma\,dB\bigg|>\varepsilon\bigg\}$$

$$\le \delta + \mathbf{P}\Big\{\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(Y(r)-Y(t_{i-1}))^2\sigma(r)^2\,dr \ge \delta\varepsilon^2\Big\}$$

$$\le \delta + \mathbf{P}\Big\{\sup_{r,s\in[0,T],\,|r-s|\le \max_{1\le i\le n} t_i-t_{i-1}}(Y(r)-Y(s))^2\int_0^T\sigma(r)^2\,dr \ge \delta\varepsilon^2\Big\}$$

$$\to \delta \quad \text{as} \quad \max_{1\le i\le n} t_i - t_{i-1}\downarrow 0 \quad \text{for each choice of } \delta,\varepsilon>0. \qquad \square$$

7.2 Itô formula

Here is a first version of the immensely important Itô formula (Itô lemma):

Theorem 7.5. (ITÔ FORMULA) *For a function $f\in C^2(\mathbb{R})$[3] we have*

$$df(B(t)) = f'(B(t))\,dB(t) + \frac{1}{2}\,f''(B(t))\,dt. \tag{7.3}$$

Note that what (7.3) really means is that (recall Definition 7.1)

$$f(B(t)) = f(B(0)) + \int_0^t f'(B(r))\,dB(r) + \frac{1}{2}\int_0^t f''(B(r))\,dr. \tag{7.4}$$

Proof of Theorem 7.5. Consider partitions $0=t_0<t_1<\ldots<t_n=t$ of the interval $[0,t]$ such that $\max_{1\le i\le n} t_i - t_{i-1}\downarrow 0$. By Taylor expansion, we have

$$\begin{aligned}
f(B(t)) - f(B(0)) &= \sum_{i=1}^{n} f(B(t_i)) - f(B(t_{i-1}))\\
&= \sum_{i=1}^{n} f'(B(t_{i-1}))\,(B(t_i)-B(t_{i-1}))\\
&\qquad + \frac{1}{2}\sum_{i=1}^{n} f''(B(t_{i-1}))\,(B(t_i)-B(t_{i-1}))^2\\
&\qquad + \sum_{i=1}^{n}\int_{B(t_{i-1})}^{B(t_i)}(B(t_i)-r)\,(f''(r)-f''(B(t_{i-1})))\,dr,
\end{aligned} \tag{7.5}$$

see also Exercise 86 below. The first term on the right-hand side converges to $\int_0^t f'(B)\,dB$ in probability by Theorem 6.7, as $f(B)$ is a continuous and adapted process. Recalling from Theorem 1.10 that the quadratic variation of BM over an interval is the length of that interval (in the sense of convergence in mean-square), see also Example 7.9 and Theorem 8.6 below, the second term on the right-hand side of (7.5) converges to $\frac{1}{2}\int_0^t f''(B(r))\,dr$ in probability by uniform continuity of $f''(B)$, see Exercise 87 below. Moreover, the third term on the right-hand side of (7.5) is bounded by

$$\sup_{r,s\in[0,T],\,|r-s|\le \max_{1\le i\le n} t_i-t_{i-1}} |f''(B(r))-f''(B(s))|\sum_{i=1}^{n}\int_{B(t_{i-1})}^{B(t_i)}(B(t_i)-r)\,dr, \tag{7.6}$$

[3] The class of two times continuously differentiable functions $f:\mathbb{R}\to\mathbb{R}$.

which goes to 0 with probability 1, see Exercise 88 below. This establishes (7.4). $\square$

Exercise 86. Prove the Taylor expansion (7.5).

Exercise 87. Prove that the second term on the right-hand side of (7.5) converges to $\frac{1}{2}\int_0^t f''(B(r))\,dr$ in probability, e.g., by means of introducing a second cruder grid $\{t'_j\}_{j=1}^m$, approximating the value of $f''(B(t_{i-1}))$ by $f''(B(t'_{j-1}))$ for an appropriate j, and sending first $\max_{1\leq i\leq n} t_i - t_{i-1}\downarrow 0$ and then $\max_{1\leq j\leq m} t'_j - t'_{j-1}\downarrow 0$ afterwards, the idea being that this makes it possible to replace $(B(t_i)-B(t_{i-1}))^2$ with $t_i - t_{i-1}$ in the first limit as $\max_{1\leq i\leq n} t_i - t_{i-1}\downarrow 0$, and that the approximation of $f''(B(t_{i-1}))$-values by $f''(B(t'_{j-1}))$-values is accurate in the second limit as $\max_{1\leq j\leq m} t'_j - t'_{j-1}\downarrow 0$, by the continuity of $f''(B)$.

Exercise 88. Explain why the third term on the right-hand side of (7.5) is bounded by the expression (7.6), and why that expression goes to 0.

Theorem 7.6. (Itô formula) *For a function $f\in C^2(\mathbb{R})$ and an Itô processes X with stochastic differential $dX(t)=\mu(t)\,dt+\sigma(t)\,dB(t)$, we have*

$$df(X(t)) = f'(X(t))\,dX(t) + \frac{1}{2}\,f''(X(t))\,\sigma(t)^2\,dt. \tag{7.7}$$

Proof. Replace B with X everywhere in (7.5). From an inspection of the resulting equation and the proof of Theorem 7.5, we readily conclude that it is sufficient to prove that, in the sense of convergence if probability, X has quadratic variation

$$\lim_{\max_{1\leq i\leq n} t_i-t_{i-1}\downarrow 0}\sum_{i=1}^n (X(t_i)-X(t_{i-1}))^2 = \int_s^t \sigma(r)^2\,dr \quad \text{for } 0\leq s\leq t, \tag{7.8}$$

where $s=t_0<t_1<\ldots<t_n=t$ are finer and finer partitions of the interval $[s,t]$. However, (7.8) in turn coincides with Corollary 8.8 below. $\square$.

Exercise 89. Elaborate on those details of the proof of Theorem 7.6 that was left out above. In particular, explain in detail how Theorem 7.6 follows from (7.8). (See also the proof of Theorem 9.10 in Section 11.2 below.)

7.3 Introduction to quadratic variation and variation

We will now start to investigate in greater depth the topics of quadratic variation and variation that were introduced in Theorems 1.10-1.11 and Exercises 12-13.

Definition 7.7. *The* quadratic variation $[f]([s,t])$ *over an interval* $[s,t]\subseteq[0,T]$ *of a function* $f:[0,T]\to\mathbb{R}$ *is defined as the limit*

$$[f]([s,t]) = \lim_{\max_{1\leq i\leq n} t_i-t_{i-1}\downarrow 0} \sum_{i=1}^{n}(f(t_i)-f(t_{i-1}))^2$$

where $s=t_0<t_1<\ldots<t_n=t$ *are finer and finer partitions of* $[s,t]$*, whenever these limits have well-defined values in* $[0,\infty]$ *for all intervals* $[s,t]\subseteq[0,T]$ *(which they a priori need not have). We use* $[f](t)$ *as short-hand notation for* $[f]([0,t])$*.*

The quadratic covariation $[f,g]([s,t])$ *over an interval* $[s,t]\subseteq[0,T]$ *between two functions* $f,g:[0,T]\to\mathbb{R}$ *is defined as the limit*

$$[f,g]([s,t]) = \lim_{\max_{1\leq i\leq n} t_i-t_{i-1}\downarrow 0} \sum_{i=1}^{n}(f(t_i)-f(t_{i-1}))\,(g(t_i)-g(t_{i-1})),$$

where $s=t_0<t_1<\ldots<t_n=t$*, whenever these limits have well-defined values in* $\mathbb{R}$ *for all intervals* $[s,t]\subseteq[0,T]$*. We use* $[f,g](t)$ *as short-hand notation for* $[f,g]([0,t])$*.*

Definition 7.8. *The* variation $V_f([s,t])$ *over an interval* $[s,t]\subseteq[0,T]$ *of a function* $f:[0,T]\to\mathbb{R}$ *is defined as the limit*

$$V_f([s,t]) = \lim_{\max_{1\leq i\leq n} t_i-t_{i-1}\downarrow 0} \sum_{i=1}^{n}|f(t_i)-f(t_{i-1})|$$

where $s=t_0<t_1<\ldots<t_n=t$ *are finer and finer partitions of* $[s,t]$*, whenever these limits have well-defined values in* $[0,\infty]$ *for all intervals* $[s,t]\subseteq[0,T]$*. We use* $V_f(t)$ *as short-hand notation for* $V_f([0,t])$*.*

Note the obvious facts that $V_f([s,t])$ and $[f]([s,t])$ are non-increasing functions of s and non-decreasing functions of t, and that $[f,f]=[f]$.

We will use the concepts of quadratic variation, quadratic covariation and variation for stochastic processes rather than just ordinary functions. In doing so the questions arises in what sense we should understand the convergences in Definitions 7.7 and 7.8. The answer to that in turn is in the sense of convergence in probability.

Example 7.9. Recall from Theorems 1.10 and 1.11 that BM has quadratic variation $[B]([s,t])=t-s$ and infinite variation $V_B([s,t])=\infty$ with probability 1 for $0\leq s<t<\infty$.

When a stochastic process $\{X(t)\}_{t\in[0,T]}$ has a quadratic variation process $\{[X](t)\}_{t\in[0,T]}$ (obtained as a limit in probability for each $t\in[0,T]$) which possesses a version that is continuous (with probability 1), then we will always select such a continuous (with probability 1) version of the quadratic variation. (All quadratic variation processes that we will encounter possess continuous versions.) Besides ruling out some uninteresting pathological versions of quadratic variations, see Exercise 90 below, the importance of this convention lies in that it ensures that (continuous with probability 1) quadratic variation processes that are versions of each other do agree (along the whole trajectory of the process) with probability 1, see the discussion after Definition and Theorem 4.5 and Exercise 48.

Exercise 90. Albeit we know that BM B has qudratic variation process $[B](t) = t$ for $t \geq 0$, which obviuously is an example of a continuous quadratic varation process, show that if we had not yielded to the above convention to always select a continuous quadratic variation whenever possible, then, for example, BM would also have had quadratic variation process $[B](t) = t\, I_{\{t \neq \xi\}}$ for $t \geq 0$, where ξ is a unit mean exponentially distributed random variable that is independent of B. (Just to make life complicated, that is!)

Even if we only accept continuous (with probability 1) versions of quadratic variation processes, when such continuous versions exist, the fact that we are dealing with stochastic processes that are defined as limits in probability (such as, for example, Itô integral processes or quadratic variation processes) still allows some pathologies to occur, as is illustrated by the next exercise:

Exercise 91. Show that BM B has qudratic variation process $[B](\omega, t) = t\, I_U(\omega)$ for $t \geq 0$, whenever U is a null event.

8 Lecture 8, Monday May 4

8.1 Introduction to stochastic differential equations

Equipped with Itô's formula for Itô processes Theorem 7.6, we can consider a first introductory example of a *stochastic differential equation* (SDE): An *SDE of diffusion type* (there are other more general types of SDE as well) takes the form

$$dX(t) = \mu(t, X(t))\, dt + \sigma(t, X(t))\, dB(t) \quad \text{for } t \in [0,T], \quad X(0) = X_0,$$

where $\mu, \sigma : [0,T] \times \mathbb{R} \to \mathbb{R}$ are measurable functions and X_0 is an $\mathcal{F}_0$-measurable random variable. A solution to this SDE is any process $\{X(t)\}_{t\in[0,T]}$ that satisfies

$$X(t) = X_0 + \int_0^t \mu(s, X(s))\, ds + \int_0^t \sigma(s, X(s))\, dB(s) \quad \text{for } t \in [0,T].$$

Note that in particular, to make the integrals on the right-hand side well-defined, X has to be measurable and adapted with

$$\mathbf{P}\Big\{\int_0^T |\mu(t, X(t))|\, dt < \infty\Big\} = \mathbf{P}\Big\{\int_0^T \sigma(t, X(t))^2\, dt < \infty\Big\} = 1.$$

Starting next lecture, our focus will be more or less exclusively on such SDE's.

Example 8.1. (Stochastic exponential) Take $\mu(t,x) = 0$ and $\sigma(t,x) = \sigma\, x$ for a constant $\sigma > 0$ to obtain the SDE

$$dX(t) = \sigma\, X(t)\, dB(t) \quad \text{for } t \in [0,T], \quad X(0) = X_0, \tag{8.1}$$

where X_0 is a strictly positive initial value. Apply the Itô formula to obtain an SDE for the transformed process $Y(t) = f(X(t))$ with $f(x) = \ln(x)$, as

$$dY(t) = f'(X(t))\, dX(t) + \frac{1}{2} f''(X(t))\, \sigma(t, X(t))^2\, dt = \ldots = \sigma\, dB(t) - \frac{1}{2}\sigma^2\, dt.$$

From this we conclude that

$$Y(t) = Y(0) + \int_0^t \sigma\, dB(s) - \frac{1}{2}\int_0^t \sigma^2\, ds = \ln(X_0) + \sigma\, B(t) - \frac{1}{2}\sigma^2\, t.$$

Now a bit of care is called for as f is only two times continuously differentiable on the interval $(0,\infty)$ rather than on whole $\mathbb{R}$ as required in Theorem 7.6. However, clearly $Y(t) = \ln(X_0) + \sigma\, dB(t) - \frac{1}{2}\sigma^2\, dt$ satisfies the equation $dY(t) = \sigma\, dB(t) - \frac{1}{2}\sigma^2\, dt$. Hence we may apply the two times continuously differentiable function $g(x) = \mathrm{e}^x$ to the Itô process Y to obtain an SDE for $X(t) = \mathrm{e}^{Y(t)}$ as

$$dX(t) = g'(Y(t))\, dY(t) + \frac{1}{2} g''(Y(t))\, \sigma^2\, dt = \ldots = \sigma\, X(t)\, dB(t).$$

Observing that $X(0) = \mathrm{e}^{Y(0)} = \mathrm{e}^{\ln(X_0)} = X_0$ we conclude that $X(t) = \mathrm{e}^{Y(t)} = \mathrm{e}^{\ln(X_0) + \sigma\, dB(t) - \frac{1}{2}\sigma^2\, dt} = X_0\, \mathrm{e}^{\sigma\, dB(t) - \frac{1}{2}\sigma^2\, dt}$ solves the SDE (8.1).

8.2 Quadratic variation and variation

The time has come to study quadratic variation and variation of Itô processes.

Theorem 8.2. *A continuous function $f:[0,T]\to\mathbb{R}$ with non-zero quadratic variation $[f]([s,t])>0$ for some $[s,t]\subseteq[0,T]$ has infinite variation $V_f([s,t])=\infty$.*

Exercise 92. Prove Theorem 8.2.

Exercise 93. Exemplify that the statement of Theorem 8.2 need not be true for non-continuous functions.

Exercise 94. Find the quadratic variation and the variation of a Poisson process.

Exercise 95. Show that a continuously differentiable function $f:[0,T]\to\mathbb{R}$ has zero quadratic variation $[f](T)=0$ and finite variation $V_f(T)<\infty$.

Theorem 8.3. *The quadratic covariation between a continuous function $f:[0,T]\to\mathbb{R}$ and a function $g:[0,T]\to\mathbb{R}$ with finite variation $V_g([s,t])<\infty$ over $[s,t]\subseteq[0,T]$ is zero $[f,g]([s,t])=0$.*

Exercise 96. Prove Theorem 8.3.

Exercise 97. Find the quadratic covariation between Brownian motion and a Poisson process (defined on a mutual probability space).

In Exercise 40 we found the following characterization of finite variation functions:

Theorem 8.4. *A function $f:[0,T]\to\mathbb{R}$ has finite variation $V_f(T)<\infty$ if and only if f can be written as the diffrerence between two non-decreasing functions.*

Unsuprisingly by inspection of the definition, quadratic variations have several properties in common with covariances. Such properties include symmetry, linearity in each of the arguments, as well as the so called polarization identity:

Exercise 98. Make statements about symmetry and linearity of quadratic covariations and prove them.

Theorem 8.5. (POLARIZATION) *For functions $f, g : [0,T] \to \mathbb{R}$ with well-defined quadratic variations and a well-defined mutual quadratic covariation, we have*

$$[f,g] = \frac{1}{2}\big([f+g] - [f] - [g]\big) = \frac{1}{4}\big([f+g] - [f-g]\big).$$

Exercise 99. Prove Theorem 8.5.

Theorems 8.2-8.5 carry over with only obvious modifications from ordinary functions to stochastic process. The statements should then be understood to hold with probability 1, and conditions on the processes involved (such as continuity, finite variation, etc.) should be satisfied probability 1.

The following result is an absolutely crucial part of the build up of the theory:

Theorem 8.6. *The quadratic variation of an Itô integral process $\{X(t)\}_{t\in[0,T]}$ with stochastic differential $dX(t) = \sigma(t)\,dB(t)$ is given (with probability 1) by*

$$[X]([s,t]) = \int_s^t \sigma(r)^2\,dr \quad \textit{for } [s,t] \subseteq [0,T].$$

Proof. It is sufficient to prove the theorem for the interval $[s,t] = [0,T]$, see Exercise 100 below. We begin with the case when $\sigma \in S_T$, so that for some constants $0 = s_0 < s_1 < \ldots < s_m = T$ and some square-integrable random variables $\sigma(0), \sigma_{s_0}, \ldots, \sigma_{s_{m-1}}$ that are adapted to $\mathcal{F}_0, \mathcal{F}_{t_0}, \ldots, \mathcal{F}_{t_{m-1}}$, respectively, we have

$$\sigma(t) = \sigma(0) I_{\{0\}}(t) + \sum_{i=1}^m \sigma_{s_{i-1}} I_{(s_{i-1},s_i]}(t) \quad \text{for } t \in [0,T].$$

Now, any given grid $0 = t_0 < t_1 < \ldots < t_n = T$ may be refined to a grid $0 = t'_0 < t'_1 < \ldots < t'_k = T$ with at most $n+m-1$ members that also includes the times $0 < s_1 < \ldots < s_{m-1} < T$. Writing $s_j = t'_{i(j)}$ for $j = 1, \ldots, m-1$, we then have

$$\begin{aligned}
&\bigg|\sum_{i=1}^n (X(t_i) - X(t_{i-1}))^2 - \sum_{i=1}^k (X(t'_i) - X(t'_{i-1}))^2\bigg| \\
&\leq \sum_{j=1}^{m-1}\Big(\big(X(t'_{i(j)+1}) - X(t'_{i(j)-1})\big)^2 + \big(X(t'_{i(j)+1}) - X(t'_{i(j)})\big)^2 + \big(X(t'_{i(j)}) - X(t'_{i(j)-1})\big)^2\Big) \\
&\leq 3\sum_{j=1}^{m-1}\Big(\big(X(t'_{i(j)+1}) - X(t'_{i(j)})\big)^2 + \big(X(t'_{i(j)}) - X(t'_{i(j)-1})\big)^2\Big) \\
&\to 0 \quad \text{with probability 1 as } \max_{1\leq i\leq k} t'_i - t'_{i-1} \leq \max_{1\leq i\leq n} t_i - t_{i-1} \downarrow 0
\end{aligned} \tag{8.2}$$

by the continuity of X, see also Exercise 101 below. To prove the theorem for $\sigma \in S_T$ it is therefore sufficient to prove that, in the sense of convergence in probability,

$$\lim_{\max_{1\le i\le n} t_i - t_{i-1}\downarrow 0} \sum_{i=1}^{n} (X(t_i) - X(t_{i-1}))^2 = \int_0^T \sigma(r)^2\, dr \tag{8.3}$$

for grids $0 = t_0 < t_1 < \ldots < t_n = T$ that include $0 < s_1 < \ldots < s_{m-1} < T$. That this is so in turn follows from the fact that $[B]([s,t]) = t - s$, see Exercise 101 below.

Now assume that we have extended the theorem from S_T to any $\sigma \in E_T$. For a stochastic differential $dX(t) = \sigma(t)\, dB(t)$ with $\sigma \in P_T$ we may then approximate the processes σ and X with the processes $\sigma_n(t) = \sigma(t) I_{[0,\tau_n]}(t)$ and $X_n(t) = \int_0^t \sigma_n(r)\, dB(r)$, respectively, where

$$\tau_n = T \wedge \inf\left\{ t \in [0,T] : \int_0^t \sigma(s)^2\, ds \ge n \right\},$$

cf. the proofs of Theorems 5.1 and 5.2. Since σ and $\sigma_n \in E_T$ trivially coincide when $\tau_n = T$, as do X and X_n by Theorem 6.5, we have

$$[X](T) \leftarrow [X_n](T) = \int_0^T \sigma_n(r)^2\, dr \to \int_0^T \sigma(r)^2\, dr \quad \text{as } n \to \infty$$

in the sense of convergence in probability. This in turn is so because

$$\begin{aligned} &\mathbf{P}\{|[X](T) - [X_n](T)| > \delta\} \vee \mathbf{P}\left\{\left|\int_0^T \sigma_n(r)^2\, dr - \int_0^T \sigma(r)^2\, dr\right| > \delta\right\} \\ &\quad \le \mathbf{P}\{[X](T) \ne [X_n](T)\} \vee \mathbf{P}\left\{\int_0^T \sigma_n(r)^2\, dr \ne \int_0^T \sigma(r)^2\, dr\right\} \\ &\quad \le \mathbf{P}\{\tau_n < T\} \\ &\quad \to 0 \quad \text{as } n \to \infty \quad \text{for any } \delta > 0, \end{aligned}$$

since $\sigma \in P_T$. As for the assumed extension of the theorem from S_T to E_T, which is the most difficult part of the proof, we will return to in Section 11.1 below. □

Exercise 100. Explain why it is sufficient to prove Theorem 8.6 for the interval $[s,t] = [0,T]$.

Exercise 101. Explain the details of equations (8.2) and (8.3).

Corollary 8.7. *If* $\mathbf{P}\{X(t) = Y(t)$ *for* $t \in [0,T]\} = 1$ *for two Itô processes* $\{X(t)\}_{t\in[0,T]}$ *and* $\{Y(t)\}_{t\in[0,T]}$ *with stochastic differentials* $dX(t) = \mu_X(t)\, dt + \sigma_X(t)\, dB(t)$ *and* $dY(t) = \mu_Y(t)\, dt + \sigma_Y(t)\, dB(t)$*, respectively, then we have*

$$\mathbf{P}\{\sigma_X(t) = \sigma_Y(t) \text{ a.e. for } t \in [0,T]\} = \mathbf{P}\{\mu_X(t) = \mu_Y(t) \text{ a.e. for } t \in [0,T]\} = 1.$$

Proof. From $\mathbf{P}\{X(t)=Y(t) \text{ for } t\in[0,T]\}=1$ we get by rearrangement that

$$\int_0^t (\sigma_X-\sigma_Y)\,dB=\int_0^t (\mu_Y(r)-\mu_X(r))\,dr \quad \text{for } t\in[0,T] \tag{8.4}$$

with probability 1. As the processes on both sides in (8.4) are continuous and have finite variation, see also Exercise 102 below, Theorem 8.3 shows that these processes have zero quadratic variation over the interval $[0,T]$ with probability 1. From this in turn we get $\int_0^T(\sigma_X(r)-\sigma_Y(r))^2\,dr=0$ with probability 1 by Theorem 8.6. Hence $\int_0^t(\sigma_X-\sigma_Y)\,dB=0$ for $t\in[0,T]$ with probability 1, so that the process on both sides in (8.4) are zero with probability 1. Now we have established that

$$\mathbf{P}\bigg\{\int_0^t(\sigma_X(r)-\sigma_Y(r))^2\,dr=\int_0^t(\mu_Y(r)-\mu_X(r))\,dr=0 \text{ for } t\in[0,T]\bigg\}=1.$$

This in turn happens if and only if the conclusion of the theorem holds. □

Exercise 102. Prove that the stochastic process $[0,T]\ni t\curvearrowright\int_0^t\mu(r)\,dr\in\mathbb{R}$ has finite variation over any interval $[s,t]\subseteq[0,T]$ when $\mathbf{P}\{\int_0^T|\mu(r)|\,dr<\infty\}=1$ for a measurable and adapted process $\{\mu(t)\}_{t\in[0,T]}$.

Corollary 8.8. *The quadratic variation of an Itô process $\{X(t)\}_{t\in[0,T]}$ with stochastic differential $dX(t)=\mu(t)\,dt+\sigma(t)\,dB(t)$ is given (with probability 1) by*

$$[X]([s,t])=\int_s^t\sigma(r)^2\,dr \quad \textit{for } [s,t]\subseteq[0,T].$$

Exercise 103. Prove Corollary 8.8.

Exercise 104. Show that an Itô process has zero quadratic variation if and only if it has finite variation.

Corollary 8.9. *The quadratic covariation between two Itô processes $\{X(t)\}_{t\in[0,T]}$ and $\{Y(t)\}_{t\in[0,T]}$ with stochastic differentials $dX(t)=\mu_X(t)\,dt+\sigma_X(t)\,dB(t)$ and $dY(t)=\mu_Y(t)\,dt+\sigma_Y(t)\,dB(t)$, respectively, is given (with probability 1) by*

$$[X,Y]([s,t])=\int_s^t\sigma_X(r)\sigma_Y(r)\,dr \quad \textit{for } [s,t]\subseteq[0,T].$$

Exercise 105. Prove Corollary 8.9.

For future needs, we also adress the topic of quadratic covariation between Itô processes driven by different independent Brownian motions.

Exercise 106. Show that the quadratic covariation between two independent Brownian motions (defined on a mutual probability space) is zero.

Theorem 8.10. *The quadratic covariation between two Itô processes driven by independent Brownian motions is zero.*

Exercise 107. Prove Theorem 8.10.

9 Lecture 9, Wednesday May 6

9.1 Stochastic differential equations (SDE)

Definition 9.1. *A* stochastic differential equation (SDE) of diffusion type *is given by*

$$dX(t) = \mu(t, X(t))\, dt + \sigma(t, X(t))\, dB(t) \quad \textit{for } t \in [0,T], \quad X(0) = X_0, \tag{9.1}$$

where $\mu, \sigma : [0,T] \times \mathbb{R} \to \mathbb{R}$ *are measurable so called* coefficient functions *and* X_0 *is an* $\mathcal{F}_0$*-measurable random variable called the* initial value.

As SDE of diffusion type are by far the most common type of SDE, it is customary to call them just SDE, with the understanding that when other types of SDE are considered we point that out explicitly. This is a custom that also we will adopt.

Definition 9.2. *A* time-homogeneous SDE *(of diffusion type) is given by*

$$dX(t) = \mu(X(t))\, dt + \sigma(X(t))\, dB(t) \quad \textit{for } t \in [0,T], \quad X(0) = X_0,$$

where $\mu, \sigma : \mathbb{R} \to \mathbb{R}$ *are measurable functions and* X_0 *is an* $\mathcal{F}_0$*-measurable random variable.*

In other words, a time-homogeneous SDE is the special case of an SDE with coefficients $\mu(t,x) = \mu(x)$ and $\sigma(t,x) = \sigma(x)$ that do not depend on $t \in [0,T]$.

Exercise 108. Show that the equation $dX(t) = X(t)^2\, dt$ for $t \in [0,T]$, $X(0) = 1$, has no solution for $T \geq 1$.

Definition 9.3. *An* SDE of non-diffusion type *is given by*

$$dX(t) = \mu(t, X(t))\, dt + \sigma(t, X(t))\, dB(t) \quad \textit{for } t \in [0,T], \quad X(0) = X_0, \tag{9.2}$$

where $\{\mu(t,x)\}_{(t,x)\in[0,T]\times\mathbb{R}}$ *and* $\{\sigma(t,x)\}_{(t,x)\in[0,T]\times\mathbb{R}}$ *are stochastic processes and* X_0 *is an* $\mathcal{F}_0$*-measurable random variable.*

A solution to an SDE is any process $\{X(t)\}_{t\in[0,T]}$ that satisfies

$$X(t) = X_0 + \int_0^t \mu(s, X(s))\, ds + \int_0^t \sigma(s, X(s))\, dB(s) \quad \text{for } t \in [0,T]. \tag{9.3}$$

This means that X is an adapted process that is continuous with probability 1 such that $\{\mu(t,X(t))\}_{t\in[0,T]}$ and $\{\sigma(t,X(t))\}_{t\in[0,T]}$ are adapted and measurable stochastic processes that satisfies

$$\mathbf{P}\bigg\{\int_0^T |\mu(t,X(t))|\,dt<\infty\bigg\}=\mathbf{P}\bigg\{\int_0^T \sigma(t,X(t))^2\,dt<\infty\bigg\}=1.$$

For a diffusion type SDE, the adaptedness and measurability of $\{\mu(t,X(t))\}_{t\in[0,T]}$ and $\{\sigma(t,X(t))\}_{t\in[0,T]}$ follow from the adaptedness and continuity of X, as μ and σ are measurable functions, see Exercise 109 below.

Exercise 109. Show that the processes $\{\mu(t,X(t))\}_{t\in[0,T]}$ and $\{\sigma(t,X(t))\}_{t\in[0,T]}$ are adapted and measurable when $\mu,\sigma:[0,T]\times\mathbb{R}\to\mathbb{R}$ are measurable functions and $\{X(t)\}_{t\in[0,T]}$ is an adapted process that is continuous with probability 1.

It turns ot that it is fruitful to consider two kinds of solutions to SDE, the theories of which, perhaps somewhat surprisingly, turns out to be quite different:

Definition 9.4. *A process* $\{X(t)\}_{t\in[0,T]}$ *is a* strong solution *to the SDE* (9.1) *[*(9.2)*] for a given BM B, given coefficients μ and σ, and a given initial value X_0, if X satisfies* (9.3) *for these choices of B, μ and σ with $X(0)=X_0$ with probability 1.*

Definition 9.5. *A process* $\{X(t)\}_{t\in[0,T]}$ *is a* weak solution *to the SDE* (9.1) *if there exists a BM B such that X satisfies* (9.3) *with $X(0)=_D X_0$ (where $=_D$ denotes equality of probability distributions).*

Definition 9.6. *Solutions to the SDE* (9.1) *have* strong uniqueness[4] *if whenever* $\{X_1(t)\}_{t\in[0,T]}$ *and* $\{X_2(t)\}_{t\in[0,T]}$ *are strong solutions for a common given BM B and a common given initial value X_0, it holds that*

$$\mathbf{P}\big\{X_1(t)=X_2(t)\ \textit{for all}\ t\in[0,T]\big\}=1. \tag{9.4}$$

Note that if $\{X_1(t)\}_{t\in[0,T]}$ and $\{X_2(t)\}_{t\in[0,T]}$ are strong solutions to (9.1), then as X_1 and X_2 are continuous with probability 1 (being Itô processes), Exercise 48 shows that (9.4) holds if and only if X_1 and X_2 are versions of each other.

Definition 9.7. *Solutions to the SDE* (9.1) *have* weak uniqueness *if whenever* $\{X_1(t)\}_{t\in[0,T]}$ *and* $\{X_2(t)\}_{t\in[0,T]}$ *are weak solutions they have common fidi's.*

[4] In fact our definition of strong uniqueness is what some authors call *pathwise uniqueness*.

Note that the above uniqueness concepts do not require existence of solutions (albeit uniqueness in the absence of existence argubly becomes a rather philosophical property), as is illustrated by the following exercise:

Exercise 110. Show that SDE which lack solutions have uniquness of solutions.

9.2 Stochastic exponential

We now return to a more complete treatment of the SDE considered in Example 8.1

$$dX(t) = \sigma\, X(t)\, dB(t) \quad \text{for } t \in [0,T], \quad X(0) = X_0 > 0.$$

We generalize this equation to the (not necessarily diffusion type) SDE

$$dX(t) = \sigma(t)\, X(t)\, dB(t) \quad \text{for } t \in [0,T], \quad X(0) = X_0 > 0, \tag{9.5}$$

where $\sigma \in P_T$. By application of Itô's formula to $Y(t) = \ln(X(t))$ we get

$$dY(t) = \frac{dX(t)}{X(t)} - \frac{d[X](t)}{2\,X(t)^2} = \sigma(t)\, dB(t) - \frac{1}{2}\,\sigma(t)^2\, dt$$

(recall Theorem 8.6), so that [noting that $Y(0) = \ln(X_0)$]

$$Y(t) = \ln(X_0) + \int_0^t \sigma\, dB - \frac{1}{2}\int_0^t \sigma(r)^2\, dr \quad \text{for } t \in [0,T].$$

As the function ln is not really two times continuously differentiable on the whole of $\mathbb{R}$, we must check that our solutiuon really works. To that end we note that an application of Itô's formula to $X(t) = \mathrm{e}^{Y(t)}$ with $Y(t)$ as above gives

$$dX(t) = \mathrm{e}^{Y(t)}\, dY(t) + \frac{1}{2}\,\mathrm{e}^{Y(t)}\, d[Y](t) = \ldots = \sigma(t)\, X(t)\, dB(t).$$

And so

$$X(t) = \mathrm{e}^{Y(t)} = X_0 \exp\biggl\{\int_0^t \sigma\, dB - \frac{1}{2}\int_0^t \sigma(r)^2\, dr\biggr\} \quad \text{for } t \in [0,T]$$

is a strong solution to the SDE (9.5).

In fact, Example 8.1 may be generalized even further than to (9.5) as follows:

Definition 9.8. *The* stochastic exponential $\{\mathcal{E}(X)(t)\}_{t\in[0,T]}$ *of an Itô process* $\{X(t)\}_{t\in[0,T]}$ *is a solution to the (not necessarily diffusion type) SDE*

$$d\mathcal{E}(X)(t) = \mathcal{E}(X)(t)\, dX(t) \quad \textit{for } t \in [0,T], \quad \mathcal{E}X(0) = 1. \tag{9.6}$$

Theorem 9.9. *The stochastic exponential of an Itô process* $\{X(t)\}_{t\in[0,T]}$ *is given by*

$$\mathcal{E}(X)(t) = \mathrm{e}^{X(t) - X(0) - \frac{1}{2}[X](t)} \quad \textit{for } t \in [0,T]. \tag{9.7}$$

Exercise 111. Prove that the process $\mathcal{E}(X)$ given by (9.7) is a strong solution to the equation (9.6).

9.3 Langevin equation and Ornstein-Uhlenbeck process

The so called *Langevin equation* is the (not necessarily diffusion type) SDE given by

$$dX(t) = -\beta(t)\,X(t)\,dt + \sigma(t)\,dB(t) \quad \text{for } t \in [0,T], \quad X(0) = X_0, \tag{9.8}$$

where $\{\beta(t)\}_{t\in[0,T]}$ and $\{\sigma(t)\}_{t\in[0,T]}$ are measurable and adpated stochastic processes such that

$$\mathbf{P}\Big\{\int_0^T |\beta(t)|\,dt < \infty\Big\} = \mathbf{P}\Big\{\int_0^T \sigma(t)^2\,dt < \infty\Big\} = 1.$$

We will need a new Itô formula for functions $f(t, X(t))$ of time t and an Itô process $\{X(t)\}_{t\in[0,T]}$ where $f : [0,T]\times\mathbb{R} \to \mathbb{R}$ is continuously differentiable in its first argument and two times continuously differentiable in its second argument. That formula is given in Theorem 9.10 below and says that

$$df(X(t)) = f'_t(t,X(t))\,dt + f'_x(t,X(t))\,dX(t) + \frac{1}{2}\,f''_{xx}(t,X(t))\,d[X](t). \tag{9.9}$$

To solve (9.8) we note that $Y(t) = \mathrm{e}^{\int_0^t \beta(s)\,ds}\,X(t)$ by Itô's formula (9.9) satisfies

$$dY(t) = \beta(t)\,Y(t)\,dt + \mathrm{e}^{\int_0^t \beta(s)\,ds}\,dX(t) = \ldots = \mathrm{e}^{\int_0^t \beta(s)\,ds}\,\sigma(t)\,dB(t).$$

It follows that

$$Y(t) = Y(0) + \int_0^t \mathrm{e}^{\int_0^s \beta(r)\,dr}\,\sigma(s)\,dB(s) = X_0 + \int_0^t \mathrm{e}^{\int_0^s \beta(r)\,dr}\,\sigma(s)\,dB(s) \quad \text{for } t \in [0,T].$$

Transforming back we conclude that the equation (9.8) has strong solution

$$X(t) = \mathrm{e}^{-\int_0^t \beta(s)\,ds}\Big(X_0 + \int_0^t \mathrm{e}^{\int_0^s \beta(r)\,dr}\,\sigma(s)\,dB(s)\Big) \quad \text{for } t \in [0,T]. \tag{9.10}$$

In our application of the Itô formula (9.9) above, the function $f(t,x) = \mathrm{e}^{\int_0^t \beta(s)\,ds}x$ is a stochastic process, which is not permitted in (9.9). Further, f need not be continuously differentiable in its first argument (without additional regularity assumptions for the β process). Hence we have not really proved that the Itô process X in (9.10) satisfies (9.8). To check that this is the case we need yet another Itô formula given in Theorem 9.12 below, which asserts that for two Itô processes X and Y and a two times continuously differentiable (in both arguments) function $f : \mathbb{R}^2 \to \mathbb{R}$, we have

$$\begin{aligned} &df(X,Y) \\ &= f'_x(X,Y)\,dX + f'_y(X,Y)\,dY + \frac{1}{2}f''_{xx}(X,Y)\,d[X] + \frac{1}{2}f''_{yy}(X,Y)\,d[Y] + f''_{xy}(X,Y)\,d[X,Y]. \end{aligned} \tag{9.11}$$

Exercise 112. Use the Itô formula (9.11) to check that the Itô process X in (9.10) really satisfies (9.8).

In the particular case when the coefficient processes of the Langevin equation (9.8) are constants $\beta(t) = \beta > 0$ and $\sigma(t) = \sigma > 0$, the solution

$$X(t) = \mathrm{e}^{-\beta t} X_0 + \sigma\, \mathrm{e}^{-\beta t} \int_0^t \mathrm{e}^{\beta r}\, dB(r) \quad \text{for } t \in [0,T] \tag{9.12}$$

is called an *Ornstein-Uhlenbeck process*. If we take $X_0 = 0$ it follows from Exercise 58 that the Ornstein-Uhlenbeck process is zero-mean Gaussian with covariance function

$$\mathbf{Cov}\{X(s), X(t)\} = \sigma^2 \mathrm{e}^{-\beta(s+t)} \int_0^{\min\{s,t\}} \mathrm{e}^{2\beta s}\, ds = \ldots = \frac{\sigma^2}{2\beta}\big(\mathrm{e}^{-\beta|t-s|} - \mathrm{e}^{-\beta(s+t)}\big).$$

Whatever you do, don't miss the next exercise:

Exercise 113. Consider the so called *Euler (Euler-Maruyama)* numerical approximation scheme of the Ornstein-Uhlenbeck process (9.12) given by $X(t) = X(t_i)$ for $t \in [t_i, t_{i+1})$, where

$$X(t_{i+1}) = X(t_i) - \beta\, X(t_i)\,(t_{i+1} - t_i) + \sigma\,(B(t_{i+1}) - B(t_i)) \quad \text{for } i = 0, 1, \ldots\ .$$

Take $t_i = T\,(i/n)$ for $i = 0, \ldots, n$. Show by means of direct calculations that the Euler scheme converges in mean-square to the solution (9.12) of (9.8) as $n \to \infty$.

9.4 General linear equation

The stochastic exponential equation for BM (9.5) and the Langevin equation (9.8) both are special cases of the so called *general linear equation* given by

$$dX(t) = (\alpha(t) + \beta(t)\, X(t))\, dt + (\gamma(t) + \delta(t)\, X(t))\, dB(t) \quad \text{for } t \in [0,T], \quad X(0) = X_0, \tag{9.13}$$

where $\{\alpha(t)\}_{t\in[0,T]}$, $\{\beta(t)\}_{t\in[0,T]}$, $\{\gamma(t)\}_{t\in[0,T]}$ and $\{\delta(t)\}_{t\in[0,T]}$ are measurable and adapted stochastic processes that satisfy suitable integrability conditions. It turns out that also this equation has a strong solution that can be given in explicit form and that can be derived in the same fashion as we dealt with the equations (9.5) and (9.8). The details are left to the very interested reader.

Exercise 114. Solve the general linear equation (9.13).

9.5 Itô formula

The time has come to discuss the two new Itô formulas that we employed in the study of the Langevin equation in Section 9.3.

Theorem 9.10. (ITÔ FORMULA) *For an Itô process* $\{X(t)\}_{t\in[0,T]}$ *and a function* $f\in C^{1,2}([0,T]\times\mathbb{R})$[5]*, we have*

$$df(t,X(t)) = f_t'(t,X(t))\,dt + f_x'(t,X(t))\,dX(t) + \frac{1}{2}\,f_{xx}''(t,X(t))\,d[X](t). \qquad (9.14)$$

We defer the proof of Theorem 9.10 to Section 11.2 below.

Example 9.11. If X is an Itô processes and if $f(t,x)=g(t)\,x$ where $g:[0,T]\to\mathbb{R}$ is continuously differentiable, then Theorem 9.10 gives

$$df(t,X(t)) = g'(t)\,X(t)\,dt + g(t)\,dX(t). \qquad (9.15)$$

We used (9.15) in Section 9.3 with $g(t)=\mathrm{e}^{\int_0^t \beta(s)\,ds}$ to derive (9.10). Note that this usage (9.15) is rigorous only if $\beta:\mathbb{R}\to\mathbb{R}$ is a continuous (non-random) function.

Here is yet another useful Itô formula:

Theorem 9.12. (ITÔ FORMULA) *For two Itô processes* X *and* Y *and a function* $f\in C^2(\mathbb{R}^2)$[6]*, we have*

$$\begin{aligned} &df(X,Y)\\ &= f_x'(X,Y)\,dX + f_y'(X,Y)\,dY + \frac{1}{2}f_{xx}''(X,Y)\,d[X] + \frac{1}{2}f_{yy}''(X,Y)\,d[Y] + f_{xy}''(X,Y)\,d[X,Y]. \end{aligned}$$

Exercise 115. Prove Theorem 9.12.

Exercise 116. Explain how Theorem 9.10 in a way (but not in full generality) follows from Theorem 9.12. (See also the proof of Theorem 9.10 in Section 11.2.)

Corollary 9.13. (INTEGRATION BY PARTS) *For Itô processes* X *and* Y *we have*

$$d(X(t)Y(t)) = X(t)\,dY(t) + Y(t)\,dX(t) + d[X,Y](t).$$

Proof. Take $f(x,y)=xy$ in Theorem 9.12. □

Example 9.14. For an Itô process $\{X(t)\}_{t\in[0,T]}$ we have

$$\int_0^t X\,dX = \frac{1}{2}\big(X(t)^2 - [X](t)\big) \quad \text{for } t\in[0,T].$$

---

[5] The class of functions $f:[0,T]\times\mathbb{R}\to\mathbb{R}$ that are continuously differentiable in the first argument and two times continuously differentiable in the second argument.

[6] The class of functions $f:\mathbb{R}^2\to\mathbb{R}$ for which all second order derivatives exist and are continuous.

10 Lecture 10, Monday May 11

10.1 Paul Lévy's characterization of BM

The implication to the right of the following famous result we encountered already in Theorem 1.10 during the first lecture. It has been a vital tool in our build up of the theory. However, the converse implication to the left, which we are now equipped to prove, is almost as important as that to the right:

Theorem 10.1. (PAUL LÉVY'S CHARACTERIZATION OF BM) *An Itô integral process $\{X(t)\}_{t\in[0,T]}$ is BM if and only if it has quadratic variation process $[X](t)=t$ for $t\in[0,T]$ with probability 1.*

Proof. $\boxed{\Leftarrow}$ Consider an Itô integral process $X(t)=\int_0^t \sigma\,dB$ for $t\in[0,T]$, for some $\sigma\in P_T$, such that $[X](t)=t$ for $t\in[0,T]$ with probability 1. From Theorem 8.6 it follows that, given any $\theta\in\mathbb{R}$, the processes

$$\sigma_1(t)=\sigma(t)\,\cos(\theta X(t))\,\mathrm{e}^{\frac{1}{2}\theta^2 t} \quad\text{and}\quad \sigma_2(t)=\sigma(t)\,\sin(\theta X(t))\,\mathrm{e}^{\frac{1}{2}\theta^2 t} \quad\text{for}\ \ t\in[0,T]$$

belong to E_T, since

$$\int_0^T \big(\sigma_1(t)^2+\sigma_2(t)^2\big)\,dt=\int_0^T \sigma(t)^2\,\mathrm{e}^{\theta^2 t}\,dt=\int_0^T \mathrm{e}^{\theta^2 t}\,d[X](t)=\int_0^T \mathrm{e}^{\theta^2 t}\,dt\le T\,\mathrm{e}^{\theta^2 T}.$$

Hence the Itô integral processes

$$Y_1(t)=\int_0^t \sigma_1\,dB \quad\text{and}\quad Y_2(t)=\int_0^t \sigma_2\,dB \quad\text{for}\ \ t\in[0,T]$$

are martingales. Next note that the processes

$$Z_1(t)=\cos(\theta X(t))\,\mathrm{e}^{\frac{1}{2}\theta^2 t} \quad\text{and}\quad Z_2(t)=\sin(\theta X(t))\,\mathrm{e}^{\frac{1}{2}\theta^2 t} \quad\text{for}\ \ t\in[0,T]$$

are Itô processes, which by Itô's formula Theorem 9.10 satisfy

$$\begin{aligned} dZ_1(t)&=\frac{\theta^2 Z_1(t)\,dt}{2}-\theta\,Z_2(t)\,dX(t)-\frac{\theta^2 Z_1(t)\,d[X](t)}{2}=-\theta\,\sigma_2(t)\,dB(t)=-\theta\,dY_2(t),\\ dZ_2(t)&=\frac{\theta^2 Z_2(t)\,dt}{2}+\theta\,Z_1(t)\,dX(t)-\frac{\theta^2 Z_2(t)\,d[X](t)}{2}=\ \ \theta\,\sigma_1(t)\,dB(t)=\ \ \theta\,dY_1(t).\end{aligned}$$

Hence Z_1 and Z_2 are martingales (as Y_1 and Y_2 are martingales). It follows that

$$\mathbf{E}\big\{\mathrm{e}^{i\theta X(t)+\frac{1}{2}\theta^2 t}\big|\mathcal{F}_s\big\}=\mathbf{E}\big\{Z_1(t)\big|\mathcal{F}_s\big\}+i\,\mathbf{E}\big\{Z_2(t)\big|\mathcal{F}_s\big\}=Z_1(s)+i\,Z_2(s)=\mathrm{e}^{i\theta X(s)+\frac{1}{2}\theta^2 s}$$

for $0\le s\le t$ and $\theta\in\mathbb{R}$, which in turn (as X is adapted) by rearrangement gives

$$\mathbf{E}\big\{\mathrm{e}^{i\theta(X(t)-X(s))}\big|\mathcal{F}_s\big\}=\mathrm{e}^{-\frac{1}{2}\theta^2(t-s)} \quad\text{and}\quad \mathbf{E}\big\{\mathrm{e}^{i\theta(X(t)-X(s))}\big\}=\mathrm{e}^{-\frac{1}{2}\theta^2(t-s)} \qquad (10.1)$$

for $0\le s\le t$ and $\theta\in\mathbb{R}$. In particular, $X(t)-X(s)$ is $\mathrm{N}(0,t-s)$-distributed. Further,

$X(t)-X(s)$ is independent of $\mathcal{F}_s$ for $0\leq s\leq t$, as for any event $A\in\mathcal{F}_s$ we get

$$\mathbf{E}\big\{e^{i\theta(X(t)-X(s))+i\varphi I_A}\big\}=\mathbf{E}\big\{\big\{e^{i\theta(X(t)-X(s))+i\varphi I_A}\big|\mathcal{F}_s\big\}e^{i\varphi I_A}\big\}=e^{-\frac{1}{2}\theta^2(t-s)}\,\mathbf{E}\big\{e^{i\varphi I_A}\big\} \tag{10.2}$$

for $\theta,\varphi\in\mathbb{R}$. In particular, this shows that X has independent increments, see Exercise 117 below. As $X(0)=0$ and X is continuous, we have shown that X is BM. □

Exercise 117. Explain why (10.2) implies that X has independent increments.

10.2 Tanaka's equation

The following rather famous example is intended to give a perspective on the topic of existence and uniqueness of solutions to SDE: Let $\text{sign}\,(x)=I_{[0,\infty)}(x)-I_{(-\infty,0)}(x)$ for $x\in\mathbb{R}$ and consider the so called *Tanaka equation*

$$dX(t)=\text{sign}\,(X(t))\,dB(t)\quad\text{for }\ t\in[0,T],\quad X(0)=0. \tag{10.3}$$

As Theorem 8.6 shows that any solution to (10.3) satisfies

$$[X](t)=\int_0^t\text{sign}\,(X(r))^2\,dr=\int_0^t dr=t\quad\text{for }\ t\in[0,T]\ \ \text{(with probability 1)},$$

any solution to (10.3) is BM by Theorem 10.1. Hence we have weak uniqueness.

Further, setting $dW(t)=\text{sign}\,(B(t))\,dB(t)$ we see as above from Theorem 8.6 that $[W](t)=t$. Hence W is also BM by Theorem 10.1. As we have

$$dB(t)=\text{sign}\,(B(t))^2\,dB(t)=\text{sign}\,(B(t))\,dW(t)\quad\text{for }\ t\in[0,T],\quad B(0)=0,$$

it follows that any given BM B is a weak solution to (10.3) when we select W as given above to be the driving BM of the SDE. Hence we have weak existence.

It can be shown that (10.3) does not have a strong solution, but that requires quite difficult rather special new techniques which we do not want to spend time on[7].

However, it easy to see that if (10.3) has a strong solution X, then that solution is not unique in the strong sense. This is so because $Y=-X$ satisfies

$$dY=d(-X)=-dX=-\text{sign}\,(X)\,dB=\text{sign}\,(-X)\,dB=\text{sign}\,(Y)\,dB,\quad Y(0)=0.$$

Here the equality $-\text{sign}\,(X)\,dB=\text{sign}\,(-X)\,dB$ is not obvious, as $-\text{sign}\,(x)=\text{sign}\,(-x)$ holds only for $x\neq 0$. However, note that $-\text{sign}\,(X)\,dB=\text{sign}\,(-X)\,dB$ means that $\int_0^t-\text{sign}\,(X)\,dB=\int_0^t\text{sign}\,(-X)\,dB$ for $t\in[0,T]$, which in turn holds if $-\text{sign}\,(X(t))=\text{sign}\,(-X(t))$ a.e. for $t\in[0,T]$ with probability 1. That this is so in

[7] The proof requires Tanaka's formula – a special case of a more general extension of Itô's formula to convex functions, see e.g., Karatzas and Shreve: "Brownian Motion and Stochastic Calculus", Example 5.3.5 – according to which $|X|(t)-|X|(0)=\int_0^t\text{sign}\,(X)\,dX+\lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\int_0^t I_{[0,\varepsilon]}(|X|(r))\,dr$ for an Itô process X. If that Itô process solves (10.3) it follows that $|X|(t)=\int_0^t\text{sign}\,(X)^2\,dB+\lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\int_0^t I_{[0,\varepsilon]}(|X|(r))\,dr$, so that $B(t)=|X(t)|-\lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\int_0^t I_{[0,\varepsilon]}(|X|(r))\,dr$. This in turn implies that $\sigma(X(s):s\leq t)\subseteq\sigma(B(s):s\leq t)\subseteq\sigma(|X(s)|:s\leq t)$, which cannot be true as X is BM.

turn follows from the fact that X is BM, see Exercise 118 below.

Exercise 118. Show that BM B satisfies $\mathbf{P}\{B(t)\neq 0 \text{ a.e. for } t\geq 0\}=1$.

Example 10.2. The non-random ODE version of the Tanaka SDE

$$dX(t)=\operatorname{sign}(X(t))\,dt \quad \text{for } t\in[0,T], \quad X(0)=0, \tag{10.4}$$

has the strong solution $X(t)=t$, see also Exercise 119 below.

Exercise 119. Explain why the solution $X(t)=t$ to (10.4) is unique in the ordinary ODE sense, that is, when we require $X'(t)=\operatorname{sign}(X(t))$ for all $t\in[0,T]$.

It is not always possible to carry over intuition from ODE to SDE and vice versa: Solutions to some SDE simply behave more or less like randomly perturbed versions of solutions to the corresponding ODE (with σ coefficient zero), as is argubly the case, for example, with the stochastic exponential equation (9.6) and its solution (9.7). On the other hand, some other SDE does not behave like randomly perturbed versions of the corresponding ODE (with $\sigma=0$) at all, as is argubly the case, for example, with the Langevin equation (9.8) and its solution (9.10).

10.3 Strong uniqueness

Although Tanaka's SDE does not have a unique strong solution if a strong solution exists, as a strong solution to this SDE does not exist it is interesting to find an example of an SDE that has a strong solutions that is not unique (recall Exercise 110):

Exercise 120. Show that the ODE version of Tanaka's SDE (10.4) does not have a unique strong solution in the SDE sense, that is, when we only require $X(t)=\int_0^t \operatorname{sign}(X(s))\,ds$ for all $t\in[0,T]$, as both $X(t)=t$ and $X(t)=-t$ solve this SDE.

We are now going to state and prove a standard type of uniqueness result. To that end we need the following simple inequality:

Theorem 10.3. (Grönwall's lemma) *If $u,v:[0,T]\to[0,\infty)$ are continuous functions such that*

$$v(t)\leq C+\int_0^t u(r)v(r)\,dr \quad \textit{for } t\in[0,T],$$

for some constant $C\geq 0$, then we have

$$v(t)\leq C+\int_0^t u(r)v(r)\,dr\leq C\exp\left\{\int_0^t u(r)\,dr\right\} \quad \textit{for } t\in[0,T].$$

Proof. Given any constant $\varepsilon > 0$, we have

$$v(t) \le (C+\varepsilon) + \int_0^t u(r)v(r)\,dr \equiv V_\varepsilon(t) \quad \text{for } t \in [0,T].$$

Since V_ε is strictly positive it follows that

$$\frac{d}{dt}\,(\ln(V_\varepsilon(t)) = \frac{V_\varepsilon'(t)}{V_\varepsilon(t)} = \frac{u(t)v(t)}{V_\varepsilon(t)} \le u(t) \quad \text{for } t \in (0,T).$$

This in turn gives

$$\ln(V_\varepsilon(t)) = \ln(V_\varepsilon(0)) + \int_0^t \frac{d}{dr}\ln(V_\varepsilon(r)\,dr \le \ln(C+\varepsilon) + \int_0^t u(r)\,dr \quad \text{for } t \in [0,T],$$

so that

$$v(t) \le V_\varepsilon(t) \le \exp\biggl\{\ln(C+\varepsilon) + \int_0^t u(r)\,dr\biggr\} = (C+\varepsilon)\exp\biggl\{\int_0^t u(r)\,dr\biggr\} \quad \text{for } t \in [0,T].$$

□

Definition 10.4. *The coefficients $\mu, \sigma : [0,T]\times\mathbb{R} \to \mathbb{R}$ of the diffusion type SDE* (9.1) *are said to satisfy a* local Lipschitz condition *if to each $n \in \mathbb{N}$ there exists a constant $K_n > 0$ such that*

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| \le K_n\,|x-y| \quad \textit{for } t \in [0,T] \textit{ and } x,y \in [-n,n].$$

Theorem 10.5. *If the coefficients of the diffusion type SDE* (9.1) *satisfy a local Lipschitz condition, then we have strong uniqueness for solutions to the SDE (but not necessarily existence).*

See Theorem 11.3 below for a useful improvement of the above uniqueness result.

Proof of Theorem 10.5. Consider two strong solutions X_1 and X_2 to (9.1). To prove uniqeness it is enough to prove that

$$\mathbf{P}\{X_1(t) = X_2(t)\} = 1 \quad \text{for each } t \in [0,T]$$

(recall Exercise 48). Consider the stopping time

$$\tau_n = \inf\bigl\{t \in [0,T] : |X_1(t)| \ge n\bigr\} \wedge \inf\bigl\{t \in [0,T] : |X_2(t)| \ge n\bigr\}.$$

Then the process $X_i^{(n)}(t) = X_i(t \wedge \tau_n)$ satisfies $|X_i^{(n)}(t)| \le n$ for $t \in [0,T]$ by the continuity of X_i for $i = 1,2$. Further, we have

$$X_i^{(n)}(t) = \int_0^t \mu(r, X_i(r))\, I_{[0,\tau_n]}(r)\,dr + \int_0^t \sigma(\,\cdot\,, X_i)\, I_{[0,\tau_n]}\,dB \quad \text{for } t \in [0,T],$$

for $i=1,2$ by Theorem 6.5. Since $\tau_n\uparrow\infty$ as $n\to\infty$ by the continuity of X_1 and X_2, it suffices to prove that $\mathbf{P}\{X_1^{(n)}(t)=X_2^{(n)}(t)\}=1$ for $t\in[0,T]$ and $n\in\mathbb{N}$, see Exercise 121 below. However, by the inequality $(x+y)^2\le 2x^2+2y^2$ and isometry together with the Cauchy-Schwarz inequality and the local Lipschitz condition, we have

$$\begin{aligned}
&\mathbf{E}\big\{(X_1^{(n)}(t)-X_2^{(n)}(t))^2\big\}\\
&\quad\le 2\,\mathbf{E}\bigg\{\bigg(\int_0^t \big(\mu(r,X_1(r))-\mu(r,X_2(r))\big)I_{[0,\tau_n]}(r)\,dr\bigg)^2\bigg\}\\
&\qquad+2\,\mathbf{E}\bigg\{\bigg(\int_0^t \big(\sigma(\,\cdot\,,X_1)-\sigma(\,\cdot\,,X_2)\big)I_{[0,\tau_n]}\,dB\bigg)^2\bigg\}\\
&\quad\le 2\,\mathbf{E}\bigg\{\bigg(\int_0^t dr\bigg)\bigg(\int_0^t \big(\mu(r,X_1(r))-\mu(r,X_2(r))\big)^2 I_{[0,\tau_n]}(r)\,dr\bigg)\bigg\}\\
&\qquad+2\,\mathbf{E}\bigg\{\int_0^t \big(\sigma(r,X_1(r))-\sigma(r,X_2(r))\big)^2 I_{[0,\tau_n]}(r)\,dr\bigg\}\\
&\quad\le 2\,(T+1)\,K_n^2\int_0^t \mathbf{E}\big\{(X_1^{(n)}(r)-X_2^{(n)}(r))^2\big\}\,dr \qquad \text{for } t\in[0,T].
\end{aligned}$$

Hence an application of Grönwall's lemma with $C=0$, $u(t)=2\,(T+1)\,K_n^2$ and $v(t)=\mathbf{E}\{(X_1^{(n)}(t)-X_2^{(n)}(t))^2\}$ (which is continuous by continuity of $X_1^{(n)}$ and $X_2^{(n)}$ together with the bounded convergence theorem) gives $v(t)=0$. □

Exercise 121. Why is it sufficient to prove that $\mathbf{P}\{X_1^{(n)}(t)=X_2^{(n)}(t)\}=1$ for $t\in[0,T]$ and $n\in\mathbb{N}$ in the proof of Theorem 10.5.

Remark 10.6. Recall from Exercise 108 that the equation (10.7) does not have a strong solution if $T\ge 1$ although it satisfies a local Lipschitz condition, which in turn illustrates the fact that it is only uniqueness that is adressed in Theorem 10.5, but not existence. See also Example 12.1 below for more information.

Example 10.7. The following equation has uniqueness for strong solutions

$$dX(t)=X(t)^2\,dt \quad \text{for } t\in[0,T], \qquad X(0)=1,$$

by Theorem 10.5, since the drift $\mu(x)=x^2$ satisfies the local Lipschitz condition

$$|x^2-y^2|^2=|x+y|\,|x-y|\le 2\,n\,|x-y| \quad \text{for } |x|,|y|\le n.$$

Exercise 122. Show that a genereal linear SDE of the type

$$dX(t)=(\alpha+\beta\,X(t))\,dt+(\gamma+\delta\,X(t))\,dB(t)$$

where $\alpha,\beta,\gamma,\delta\in\mathbb{R}$ are real constants have strong uniqueness.

Exercise 123. Does strong uniqueness trivially imply weak uniqueness? Does weak uniqueness trivially imply strong uniqueness?

11 Lecture 11, Wednesday May 13

11.1 Conclusion of the proof of Theorem 8.6

It remains to extend the range of the theorem from S_T to E_T. To that end, take $\sigma \in E_T$ and use Theorem 4.4 to find a sequence $\{\sigma_m\}_{m=1}^{\infty} \subseteq S_T$ such that

$$\lim_{m\to\infty} \mathbf{E}\bigg\{\int_0^T \big(\sigma(r)-\sigma_m(r)\big)^2\,dr\bigg\} = 0. \tag{11.1}$$

By the elementary inequality $(x+y)^2 \leq (1+\varepsilon)\,x^2 + (1+1/\varepsilon)\,y^2$ for $x,y\in\mathbb{R}$ and $\varepsilon>0$ together with Boole's inequality, we have, given any constants $\delta,\varepsilon>0$ and $m\in\mathbb{N}$,

$$\begin{aligned}
&\mathbf{P}\bigg\{\sum_{i=1}^n (X(t_i)-X(t_{i-1}))^2 - \int_0^T \sigma(r)^2\,dr > \delta\bigg\}\\
&\leq \mathbf{P}\bigg\{(1+\varepsilon)\sum_{i=1}^n\Big(\int_{t_{i-1}}^{t_i}\sigma_m\,dB\Big)^2 + (1+1/\varepsilon)\sum_{i=1}^n\Big(\int_{t_{i-1}}^{t_i}(\sigma-\sigma_m)\,dB\Big)^2 - \int_0^T\sigma(r)^2\,dr > \delta\bigg\}\\
&\leq \mathbf{P}\bigg\{(1+\varepsilon)\sum_{i=1}^n\Big(\int_{t_{i-1}}^{t_i}\sigma_m\,dB\Big)^2 - (1+\varepsilon)\int_0^T\sigma_m(r)^2\,dr > \frac{\delta}{4}\bigg\}\\
&\qquad + \mathbf{P}\bigg\{(1+1/\varepsilon)\sum_{i=1}^n\Big(\int_{t_{i-1}}^{t_i}(\sigma-\sigma_m)\,dB\Big)^2 > \frac{\delta}{4}\bigg\}\\
&\qquad + \mathbf{P}\bigg\{(1+\varepsilon)\int_0^T\big(\sigma_m(r)^2-\sigma(r)^2\big)\,dr > \frac{\delta}{4}\bigg\}\\
&\qquad + \mathbf{P}\bigg\{\varepsilon\int_0^T\sigma(r)^2\,dr > \frac{\delta}{4}\bigg\}.
\end{aligned} \tag{11.2}$$

Here the first term on the right-hand side goes to zero as $\max_{1\leq i\leq n} t_i - t_{i-1} \downarrow 0$ by application of what we have proved already for $\sigma_m \in S_T$. By Tjebysjev's inequality and isometry, the second term on the right-hand side of (11.2) is bounded by

$$\frac{16\,(1+1/\varepsilon)^2}{\delta^2}\,\mathbf{E}\bigg\{\int_0^T\big(\sigma(r)-\sigma_m(r)\big)^2\,dr\bigg\},$$

which goes to 0 as $m\to\infty$ by (11.1). As $\sigma_m^2-\sigma^2 = (\sigma-\sigma_m)^2 + 2\,(\sigma_m-\sigma)\,\sigma$, Markov's inequality together with the Cauchy-Schwarz inequality show that the third term on the right-hand side of (11.2) is bounded by

$$\frac{4\,(1+\varepsilon)}{\delta}\Bigg(\mathbf{E}\bigg\{\int_0^T(\sigma(r)-\sigma_m(r))^2\,dr\bigg\} + 2\sqrt{\mathbf{E}\bigg\{\int_0^T(\sigma_m(r)-\sigma(r))^2\,dr\bigg\}\,\mathbf{E}\bigg\{\int_0^T\sigma(r)^2\,dr\bigg\}}\Bigg),$$

which goes to 0 as $m\to\infty$ by (11.1). Finally, by Markov's inequality the fourth term on the right-hand side of (11.2) is bounded by

$$\frac{4\,\varepsilon}{\delta}\,\mathbf{E}\bigg\{\int_0^T\sigma(r)^2\,dr\bigg\}.$$

Hence, given any $\delta,\epsilon>0$, we may first select $\varepsilon>0$ sufficiently small to make the fourth term on the right-hand side smaller than $\epsilon/3$. With that choice of $\varepsilon > 0$ we may then select $m \in \mathbb{N}$ sufficiently large to make each of the second term and the third term on the right-hand side smaller than $\epsilon/3$. It follows that the limit (superior) as $\max_{1\le i\le n} t_i - t_{i-1} \downarrow 0$ of the left-hand side of (11.2) is at most ϵ for any $\epsilon > 0$, so that the limit must be zero. Using the elementary inequality $(x+y)^2 \ge (1-\varepsilon)\, x^2 - (1/\varepsilon - 1)\, y^2$ for $x, y \in \mathbb{R}$ and $\varepsilon \in (0,1)$, we can in an entirely similar way show that

$$\mathbf{P}\bigg\{\sum_{i=1}^n (X(t_i) - X(t_{i-1}))^2 - \int_0^T \sigma(r)^2\, dr < -\delta\bigg\} \to 0 \quad \text{as} \quad \max_{1\le i\le n} t_i - t_{i-1} \downarrow 0. \qquad \square$$

Exercise 124. Prove Theorem 8.10.

11.2 Itô formula

Proof of Theorem 9.10. Consider partitions $0 = t_0 < t_1 < \ldots < t_n = t$ of the interval $[0,t]$ such that $\max_{1\le i\le n} t_i - t_{i-1} \downarrow 0$. By Taylor expansion with remainder, we have

$$\begin{aligned}
&f(t, X(t)) - f(0, X(0)) \\
&= \sum_{i=1}^n \big(f(t_i, X(t_i)) - f(t_{i-1}, X(t_i))\big) + \sum_{i=1}^n \big(f(t_{i-1}, X(t_i)) - f(t_{i-1}, X(t_{i-1}))\big) \\
&= \sum_{i=1}^n f_t'(t_i, X(t_i))\, (t_i - t_{i-1}) \\
&\qquad\quad + \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \big(f_t'(s, X(t_i)) - f_t'(t_i, X(t_i))\big)\, ds \\
&\qquad + \sum_{i=1}^n f_x'(t_{i-1}, X(t_{i-1}))\, (X(t_i) - X(t_{i-1})) \\
&\qquad\quad + \frac{1}{2} \sum_{i=1}^n f_{xx}''(t_{i-1}, X(t_{i-1}))\, (X(t_i) - X(t_{i-1}))^2 \\
&\qquad\quad + \sum_{i=1}^n \int_{X(t_{i-1})}^{X(t_i)} (X(t_i) - s)\, (f''(t_{i-1}, s) - f''(t_{i-1}, X(t_{i-1})))\, ds \\
&\to \int_0^t f_t'(s, X(s))\, ds + 0 + \int_0^t f_x'(s, X(s))\, dX(s) + \frac{1}{2} \int_0^t f_{xx}''(s, X(s))\, d[X](s) + 0
\end{aligned} \tag{11.3}$$

in probability as $\max_{1\le i\le n} t_i - t_{i-1} \downarrow 0$ by more or less precisely the same reasons that we invoked to complete the proofs of Theorems 7.5 and 7.6. $\square$

We have already seen four different Itô formulas. The requirement in Itô's formula that the function we apply to an Itô process has to be two times continuously differentiable over the whole world of values of the Itô process can in fact be weakened to exactly that requirement, resulting in the following improved versions of our previously considered Itô formulas in Theorems 9.10 and 9.12:

Theorem 11.1. (ITÔ FORMULA) *For an Itô process X all values of which belong to an open interval $I\subseteq\mathbb{R}$ with probability 1 and a function $f\in C^{1,2}([0,T]\times I)$*[8]*, we have*

$$df(t,X(t)) = f_t'(t,X(t))\,dt + f_x'(t,X(t))\,dX(t) + \frac{1}{2}\,f_{xx}''(t,X(t))\,d[X](t).$$

Proof. Let X have stochastic differential $dX(t)=\mu(t)\,dt+\sigma(t)\,dB(t)$. Consider the stopped process $X^{(n)}(t)=X(t\wedge\tau_n)$ for $t\in[0,T]$, where (with obvious notation)

$$\tau_n = \inf\big\{t\in[0,T] : |X(t)-I^c|\leq 1/n\big\}.$$

Then the almost sure continuity of X gives $|X^{(n)}(t)-I^c|\geq 1/n$ for $t\in[0,T]$ with probability 1, while the stopping property Theorem 6.5 gives

$$X^{(n)}(t) = \int_0^t I_{[0,\tau_n]}(r)\mu(r)\,dr + \int_0^t I_{[0,\tau_n]}\,\sigma\,dB \quad \text{for } t\in[0,T]. \tag{11.4}$$

Now the Itô formula in Theorem 9.10 applies to the Itô process $X^{(n)}$ by Exercise 125 below. Using this together with (11.4) and the stopping property Theorem 6.5 we get

$$\begin{aligned}
&f(t,X(t\wedge\tau_n)) - f(t,X(0))\\
&= f(t,X^{(n)}(t)) - f(t,X^{(n)}(0))\\
&= \int_0^t f_r'(r,X^{(n)}(r))\,dr + \int_0^t f_x'(\,\cdot\,,X^{(n)})\,dX^{(n)} + \frac{1}{2}\int_0^t f_{xx}''(\,\cdot\,,X^{(n)})\,d[X^{(n)}] \qquad (11.5)\\
&= \int_0^t f_r'(r,X^{(n)}(r))\,dr + \int_0^{t\wedge\tau_n} f_x'(\,\cdot\,,X^{(n)})\,dX + \frac{1}{2}\int_0^{t\wedge\tau_n} f_{xx}''(\,\cdot\,,X^{(n)})\,d[X]
\end{aligned}$$

for $t\in[0,T]$. As the values of X belong to I with probability 1 and X is continuous with probability 1, there exists a null event E such that $\{\tau_n=T\}\uparrow\Omega\backslash E$ as $n\to\infty$. For $\omega\in\Omega\backslash E$ we may therefore select n sufficiently large to ensure that $\tau_n=T$ in order to conclude from (11.5) that

$$\begin{aligned}
&f(t,X(t)) - f(t,X(0))\\
&= \int_0^t f_r'(r,X(r))\,dr + \int_0^t f_x'(\,\cdot\,,X^{(n)})\,dX + \frac{1}{2}\int_0^t f_{xx}''(\,\cdot\,,X)\,d[X] \quad \text{for}\, t\in[0,T],
\end{aligned}$$

for n sufficiently large, for $\omega\in\Omega\backslash E$. Hence it is sufficient to prove that

$$\begin{aligned}
\int_0^t f_x'(\,\cdot\,,X^{(n)})\,dX &= \int_0^t f_x'(r,X^{(n)}(r))\mu(r)\,dr + \int_0^t f_x'(\,\cdot\,,X^{(n)})\sigma\,dB\\
&\to \int_0^t f_x'(r,X(r))\mu(r)\,dr + \int_0^t f_x'(\,\cdot\,,X)\sigma\,dB\\
&= \int_0^t f_x'(\,\cdot\,,X)\,dX \quad \text{as } n\to\infty \quad \text{for } t\in[0,T]
\end{aligned}$$

[8]The class of functions $f:[0,T]\times I\to\mathbb{R}$ that are continuously differentiable in the first argument and two times continuously differentiable in the second argument.

in the sense of convergence in probability. Here the Lebesgue integrals $\int_0^t f_x'(r,X^{(n)}(r))\mu(r)\,dr$ and $\int_0^t f_x'(r,X(r))\mu(r)\,dr$ agree for $t\in[0,T]$, for n sufficiently large, for $\omega\in\Omega\backslash E$, by what we have proved already, so that it is sufficient to prove that

$$\int_0^t f_x'(\cdot,X^{(n)})\sigma\,dB \to \int_0^t f_x'(\cdot,X)\sigma\,dB \quad \text{as } n\to\infty \text{ for } t\in[0,T] \tag{11.6}$$

in the sense of convergence in probability. By Theorem 5.4, (11.6) in turn holds if

$$\int_0^T \big(f_x'(r,X^{(n)}(r))\sigma(r)-f_x'(r,X(r))\sigma(r)\big)^2\,dr \to 0 \quad \text{as } n\to\infty \tag{11.7}$$

in the sense of convergence in probability. However, (11.7) does in fact hold with almost sure convergence since the integral on the left-hand side is equal to zero for n sufficiently large, for $\omega\in\Omega\backslash E$, by what we have proved already. □

Exercise 125. Show that the Itô formula in Theorem 9.10 applies to the Itô process $X^{(n)}$ in the proof of Theorem 11.1.

Theorem 11.2. (ITÔ FORMULA) *For two Itô processes X and Y all values of which belong to two open intervals $I_X\subseteq\mathbb{R}$ and $I_Y\subseteq\mathbb{R}$, respectively, with probability 1 and a function $f\in C^2(I_X\times I_Y)$*[9]*, we have*

$$\begin{aligned}&df(X,Y)\\ &= f_x'(X,Y)\,dX+f_y'(X,Y)\,dY+\frac{1}{2}f_{xx}''(X,Y)\,d[X]+\frac{1}{2}f_{yy}''(X,Y)\,d[Y]+f_{xy}''(X,Y)\,d[X,Y].\end{aligned}$$

Exercise 126. Prove Theorem 11.2.

11.3 Strong uniqueness

By inspection of the proof of Theorem 11.1 we readily get the following improvement of the strong uniqueness result Theorem 10.5:

Theorem 11.3. *Let $I\subseteq\mathbb{R}$ be an open interval and let $\{X(t)\}_{t\in[0,T]}$ be an Itô process that takes values in I and solves the diffusion type SDE (9.1) with coefficients $\mu,\sigma:[0,T]\times I\to\mathbb{R}$ that satisfy the local Lipschitz condition*

$$|\mu(t,x)-\mu(t,y)|+|\sigma(t,x)-\sigma(t,y)|\le K_{a,b}\,|x-y| \quad \text{for } t\in[0,T] \text{ and } x,y\in[a,b],$$

for some constant $K_{a,b}>0$ for every closed interval $[a,b]\subseteq I$. Then every other solution to the SDE must agree with X with probability 1.

[9] The class of functions $f:I_X\times I_Y\to\mathbb{R}$ for which all second order derivatives exist and are continuous.

Exercise 127. Prove Theorem 11.3.

11.4 Stochastic logarithm

Here we give a star application of our new Itô formula Theorem 11.1.

Definition 11.4. *The* stochastic logarithm $\{\mathcal{L}(X)(t)\}_{t\in[0,T]}$ *of an Itô process* $\{X(t)\}_{t\in[0,T]}$ *that is strictly positive with probability 1 is a solution to the (not necessarily diffusion type) SDE*

$$dX(t) = X(t)\, d\mathcal{L}(X)(t) \quad \textit{for } t \in [0,T], \quad \mathcal{L}(X)(0) = 0.$$

Theorem 11.5. *The stochastic logarithm of an Itô process* $\{X(t)\}_{t\in[0,T]}$ *that is strictly positive with probability 1 is given by*

$$\mathcal{L}(X)(t) = \log(X(t)) - \log(X(0)) + \frac{1}{2}\int_0^t \frac{d[X](r)}{X(r)^2} \quad \textit{for } t \in [0,T]. \tag{11.8}$$

Proof. By Itô's formula Theorem 11.1 the process $Y(t) = \log(X(t))$ satisfies

$$dY(t) = \frac{dX(t)}{X(t)} - \frac{1}{2}\frac{d[X](t)}{X(t)^2},$$

so that

$$X(t)\, d\left(Y(t) + \frac{1}{2}\int_0^t \frac{d[X](r)}{X(r)^2}\right) = dX(t),$$

see also Exercise 128 below. □

Exercise 128. Why is the integral in (11.8) well-defined?

Exercise 129. Show that $\mathcal{L}(\mathcal{E}(X)) = X$ for any Itô process X. Is it also true that $\mathcal{E}(\mathcal{L}(X)) = X$?

12 Lecture 12, Monday May 18

12.1 Strong existence

Let us examplify that the local Lipschitz condition for the coefficients that is sufficient for strong uniqueness by Theorem 10.5 is not sufficient for strong existence:

Example 12.1. Recall from Exercise 108 the equation

$$dX(t) = X(t)^2\,dt \quad \text{for } t \in [0,T], \quad X(0) = 1,$$

where now $T = 1$. We will give this exercise a completely rigorous treatment: Assume that the equation has a strong solution $\{X(t)\}_{t\in[0,1]}$ when $T = 1$. This solution X must satisfy with probability 1 that

$$\int_0^1 |\mu(t, X(t))|\,dt + \int_0^1 \sigma(t, X(t))^2\,dt = \int_0^1 X(t)^2\,dt < \infty. \tag{12.1}$$

As the coefficient $\mu(t,x) = x^2$ and $\sigma(t,x) = 0$ are locally Lipschitz, the solution X is unique by Theorem 10.5. In particular we have a unique strong solution $\{X_T(t)\}_{t\in[0,T]}$ for every $T < 1$ that coincides with $\{X(t)\}_{t\in[0,1]}$ for $t \in [0,T]$. By solving the SDE we see that this solution is $X(t) = X_T(t) = 1/(1-t)$ for $t \leq T < 1$, for each $T = 1$. Hence we have $X(t) = 1/(1-t)$ for $t < 1$. But this contradicts (12.1), so that our initial assumption that the solution X exists is false.

It turns out that it is the über-linear growth of the coefficient of the equation in Example 12.1 that makes impossible the existence. And here is what to do about it:

Definition 12.2. *The coefficients* $\mu, \sigma : [0,T]\times\mathbb{R} \to \mathbb{R}$ *of the SDE* (9.1) *are said to satisfy a* global linear growth condition *if there exists a constant* $C > 0$ *such that*

$$\mu(t,x)^2 + \sigma(t,x)^2 \leq C\,(1+x^2) \quad \textit{for } t \in [0,T] \textit{ and } x \in \mathbb{R}.$$

Theorem 12.3. *If the coefficients of the SDE* (9.1) *satisfy a local Lipschitz condition and a global linear growth condition, then there exists a strong solution to the SDE for every given BM B and any given initial value X_0 that is strongly unique.*

Definition 12.4. *The coefficients* $\mu, \sigma : [0,T]\times\mathbb{R} \to \mathbb{R}$ *of the SDE* (9.1) *are said to satisfy a* global Lipschitz condition *if there exists a constant* $K > 0$ *such that*

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq K\,|x-y| \quad \textit{for } t \in [0,T] \textit{ and } x, y \in \mathbb{R}.$$

Corollary 12.5. *If the coefficients of the SDE* (9.1) *satisfy a global Lipschitz condition, then there exists a strong solution to the SDE for every given BM B and initial value X_0 that is strongly unique.*

Exercise 130. Derive Corollary 12.5 from Theorem 12.3.

Proof of Theorem 12.3. As we have uniqueness by Theorem 10.5 it is sufficient to prove existence. Further, it is sufficient to prove existence for any square-integrable initial value $\mathbf{E}\{X_0^2\} < \infty$. This is so because with the notation (5.1) there then exists a unique strong solution $\{{}_nX(t)\}_{t\in[0,T]}$ for every truncation $X_0^{(n)}$, $n \in \mathbb{N}$, of X_0. By the strong uniqueness of solutions, on the almost sure event $\bigcup_{n\geq 1}\{X_0 \leq n\}$ we may define a solution $\{X(t)\}_{t\in[0,T]}$ to (9.1) by $X(t) = {}_nX(t)$ for $t \in [0,T]$ when $\omega \in \{X_0 \leq n\}$, see Exercise 131 below.

Consider a so called *Picard-Lindelöf iteration* where $X_0(t) = X_0$ for $t \in [0,T]$ and

$$X_{k+1}(t) = X_0 + \int_0^t \mu(r, X_k(r))\, dr + \int_0^t \sigma(\,\cdot\,, X_k)\, dB \quad \text{for } t \in [0,T], \tag{12.2}$$

for $k \geq 0$. To establish that the process X_{k+1} on the left-hand side of (12.2) is well-defined it is sufficient to show that the process X_k on the right-hand side satisfies

$$\mathbf{E}\Big\{\sup_{t\in[0,T]} X_k(t)^2\Big\} < \infty, \tag{12.3}$$

because by Fubini's theorem together with global linear growth this shows that

$$\mathbf{E}\Big\{\int_0^t \sigma(r, X_k(r))^2\, dr\Big\} \leq \mathbf{E}\Big\{\int_0^t C\,(1 + X_k(r)^2)\, dr\Big\} = \int_0^t C\big(1 + \mathbf{E}\{X_k(r)^2\}\big)\, dr < \infty \tag{12.4}$$

for $t \in [0,T]$, so that $\sigma(\,\cdot\,, X_k) \in E_T$, and similarly, by Cauchy-Schwarz' inequality,

$$\begin{aligned} \mathbf{E}\Big\{\int_0^t |\mu(r, X_k(r))|\, dr\Big\} &\leq \bigg[\mathbf{E}\bigg\{\Big(\int_0^t |\mu(r, X_k(r))|\, dr\Big)^2\bigg\}\bigg]^{1/2} \\ &\leq \bigg[\mathbf{E}\bigg\{\Big(\int_0^t 1^2\, dr\Big)\Big(\int_0^t \mu(r, X_k(r))^2\, dr\Big)\bigg\}\bigg]^{1/2} \\ &\leq \sqrt{T}\,\bigg[\mathbf{E}\Big\{\int_0^t C\,(1 + X_k(r)^2)\, dr\Big\}\bigg]^{1/2} \\ &= \sqrt{T}\,\bigg[\int_0^t C\big(1 + \mathbf{E}\{X_k(r)^2\}\big)\, dr\bigg]^{1/2} \\ &< \infty \quad \text{for } t \in [0,T], \end{aligned} \tag{12.5}$$

so that $\int_0^T |\mu(r, X_k(r)|\, dr < \infty$ with probability 1. Now, (12.3) holds trivially for $k = 0$ since $\mathbf{E}\{X_0^2\} < \infty$. Further, if (12.3) holds for a certain $k \in \mathbb{N}$, then by the elementary inequality $(x+y+z)^2 \leq 3\,x^2+3\,y^2+3\,z^2$ and Doob's maximal inequality Theorem 13.1 below together with (12.2), isometry and (12.4)-(12.5), we have

$$
\begin{aligned}
&\mathbf{E}\Big\{\sup_{s\in[0,t]} X_{k+1}(s)^2\Big\}\\
&\leq 3\,\mathbf{E}\{X_0^2\}+3\,\mathbf{E}\bigg\{\sup_{s\in[0,t]}\bigg(\int_0^s \mu(r,X_k(r))\,dr\bigg)^2\bigg\}+3\,\mathbf{E}\bigg\{\sup_{s\in[0,t]}\bigg(\int_0^s \sigma(\,\cdot\,,X_k)\,dB\bigg)^2\bigg\}\\
&\leq 3\,\mathbf{E}\{X_0^2\}+3\,T\int_0^t C\big(1+\mathbf{E}\{X_k(r)^2\}\big)\,dr+12\,\mathbf{E}\bigg\{\int_0^t \sigma(r,X_k(r))^2\,dr\bigg\}\\
&\leq D+D\int_0^t \mathbf{E}\Big\{\sup_{r\in[0,s]} X_k(r)^2\Big\}\,ds \quad \text{for } t\in[0,T],
\end{aligned}
\tag{12.6}
$$

where $D=\max\{3\mathbf{E}\{X_0^2\}+3T^2C+12TC,\,3TC+12C\}$. This gives (12.3) for $k+1$, and thus (12.3) holds for all $k\in\mathbb{N}$ by induction. In fact, from (12.6) we may readily deduce by means of iteration the following stronger version of (12.3)

$$
\mathbf{E}\Big\{\sup_{s\in[0,t]} X_k(s)^2\Big\} \leq (D\vee 1)\,\mathrm{e}^{D\,t} \quad \text{for } t\in[0,T] \text{ and } k\geq 0, \tag{12.7}
$$

see Exercise 132 below. (Also compare with Grönwall's lemma!)

Now assume that the global Lipschitz condition in Definition 12.4 holds – an assumption that we will relax to a local Lipschitz condition later. By application of the arguments used to obtain (12.6) [this time using the inequality $(x+y)^2\leq 2\,x^2+2\,y^2$ instead of $(x+y+z)^2\leq 3\,x^2+3\,y^2+3\,z^2$], we then readily obtain

$$
\mathbf{E}\bigg\{\sup_{s\in[0,t]}\big(X_{k+1}(s)-X_k(s)\big)^2\bigg\} \leq L\int_0^t \mathbf{E}\bigg\{\sup_{r\in[0,s]}\big(X_k(r)-X_{k-1}(r)\big)^2\bigg\}\,ds \tag{12.8}
$$

for $t\in[0,T]$ and $k\in\mathbb{N}$, where $L=2TK^2+8K^2$, see Exercise 133 below. By induction this in turn gives

$$
\mathbf{E}\bigg\{\sup_{s\in[0,t]}\big(X_{k+1}(s)-X_k(s)\big)^2\bigg\} \leq \mathbf{E}\bigg\{\sup_{s\in[0,T]}(X^{(1)}(s)-X^{(0)}(s))^2\bigg\}\,\frac{(Lt)^k}{k!} \leq M\,\frac{(LT)^k}{k!} \tag{12.9}
$$

for $t\in[0,T]$ and $k\in\mathbb{N}$, for some constant $M<\infty$ [recall (12.3)], see Exercise 134 below. By Tjebysjev's inequality, (12.9) implies that

$$
\sum_{k=0}^{\infty}\mathbf{P}\bigg\{\sup_{t\in[0,T]}\big|X_{k+1}(t)-X_k(t)\big|>\frac{1}{2^k}\bigg\} \leq \sum_{k=0}^{\infty} M\,\frac{(LT)^k}{k!}\,(2^k)^2 = \sum_{k=0}^{\infty}\frac{M\,(4LT)^k}{k!}<\infty.
$$

And so the Borel-Cantelli lemma shows that

$$
\mathbf{P}\bigg\{\sup_{t\in[0,T]}\big|X_{k+1}(t)-X_k(t)\big|>\frac{1}{2^k} \text{ for infinitely many } k\in\mathbb{N}\bigg\}=0.
$$

From this in turn we obtain by summation that

$$
\sup_{t\in[0,T]}\big|X_m(t)-X_n(t)\big| \leq \sum_{k=m\wedge n}^{m\vee n-1}\sup_{t\in[0,T]}\big|X_{k+1}(t)-X_k(t)\big| \leq \sum_{k=m\wedge n}^{\infty}\frac{1}{2^k}=\frac{2}{2^{m\wedge n}}
$$

for $m\wedge n$ sufficiently large with probability 1. Hence the processes $\{X_k\}_{k=1}^\infty$ constitute a Cauchy sequence with respect to uniform convergence of continuous functions on the interval $[0,T]$ with probability 1. It follows that there exists a stochastic process $\{X(t)\}_{t\in[0,T]}$ that is continuous with probability 1 such that $\sup_{t\in[0,T]}|X_k(t)-X(t)|\to 0$ as $k\to\infty$ with probability 1.

Using the convergence established in the previous paragraph and Fatou's lemma together with the elementary inequality $(\sum_{i=1}^n x_i)^2 \le \sum_{i=1}^n 2^i x_i^2$ and (12.9), we obtain

$$\begin{aligned}
\mathbf{E}\Big\{\sup_{t\in[0,T]}\big(X(t)-X_k(t)\big)^2\Big\} &= \mathbf{E}\Big\{\sup_{t\in[0,T]}\Big(\liminf_{\ell\to\infty} X_\ell(t)-X_k(t)\Big)^2\Big\}\\
&= \mathbf{E}\Big\{\liminf_{\ell\to\infty}\sup_{t\in[0,T]}\big(X_\ell(t)-X_k(t)\big)^2\Big\}\\
&\le \liminf_{\ell\to\infty}\mathbf{E}\Big\{\sup_{t\in[0,T]}\big(X_\ell(t)-X_k(t)\big)^2\Big\}\\
&\le \limsup_{\ell\to\infty}\sum_{n=k}^{\ell-1}2^{n-k+1}\,\mathbf{E}\Big\{\sup_{t\in[0,T]}\big(X_{n+1}(t)-X_n(t)\big)^2\Big\}\\
&\le \sum_{n=k}^{\infty} M\,2^{1-k}\,\frac{(2LT)^n}{n!}\\
&\le \frac{2M\,\mathrm{e}^{2LT}}{2^k}\quad\text{for }\ k\in\mathbb{N}.
\end{aligned}\tag{12.10}$$

In particular we have

$$\mathbf{E}\Big\{\sup_{t\in[0,T]}X(t)^2\Big\}\le 2\,\mathbf{E}\Big\{\sup_{t\in[0,T]}\big(X(t)-X_0(t)\big)^2\Big\}+2\,\mathbf{E}\{X_0^2\}<\infty,$$

so that the Itô process

$$Y(t)=X_0+\int_0^t\mu(r,X(r))\,dr+\int_0^t\sigma(\,\cdot\,,X)\,dB\quad\text{for }\ t\in[0,T]$$

is well-defined [by the argument employed to show that the left-hand of (12.2) is well-define when (12.3) holds]. Hence it is enough to show that $X=Y$ with probability 1 to establish the existence of a solution. However, by an obvious version of (12.8) together with (12.10), we have

$$\mathbf{E}\Big\{\sup_{t\in[0,T]}\big(Y(t)-X_{k+1}(t)\big)^2\Big\}\le L\int_0^T\mathbf{E}\Big\{\sup_{s\in[0,t]}\big(X(s)-X_k(s)\big)^2\Big\}\,dt\le\frac{2MLT\,\mathrm{e}^{2LT}}{2^k}$$

for $k\in\mathbb{N}$, so that by another application of (12.10)

$$\begin{aligned}
\mathbf{E}\{(Y(t)-X(t))^2\} &\le 2\,\mathbf{E}\Big\{\sup_{t\in[0,T]}\big(Y(t)-X_{k+1}(t)\big)^2\Big\}+2\,\mathbf{E}\Big\{\sup_{t\in[0,T]}\big(X_{k+1}(t)-X(t)\big)^2\Big\}\\
&\le\frac{2MLT\,\mathrm{e}^{2LT}}{2^k}+\frac{2M\,\mathrm{e}^{2LT}}{2^{k+1}}\quad\text{for }\ k\in\mathbb{N},
\end{aligned}$$

which gives $X(t) = Y(t)$ with probability 1 by sending $k \to \infty$.

It remains to show how to relax the global Lipschitz condition to a local Lipschitz condition. To that end we consider the truncated coefficients [recall (5.1)]

$$\mu^{(n)}(t,x) = \begin{cases} \mu(t,-n) & \text{for} \quad x < -n \\ \mu(t,x) & \text{for} \quad |x| \le n \\ \mu(t,n) & \text{for} \quad x > n \end{cases} \quad \text{and} \quad \sigma^{(n)}(t,x) = \begin{cases} \sigma(t,-n) & \text{for} \quad x < -n \\ \sigma(t,x) & \text{for} \quad |x| \le n \\ \sigma(t,n) & \text{for} \quad x > n \end{cases}.$$

It is easy to see that the truncated coefficients $\mu^{(n)}$ and $\sigma^{(n)}$ satisfy a global linear growth condition with the same global growth constant C as the non-truncated coefficients for each n, and that they satisfy a global Lipschitz condition with global Lipschitz constant K_n for each n, where K_n comes from the local Lipschitz condition Definition 10.4 for the non-truncated coefficients, see Exercise 135 below. Hence it follows from what we have proved already that the SDE

$$dX(t) = \mu^{(n)}(t, X(t))\, dt + \sigma^{(n)}(t, X(t))\, dB(t) \quad \text{for } t \in [0,T], \quad X(0) = X_0,$$

has a unique strong solution ${}_nX$ for each $n \in \mathbb{N}$. By the strong uniqueness we may define a solution X to (9.1) on the event $\mathcal{E} = \bigcup_{n\ge 1}\{\sup_{t\in[0,T]} |{}_nX(t)| \le n\}$ by $X = {}_nX$ when $\omega \in \{\sup_{t\in[0,T]} |{}_nX(t)| \le n\}$. Note that this is in essence the same argument as the one we employed to justify the truncation of X_0, see also Exercise 131 below. Hence it is sufficient to show that the event $\mathcal{E}$ is almost sure. However, as all processes ${}_nX$ are almost sure limits of processes $({}_nX)_k$ [given by (12.2) with μ and σ replaced with the truncated coefficients], that in turn all satisfy the inequality (12.7) with a constant D that does not depend on n, Tjebysjev's inequality and (12.7) give

$$\sum_{n=1}^{\infty} \mathbf{P}\Big\{\sup_{t\in[0,T]} |{}_nX(t)| > n\Big\} \le \sum_{n=1}^{\infty} \liminf_{k\to\infty} \mathbf{P}\Big\{\sup_{t\in[0,T]} |({}_nX)_k(t)| > n\Big\} \le \sum_{n=1}^{\infty} \frac{(D\vee 1)\,\mathrm{e}^{DT}}{n^2} < \infty.$$

Hence the Borel-Cantelli lemms shows that

$$\mathbf{P}\Big\{\sup_{t\in[0,T]} |{}_nX(t)| > n \text{ for infinitely many } n \in \mathbb{N}\Big\} = 0. \qquad \square$$

Exercise 131. The argument in the first paragraph of the proof of Theorem 12.3 that it is sufficient to prove the theorem for X_0 square-integrable relies on the fact that two strong solutions ${}_mX$ and ${}_nX$ with truncated initial values $X_0^{(m)}$ and $X_0^{(n)}$, respectively, must coincide on the event $\mathcal{E}_{m\wedge n} = \{X_0 \le m \wedge n\}$. Recalling that X_0 is independent of the BM B, show how this uniqueness follows by applying the uniqeness Theorem 10.5 to the SDE (9.1) on the modified probability space

$$\big(\Omega \cap \mathcal{E}_{m\wedge n}, \mathcal{F} \cap \mathcal{E}_{m\wedge n}, \mathbf{P}\{\,\cdot\,|\mathcal{E}_{m\wedge n}\}\big) \quad \text{with filtration} \quad \{\mathcal{F}_t \cap \mathcal{E}_{m\wedge n}\}_{t\ge 0}$$

(on which the given BM B is still a BM by independence of X_0 and B).

Exercise 132. Prove that (12.6) gives (12.7).

Exercise 133. Prove (12.8) by means of modification of the proof of (12.6).

Exercise 134. Prove that (12.8) gives (12.9).

Exercise 135. Show that the truncated coefficients $\mu^{(n)}$ and $\sigma^{(n)}$ in the proof of Theorem 12.3 satisfy a global linear growth condition with the same global growth constant C as the non-truncated coefficients for each n, and that they satisfy a global Lipschitz condition with global Lipschitz constant K_n for each n, where K_n is the local Lipschitz constant for the non-truncated coefficients.

Corollary 12.6. *If the coefficients of the SDE* (9.1) *satisfy a global Lipschitz condition and if the initial value is square-integrable* $\mathbf{E}\{X_0^2\} < \infty$*, then there exists a strong solution* $\{X(t)\}_{t\in[0,T]}$ *to the SDE for any given BM B that is strongly unique. This solution is square-integrable* $\sup_{t\in[0,T]} \mathbf{E}\{X(t)^2\} \leq \mathbf{E}\{\sup_{t\in[0,T]} X(t)^2\} < \infty$.

Exercise 136. Prove Corollary 12.6 by inspection of the proofs of Theorem 12.3 and Corollary 12.5.

13 Lecture 13, Wednesday May 20

13.1 Doob's maximal inequality

Here is another important consequence of Doob's inequality Theorem 3.1:

Theorem 13.1. (Doob's maximal inequality) *Pick a constant* $p > 1$. *For a right-continuous martingale or a non-negative submartingale* $\{X(t)\}_{t\in[0,T]}$ *such that the process* $\{|X(t)|^p\}_{t\in[0,T]}$ *is integrable, we have*

$$\mathbf{E}\Big\{\sup_{0\le t\le T}|X(t)|^p\Big\} \le q^p\,\mathbf{E}\{|X(T)|^p\} \quad \textit{where} \quad \frac{1}{p}+\frac{1}{q}=1.$$

Proof. From Exercises 22 and 23 we know that $|X|$ is a submartingale. Hence we may apply Doob's inequality Theorem 3.1 to $|X|$ as well as Hölder's inequality to obtain

$$\begin{aligned}
\mathbf{E}\Big\{\Big(\sup_{0\le t\le T}|X(t)|\wedge k\Big)^p\Big\} &= \mathbf{E}\Big\{\int_0^{\sup_{0\le t\le T}|X(t)|\wedge k} p\,\lambda^{p-1}\,d\lambda\Big\}\\
&= \mathbf{E}\Big\{\int_0^k p\,\lambda^{p-1}\,I_{\{\sup_{0\le t\le T}|X(t)|\ge\lambda\}}\,d\lambda\Big\}\\
&= \int_0^k p\,\lambda^{p-1}\,\mathbf{P}\Big\{\sup_{0\le t\le T}|X(t)|\ge\lambda\Big\}\,d\lambda\\
&\le \int_0^k p\,\lambda^{p-1}\,\frac{\mathbf{E}\big\{|X(T)|\,I_{\{\sup_{0\le t\le T}|X(t)|\ge\lambda\}}\big\}}{\lambda}\,d\lambda\\
&= p\,\mathbf{E}\Big\{|X(T)|\int_0^{\sup_{0\le t\le T}|X(t)|\wedge k}\lambda^{p-2}\,d\lambda\Big\}\\
&= q\,\mathbf{E}\Big\{|X(T)|\Big(\sup_{0\le t\le T}|X(t)|\wedge k\Big)^{p-1}\Big\}\\
&\le q\,\mathbf{E}\big\{|X(T)|^p\big\}^{1/p}\,\mathbf{E}\Big\{\Big(\sup_{0\le t\le T}|X(t)|\wedge k\Big)^p\Big\}^{(p-1)/p},
\end{aligned}$$

so that by rearrangement together with Fatou's lemma

$$\mathbf{E}\Big\{\sup_{0\le t\le T}|X(t)|^p\Big\} \le \liminf_{k\to\infty}\mathbf{E}\Big\{\Big(\sup_{0\le t\le T}|X(t)|\wedge k\Big)^p\Big\} \le q^p\,\mathbf{E}\big\{|X(T)|^p\big\}. \qquad \square$$

13.2 Uniform integrability

Definition 13.2. *A family* $\{Y_\alpha\}_{\alpha\in\mathfrak{A}}$ *of random variables (not necessarily defined on a common probability space) is* uniformly integrable *if*

$$\lim_{y\to\infty}\sup_{\alpha\in\mathfrak{A}}\mathbf{E}\big\{I_{\{|Y_\alpha|>y\}}|Y_\alpha|\big\}=0.$$

Theorem 13.3. *Each member Y_α of a family $\{Y_\alpha\}_{\alpha\in\mathfrak{A}}$ of uniformly integrable random variables is integrable.*

Exercise 137. Prove Theorem 13.3.

Theorem 13.4. *A family $\{Y_\alpha\}_{\alpha\in\mathfrak{A}}$ of random variables is uniformly integrable if $\sup_{\alpha\in\mathfrak{A}} \mathbf{E}\{|Y_\alpha|^p\} < \infty$ for some constant $p > 1$.*

Exercise 138. Derive Theorem 13.4 from the Hölder and Tjebysjev inequalities.

Theorem 13.5. *A family $\{Y_\alpha\}_{\alpha\in\mathfrak{A}}$ of random variables defined on a common probability space is uniformly integrable if $|Y_\alpha| \leq Z$ almost surely for each $\alpha \in \mathfrak{A}$, for some integrable random variable Z.*

Exercise 139. Prove Theorem 13.5.

Theorem 13.6. *If $\{Y_n\}_{n=1}^\infty$ is an uniformly integrable sequence of random variables such that $Y_n \to Y$ weakly as $n \to \infty$ for some random variable Y, then Y is integrable.*

Exercise 140. Use Fatou's lemma to prove Theorem 13.6.

The next two theorems are deeper than the rather immediate Theorems 13.3-13.6:

Theorem 13.7. *For a sequence of integrable random variables $\{Y_n\}_{n=1}^\infty$ defined on a common probability space such that $Y_n \to Y$ in probability as $n \to \infty$ for some random variable Y, the following statements are equivalent:*

(1) *$Y_n \to Y$ in $\mathbb{L}^1$ as $n \to \infty$;*

(2) *Y is integrable and $\lim_{n\to\infty} \mathbf{E}\{|Y_n|\} = \mathbf{E}\{|Y|\}$;*

(3) *$\{Y_n\}_{n=1}^\infty$ is uniformly integrable.*

Theorem 13.8. *For a sequence of integrable random variables $\{Y_n\}_{n=1}^\infty$ such that $Y_n \to Y$ weakly as $n \to \infty$ for some random variable Y, we have that Y is integrable and $\lim_{n\to\infty} \mathbf{E}\{|Y_n|\} = \mathbf{E}\{|Y|\}$ if and only if $\{Y_n\}_{n=1}^\infty$ is uniformly integrable. In either case we have $\lim_{n\to\infty} \mathbf{E}\{Y_n\} = \mathbf{E}\{Y\}$.*

In other words, uniform integrability is what you should replace dominated convergence with to get a sharp condition for convergence of your means!

We will not prove Theorems 13.7 and 13.8, although we make some use of them, because we consider them to belong to basic graduate probability theory courses.

13.3 Local martingales

We know that Itô integral processes $\{\int_0^t \sigma\, dB\}_{t\in[0,T]}$ are square-integrable continuous martingales when $\sigma \in E_T$. We have made crucial use of this fact in several proofs, e.g., to be able to use the Doob inequalities. However, we have not investigated what happens with the martingale property if $\sigma \in P_T \backslash E_T$. It turns out that for such σ the Itô intergal process is no longer necessarily a martingale. However, an Itô intergal process is always a continuous so called *local martingale*, see Definition 13.9 below. In fact, even the converse holds, that is, continuous local martingales are Itô intergal processes (although we will not make explicit use of that fact). As it turns out that it is possible to develop a rather rich theory for (continuous) local martingales on their own, it is not surprising that this theory is important in the study of Itô processes. In particular, local martingale theory is crucial for the study of weak solutions to SDE.

Definition 13.9. *An adapted stochastic process* $\{M(t)\}_{t\in[0,T]}$ *is a* local martingale/submartingale/supermartingale *if there exists a so called* localizing sequence *of stopping times* $0 \le \tau_1 \le \tau_2 \le \ldots$ *such that* $\tau_n \to \infty$ *almost surely as* $n \to \infty$ *and* $\{M(t\wedge\tau_n)\}_{t\in[0,T]}$ *is a uniformly integrable martingale/submartingale/supermartingale for each* $n \in \mathbb{N}$.

Surprisinlgy, the uniform integrability requirement in Definition 13.9 often is void:

Theorem 13.10. *A right-continuous martingale/non-negative submartingale/non-positive supermartingale* $\{M(t)\}_{t\in[0,T]}$ *is uniformly integrable.*

Proof. For a martingale M Jensen's inequality gives

$$\begin{aligned}
\sup_{t\in[0,T]} \mathbf{E}\big\{I_{\{|M(t)|>y\}}\,|M(t)|\big\} &= \sup_{t\in[0,T]} \mathbf{E}\big\{I_{\{|M(t)|>y\}}\,\big|\mathbf{E}\{M(T)\,|\,\mathcal{F}_t\}\big|\big\} \\
&\le \sup_{t\in[0,T]} \mathbf{E}\big\{I_{\{|M(t)|>y\}}\,\mathbf{E}\{|M(T)|\,|\,\mathcal{F}_t\}\big\} \\
&= \sup_{t\in[0,T]} \mathbf{E}\big\{I_{\{|M(t)|>y\}}\,|M(T)|\big\} \\
&\le \mathbf{E}\big\{I_{\{\sup_{t\in[0,T]}|M(t)|>y\}}\,|M(T)|\big\}.
\end{aligned}$$

As $|M(T)|$ is integrable the right-hand side goes to zero as $y \to \infty$ by so called absolute continuity of the integral, if $\mathbf{P}\{\sup_{t\in[0,T]}|M(t)|>y\}\to 0$ as $y\to\infty$. However, the latter fact follows from the Doob-Kolmogorov inequality. The proof for non-negative submartingales/non-positive supermartingales is Exercise 141 below. □

Exercise 141. Prove Theorem 13.10 for non-negative submartingales and non-positive supermartingales. (Now you need to use Doob's inequality Theorem 3.1 instead of the Doob-Kolmogorov inequality.)

Corollary 13.11. *A right-continuous martingale/non-negative submartingale/non-positive supermartingale $\{M(t)\}_{t\in[0,T]}$ is a local martingale/submartingale/supermartingale.*

Exercise 142. Prove Corollary 13.11.

Albeit not easy to exemplify, integrable local martingales need not be martingales. In fact, not even uniformly integrable local martingales have to be martingales. Also, local martingales may have non-constant means. (Just to not give you any ideas ...)

Theorem 13.12. *A local martingale/submartingale/supermartingale $\{M(t)\}_{t\in[0,T]}$ such that $|M(t)|\leq Z$ for each $t\in[0,T]$ almost surely, for some integrable random variable Z is a martingale/submartingale/supermartingale.*

Proof. It is enough to prove the theorem for local submartingales. To that end pick a localizing sequence $\{\tau_n\}_{n=1}^{\infty}$. As the random variables $\{M(t\wedge\tau_n)\}_{n=1}^{\infty}$ are uniformly integrable by Theorem 13.5, and as $M(t\wedge\tau_n)\to M(t)$ almost surely as $n\to\infty$, we have $M(t\wedge\tau_n)\to M(t)$ in $\mathbb{L}^1$ as $n\to\infty$ by Theorem 13.7. Hence it follows that

$$\mathbf{E}\{M(t)|\mathcal{F}_s\}\leftarrow\mathbf{E}\{M(t\wedge\tau_n)|\mathcal{F}_s\}\geq\mathbf{E}\{M(s\wedge\tau_n)\}\to\mathbf{E}\{M(s)\}\quad\text{as } n\to\infty$$

for $0\leq s\leq t$ in the sense of convergence in $\mathbb{L}^1$ (recall Exercise 53). □

Here is one simple but crucially important result:

Theorem 13.13. (LOCAL MARTINGALE) *For $\sigma\in P_T$ the Itô intergal process $\{\int_0^t\sigma\,dB\}_{t\in[0,T]}$ is a continuous local martingale.*

Proof. We know that the Itô intergal process is continuous. Now define

$$\tau_n=n\wedge\inf\bigg\{t\in[0,T]:\int_0^t\sigma(r)^2\,dr\geq n\bigg\}$$

[cf. (6.3)]. Then we have $\int_0^{t\wedge\tau_n} \sigma\, dB = \int_0^t \sigma I_{[0,\tau_n]}\, dB$ for $t\in[0,T]$ by Theorem 6.5. Recalling from the proof of Theorem 5.1 that $\sigma I_{[0,\tau_n]}\in E_T$ [as $\int_0^T \sigma(s)^2 I_{[0,\tau_n]}(s)^2\, ds \leq n$] it follows that $\int_0^{t\wedge\tau_n} \sigma\, dB$ is a square-integrable martingale, which is uniformly integrable by Theorem 13.10. □

13.4 Doob-Meyer decomposition

Now it is time for one of the most important results of martingale theory. The proof of this result is easy in dicrete time, but painfully difficult in continuous time:

Theorem 13.14. (DOOB-MEYER DECOMPOSITION[10]) *For a continuous non-negative submartingale $\{X(t)\}_{t\in[0,T]}$ there exists a unique continuous adapted integrable and non-decreasing process $\{A(t)\}_{t\in[0,T]}$ with $A(0)=0$ and a unique continuous martingale $\{M(t)\}_{t\in[0,T]}$ with $M(0)=Y(0)$ such that $X(t)=M(t)+A(t)$ for $t\in[0,T]$.*

Example 13.15. The compensator of the submartingale $B(t)^2$ (recall Exercise 22) is $A(t)=t$, because $B(t)^2-t = 2\int_0^t B(s)\, dB(s)$ (by Itô's formula for BM), where the process on the right-hand side is a martingale.

Exercise 143. In what sense is the uniqueness in Theorem 13.14?

Proof of Theorem 13.14 for a discrete time submartingale $\{X_n\}_{n=0}^N$. We have

$$X_{n+1} = X_0 + \sum_{i=0}^{n} \big(X_{i+1} - \mathbf{E}\{X_{i+1}|\mathcal{F}_i\}\big) + \sum_{i=0}^{n} \big(\mathbf{E}\{X_{i+1}|\mathcal{F}_i\} - X_i\big) \quad \text{for } n=0,\ldots,N-1,$$

where

$$M_{n+1} = X_0 + \sum_{i=0}^{n} \big(X_{i+1} - \mathbf{E}\{X_{i+1}|\mathcal{F}_i\}\big) \quad \text{and} \quad A_{n+1} = \sum_{i=0}^{n} \big(\mathbf{E}\{X_{i+1}|\mathcal{F}_i\} - X_i\big)$$

is a martingale and an adapted integrable non-decreasing process, respectively, see Exercise 144 below. □

Exercise 144. Explain why the processes M and A in the above proof is a martingale and an adapted integrable non-decreasing process, respectively.

Definition 13.16. *The process A in the Doob-Meyer decomposition Theorem 13.14 is called the* compensator *of the continuous submartingale X.*

[10] See e.g., Karatzas and Shreve: "Brownian Motion and Stochastic Calculus", Section 1.4. This is the one and only result essential for the build up of the theory that we do not prove in these notes.

The reason for the crucial importance of the Doob-Meyer decomposition in the study of SDE is the next result:

Theorem 13.17. *A continuous square-integrable martingale* $\{M(t)\}_{t\in[0,T]}$ *has a well-defined continuous, adapted and integrable quadratic variation process* $\{[M](t)\}_{t\in[0,T]}$ *(in the sense of convergence in probability) which is given by the compensator of the continuous submartingale* $\{M(t)^2\}_{t\in[0,T]}$.

Sketch of proof. We want to prove that $M^2-[M]$ is a martingale, which is to say that

$$\mathbf{E}\{(M(t)^2 - M(s)^2) - ([M](t) - [M](s)) \,|\, \mathcal{F}_s\} = 0 \quad \text{for } 0 \leq s \leq t. \tag{13.1}$$

However, for a partition $s = t_0 < t_1 < \ldots < t_n = t$ of the interval $[s,t]$ we have

$$\mathbf{E}\left\{(M(t)^2 - M(s)^2) - \sum_{i=1}^{n}(M(t_i) - M(t_{i-1}))^2 \,\middle|\, \mathcal{F}_s\right\} = 0 \tag{13.2}$$

by basic arguments (towering etc ...), see Exercise 145 below. □

Exercise 145. Prove (13.2).

Example 13.18. The process $\{(\int_0^t \sigma\, dB)^2 - \int_0^t \sigma(s)^2\, ds\}_{t\in[0,T]}$ is a martingale for $\sigma \in E_T$.

Corollary 13.19. *A pair of continuous square-integrable martingales* $\{M(t)\}_{t\in[0,T]}$ *and* $\{N(t)\}_{t\in[0,T]}$ *has a well-defined continuous, adapted and integrable quadratic co-variation process* $\{[M,N](t)\}_{t\in[0,T]}$ *(in the sense of convergence in probability) which has the property that* $\{M(t)N(t) - [M,N](t)\}_{t\in[0,T]}$ *is a continuous martingale.*

Exercise 146. Prove Corollary 13.19.

Written exam on Lectures 1-13, Thursday May 28

Exam Thursday May 28 2009, 8.30 am - 1.30 pm

HJÄLPMEDEL/AIDS: Inga/None.

LÄRARE/TEACHER: Patrik Albin 070 69 45 709.

BETYG/GRADES: 12 poäng för godkänt/12 credits to pass the exam.

RESULTAT/RESULTS: Meddelas via email/Communicated by means of email.

A. Stratonovich Stochastic Calculus. **(a)** Let $f:\mathbb{R}\to\mathbb{R}$ be a two times continuously differentiable function and let $\{X(t)\}_{t\in[0,T]}$ be an Itô process. Show that

$$d[f'(X),X](t) = f''(X)\,d[X](t) \quad \text{for } t\in[0,T]. \qquad \textbf{(1 credit)} \qquad \text{(A.1)}$$

(b) Let $\{X(t)\}_{t\in[0,T]}$ and $\{Y(t)\}_{t\in[0,T]}$ be Itô processes. The Stratonovich integral process $\{\int_0^t Y\,\partial X\}_{t\in[0,T]}$ is defined as the limit

$$\int_0^t Y(s)\,\partial X(s) = \lim_{\max_{1\le i\le n} t_i - t_{i-1}\downarrow 0} \sum_{i=1}^n \frac{Y(t_i)+Y(t_{i-1})}{2}\,(X(t_i)-X(t_{i-1})), \qquad \text{(A.2)}$$

where $0=t_0<t_1<\ldots<t_n=t$ are partitions of the interval $[0,t]$ that becomes finer and finer. Make the definition (A.2) rigorous by means of proving that

$$\int_0^t Y(s)\,\partial X(s) = \int_0^t Y(s)\,dX(s) + \frac{1}{2}\,[Y,X](t) \quad \text{for } t\in[0,T],$$

which is to say that

$$Y(t)\,\partial X(t) = Y(t)\,dX(t) + \frac{1}{2}\,d[Y,X](t) \quad \text{for } t\in[0,T]. \qquad \textbf{(1 credit)} \quad \text{(A.3)}$$

(c) Let the Itô processes $\{X(t)\}_{t\in[0,T]}$ and $\{Y(t)\}_{t\in[0,T]}$ be given as

$$X(t) = \int_0^t \mu_X(s)\,ds + \int_0^t \sigma_X(s)\,dB(s) \quad \text{and} \quad Y(t) = \int_0^t \mu_Y(s)\,ds + \int_0^t \sigma_Y(s)\,dB(s)$$

for $t\in[0,T]$ [where $\{\mu_X(t)\}_{t\in[0,T]}$ and $\{\mu_Y(t)\}_{t\in[0,T]}$ are measurable and adpated processes such that $\mathbf{P}\{\int_0^T|\mu_X(t)|\,dt<\infty\} = \mathbf{P}\{\int_0^T|\mu_Y(t)|\,dt<\infty\} = 1$, and where $\sigma_X,\sigma_Y\in P_T$]. Use (A.3) to find sharp conditions on the processes μ_X,μ_Y,σ_X and σ_Y for the Itô and Stratonovich integral processes of Y with respect to X to agree

$$\mathbf{P}\bigg\{\int_0^t Y(s)\,dX(s) = \int_0^t Y(s)\,\partial X(s) \ \text{ for } t\in[0,T]\bigg\} = 1. \qquad \textbf{(1 credit)}$$

(d) Let $f:\mathbb{R}\to\mathbb{R}$ be a two times continuously differentiable function and let $\{X(t)\}_{t\in[0,T]}$ be an Itô process. Use (A.1) and (A.3) to prove the Stratonovich chain rule

$$\partial f(X(t)) = f'(X(t))\,\partial X(t) \quad \text{for } t\in[0,T]. \qquad \textbf{(1.5 credits)} \quad \text{(A.4)}$$

(e) Use (A.4) to solve the Stratonovich stochastic exponential equation

$$dX(t) = X(t)\,\partial B(t) \quad \text{for } t\in[0,T], \quad X(0)=1. \qquad \textbf{(0.5 credits)}$$

B. Linear and Quadratic Equations. We shall solve the general linear equation

$$dX(t) = (\alpha(t)+\beta(t)X(t))\,dt + (\gamma(t)+\delta(t)X(t))\,dB(t) \quad \text{for } t\in[0,T], \quad X(0)=X_0, \tag{B.1}$$

where $\{\alpha(t)\}_{t\in[0,T]}$, $\{\beta(t)\}_{t\in[0,T]}$, $\{\gamma(t)\}_{t\in[0,T]}$ and $\{\delta(t)\}_{t\in[0,T]}$ are measurable and adapted stochastic processes that satisfy certain integrability conditions to be specified later. We shall look for a solution to (B.1) of the type $X(t)=U(t)V(t)$, where

$$dU(t) = \beta(t)U(t)\,dt + \delta(t)U(t)\,dB(t) \quad \text{for } t\in[0,T], \quad U(0)= 1\,, \tag{B.2}$$

$$dV(t) = \quad a(t)\,dt + b(t)\,dB(t) \quad \text{for } t\in[0,T], \quad V(0)=X_0. \tag{B.3}$$

(a) Show that the equation (B.2) has solution

$$U(t) = \exp\bigg\{\int_0^t\Big(\beta(s)-\frac{1}{2}\delta(s)^2\Big)\,ds + \int_0^t \delta(s)\,dB(s)\bigg\} \quad \text{for } t\in[0,T]. \qquad \textbf{(1 credit)}$$

(b) Show that $a(t) = (\alpha(t)-\gamma(t)\delta(t))/U(t)$ and $b(t) = \gamma(t)/U(t)$ in (B.3). **(1 credit)**

(c) Which integrability conditions on the coefficient processes α, β, γ and δ makes

$$X(t) = U(t)\bigg(X_0 + \int_0^t \frac{\alpha(s)-\gamma(s)\delta(s)}{U(s)}\,ds + \int_0^t \frac{\gamma(s)}{U(s)}\,dB(s)\bigg\}\bigg) \quad \text{for } t\in[0,T]$$

[with U as in Task a] a well-defined solution to (B.1)? **(1 credit)**

We now consider quadratic equations. Recall that we have seen that the equation

$$dX(t) = X(t)^2\,dt \quad \text{for } t\in[0,T], \quad X(0)=X_0,$$

has no solution for $TX_0\geq 1$. But how about an equation with a quadratic σ-coefficient

$$dX(t) = X(t)^2\,dB(t) \quad \text{for } t\in[0,T], \quad X(0)=X_0? \tag{B.4}$$

(d) Show that $\{X(t)\}_{t\in[0,T]}$ is a strictly positive solution to (B.4) if and only if $Y=1/X^2$ is a strictly positive solution to the equation

$$dY(t) = 3\,dt - 2\sqrt{Y(t)}\,dB(t) \quad \text{for } t\in[0,T], \quad Y(0)=1/X_0^2. \quad \textbf{(0.5 credits)} \tag{B.5}$$

(e) Show that $Y(t) = B_1(t)^2+B_2(t)^2+B_3(t)^2$ for $t\in[0,T]$ is a weak solution to (B.5) with $Y(0)=0$ when B_1, B_2 and B_3 are independent BM's. **(1.5 credits)**

Of course, $X=1/\sqrt{Y}$ with Y from Task e does not solve (B.4) as $Y(0)=0$. But

$$X(t) = \frac{1}{\sqrt{Y(t+1)}} = \frac{1}{\sqrt{B_1(t+1)^2+B_2(t+1)^2+B_3(t+1)^2}} \quad \text{for } t\in[0,T]$$

will be a weak solution to (B.4) with $X_0=1/\sqrt{Y(1)}$, see also Task C c below.

To complete Task e you might want to find the quadratic variation of the process

$$M(t) = \sum_{i=1}^3 \int_0^t 2B_i(s)\,dB_i(s) \quad \text{for } t\in[0,T],$$

and make use of the fact (to be proven in Lecture 14) that if a martingale $\{M(t)\}_{t\in[0,T]}$ has quadratic variation $[M](t) = \int_0^t \sigma(s)^2\,ds$ for $t\in[0,T]$, for some $\sigma\in P_T$, then it holds that $M(t)=\int_0^t \sigma\,dW$ for $t\in[0,T]$, for some Brownian motion $\{W(t)\}_{t\in[0,T]}$.

C. Martingales.

(a) Prove that $\{B(t)^2-t\}_{t\in[0,T]}$ is a martingale in two different ways. **(0.5 credits)**

(b) Show that

$$M(t)=\int_0^t \mathrm{e}^{B(s)^2}\,dB(s)\quad\text{for }\ t\in[0,T]$$

is a well-defined Itô integral process that is a square-integrable martingale for $T\in[0,\frac{1}{4}]$, but that is not a square-integrable martingale for $T\in(\frac{1}{4},1]$. **(1 credit)**

To complete the second part of Task b you might want to use the consequence of the Doob-Meyer decomposition that the submartingale $\{M(t)^2\}_{t\in[0,T]}$ has an integrable compensator that coincides with the quadratic variation $\{[M](t)\}_{t\in[0,T]}$ of M when $\{M(t)\}_{t\in[0,T]}$ is a continuous square-integrable martingale.

(c) Let B_1, B_2 and B_3 be independent BM's. Show that the weak solution

$$X(t)=\frac{1}{\sqrt{B_1(t{+}1)^2+B_2(t{+}1)^2+B_3(t{+}1)^2}}\qquad\text{for }\ t\in[0,T]$$

obtained in Task B e to the equation

$$dX(t)=X(t)^2\,dB(t)\quad\text{for }\ t\in[0,T],\quad X(0)=\frac{1}{\sqrt{B_1(1)^2+B_2(1)^2+B_3(1)^2}},$$

is a local martingale that is square-integrable but that is not a martingale. **(1 credit)**

To complete Task c you might want to recall that $B_1(1)^2{+}B_2(1)^2{+}B_3(1)^2$ has probability density function $f_{B_1(1)^2+B_2(1)^2+B_3(1)^2}(x)=\sqrt{x/(2\pi)}\,\mathrm{e}^{-x/2}$ for $x\geq 0$ and that martingales have constant means, and use that the self-similarity of BM gives

$$B_1(t{+}1)^2+B_2(t{+}1)^2+B_3(t{+}1)^2=_{\text{distribution}}(t{+}1)\,(B_1(1)^2{+}B_2(1)^2{+}B_3(1)^2)\quad\text{for }\ t\geq 0.$$

(d) Given a constant $p\geq 1$ and a right-continuous martingale $\{M(t)\}_{t\in[0,T]}$ such that the process $\{|M(t)|^p\}_{t\in[0,T]}$ is integrable, the Doob-Kolmogorov inequality says that

$$\mathbf{P}\Big\{\sup_{0\leq t\leq T}|M(t)|\geq\lambda\Big\}\leq\frac{\mathbf{E}\{|M(T)|^p\}}{\lambda^p}\quad\text{for }\ \lambda>0.$$

Show that it (unsurprisingly) is sufficient to require that $\mathbf{E}\{|M(T)|^p\}<\infty$ [rather than that $\mathbf{E}\{|M(t)|^p\}<\infty$ for all $t\in[0,T]$] for the inequality to hold. **(0.5 credits)**

(e) Show that a non-negative local supermartingale is a supermartingale. Also, show that a local submartingale $\{M(t)\}_{t\in[0,T]}$ such that $\{|M(t)|\}_{t\in[0,T]}$ is a square-integrable submartingale is a square-integrable submartingale. **(2 credits)**

To complete the first part of Task e you might want to use Fatou's lemma to prove integrability and Fatou's lemma for conditional expectations to prove the defining supermartingale inequality. To complete the second part of Task e you might want to use Doob's maximal inequality to show that $\mathbf{E}\{(\sup_{t\in[0,T]}|M(t)|)^2\}<\infty$.

D. Stationary Ornstein-Uhlenbeck Processes. A zero-mean Gaussian stochastic process $\{X(t)\}_{t\geq 0}$ is called stationary if it has covariance function

$$[0,\infty)\ni s,t \curvearrowright r(t-s) = \mathbf{E}\{X(s)X(t)\}\in\mathbb{R}$$

that depends only on the distance $t-s$ between the times $s,t\geq 0$. We shall investigate which measurable and stationary zero-mean Gaussian stochastic processes $\{X(t)\}_{t\geq 0}$ that are Markov processes. Here the Markov property means that

$$\mathbf{E}\{I_{\{X(t)\in A\}}|\mathcal{F}_s^X\} = \mathbf{E}\{I_{\{X(t)\in A\}}|X(s)\} \quad \text{for } 0\leq s\leq t \text{ and measurable sets } A\subseteq\mathbb{R},$$

where $\mathcal{F}_s^X = \sigma(X(r): r\in[0,s])$ for $s\geq 0$. Note that the measurability of X ensures that the covariance function r is measurable by Fubini's theorem.

(a) Show that a measurable stationary zero-mean Gaussian Markov processes $\{X(t)\}_{t\geq 0}$ must have covariance function $r(t)=r(0)\,\mathrm{e}^{-\alpha t}$ for $t\geq 0$, for some constant $\alpha\geq 0$. **(1 credit** (This task gives only 1 credit because it is not so much about SDE.))

To complete Task a you might want to show that $\mathbf{E}\{X(t+s)|X(s)\} = r(t)X(s)/r(0)$ and use this fact to show that by taking appropriate conditional expectations

$$\frac{r(t+s)}{r(0)} = \frac{\mathbf{E}\{X(t+s)X(0)\}}{r(0)} = \frac{r(s)}{r(0)}\,\frac{r(t)}{r(0)} \quad \text{for } s,t\geq 0.$$

This in turn is the so called Cauchy functional equation, the only measurable solutions of which take the form $r(t)/r(0) = \mathrm{e}^{-\alpha t}$ for $t\geq 0$, for some constant $\alpha\geq 0$.

(b) Show that the solution

$$X(t) = \mathrm{e}^{-\alpha t}X_0 + \sqrt{2\alpha\, r(0)}\,\mathrm{e}^{-\alpha t}\int_0^t \mathrm{e}^{\alpha r}\,dB(r) \quad \text{for } t\geq 0 \tag{D.1}$$

to the Langevin equation

$$dX(t) = -\alpha\,dt + \sqrt{2\alpha\, r(0)}\,dB(t) \quad \text{for } t\geq 0, \quad X(0)=X_0, \tag{D.2}$$

where the initial value X_0 is normal $N(0,r(0))$ distributed (and independent of B as always) is a measurable stationary zero-mean Gaussian process with the desired covariance function $r(t)=r(0)\,\mathrm{e}^{-\alpha t}$ for $t\geq 0$. **(1.5 credits)**

(c) Show from algebraic manipulations that the process X given by (D.1) satisfies

$$X(t) = \mathrm{e}^{-\alpha(t-s)}X(s) + \sqrt{2\alpha\, r(0)}\,\mathrm{e}^{-\alpha t}\int_s^t \mathrm{e}^{\alpha r}\,dB(r) \quad \text{for } 0\leq s\leq t. \tag{D.3}$$

(0.5 credits)

(d) Prove (D.3) in a different way by means of using (existence and) uniqueness for strong solutions to the equation (D.2) together with time homogeneity. **(1 credit)**

(e) Show from (D.3) that the process X given by (D.1) is Markov. [Thus X is *the* measurable stationary zero-mean Gaussian Markov process by Tasks a-b]. **(1 credit)**

E. Properties of Itô Integrals. Recall that we defined the Itô integral process $\{\int_0^t X\,dB\}_{t\in[0,T]}$ for X in the space S_T of simple processes on $[0,T]$ taking the form

$$X(t)=X(0)I_{\{0\}}(t)+\sum_{i=1}^{n}X_{t_{i-1}}\,I_{(t_{i-1},t_i]}(t)\quad\text{for }\ t\in[0,T],$$

for some constants $0=t_0<t_1<\ldots<t_n=T$ and for some random variables $X(0),X_{t_0},\ldots,X_{t_{n-1}}$ that are adapted to $\mathcal{F}_0,\mathcal{F}_{t_0},\ldots,\mathcal{F}_{t_{n-1}}$, respectively, and that satisfy $|X(0)|,|X_{t_0}|,\ldots,|X_{t_{n-1}}|\leq C$ for some (non-random) constant $C>0$, as

$$\int_0^t X\,dB=\sum_{i=1}^{m}X_{t_{i-1}}\,(B(t_i)-B(t_{i-1}))+X_{t_m}\,(B(t)-B(t_m))\quad\text{for }\ t\in(t_m,t_{m+1}],$$

for $m=0,\ldots,n-1$, with $\int_0^0 X\,dB=0$.

Next we defined the Itô integral process $\{\int_0^t X\,dB\}_{t\in[0,T]}$ for X in the space E_T of measurable and adapted processes $\{X(t)\}_{t\in[0,T]}$ such that

$$\mathbf{E}\bigg\{\int_0^T X(t)^2\,dt\bigg\}<\infty$$

as the limit in the sense of convergence in mean-square

$$\int_0^t X\,dB\leftarrow\int_0^t X_n\,dB\quad\text{as }\ n\to\infty\quad\text{for }\ t\in[0,T],$$

where $\{X_n\}_{n=1}^{\infty}$ is a sequence in S_T such that

$$\lim_{n\to\infty}\mathbf{E}\bigg\{\int_0^T\big(X_n(t)-X(t)\big)^2\,dt\bigg\}=0.$$

Finally we defined the Itô integral process $\{\int_0^t X\,dB\}_{t\in[0,T]}$ for X in the space P_T of measurable and adapted processes $\{X(t)\}_{t\in[0,T]}$ such that

$$\mathbf{P}\bigg\{\int_0^T X(t)^2\,dt<\infty\bigg\}=1$$

as the limit in the sense of convergence in probability

$$\int_0^t X\,dB\leftarrow\int_0^t X_n\,dB\quad\text{as }\ n\to\infty\quad\text{for }\ t\in[0,T],$$

where $\{X_n\}_{n=1}^{\infty}$ is a sequence in E_T such that in the sense of convergence in probability

$$\int_0^T(X_n(t)-X(t))^2\,dt\to 0\quad\text{as }\ n\to\infty.$$

(a) Prove linearity for Itô integrals on one of the spaces S_T, E_T or P_T[a]. **(1 credit)**

(b) Prove adaptivity for Itô integrals on one of the spaces S_T, E_T or P_T[a]. **(1 credit)**

(c) Prove continuity for Itô integrals on one of the spaces S_T, E_T or P_T[a]. **(1 credit)**

(d) Prove isometry for Itô integrals on one of the spaces S_T or E_T[b]. **(1 credit)**

(e) Prove that Itô integrals on one of the spaces S_T or E_T are martingales[b]. **(1 credit)**

[a]A proof for S_T must take off from scratch, while a proof for E_T may use the property for S_T, and a proof for P_T may use the property for E_T.

[b]A proof for S_T must take off from scratch, while a proof for E_T may use the property for S_T.

F. Stopping Times.

(a) Prove that $\tau_1+\tau_2$ is a stopping time when τ_1 and τ_2 are stopping times. **(1 credit)**

To complete Task a you might want to make use of the fact that

$$\{\tau_1+\tau_2>t\} = \bigcup_{q\in\mathbb{Q}\cap[0,t]} \big(\{\tau_1>q\}\cap\{\tau_2>t-q\}\big) \quad \text{for } t\geq 0.$$

(b) Given a stopping time τ it is a standard procedure in many a proof to approximate τ with the discrete random variable

$$\tau_n = \frac{\lfloor 2^n\tau+1\rfloor}{2^n} = \frac{k+1}{2^n} \quad \text{for } \tau\in\Big[\frac{k}{2^n},\frac{k+1}{2^n}\Big), \quad \text{for } k=0,1,2,\ldots\ .$$

Prove that the discrete approximation τ_n of τ is a stopping time for each $n\in\mathbb{N}$ and that $\tau_n\downarrow\tau$ as $n\to\infty$. **(1 credit)**

(c) Show how the approximation technique from Task b can be employed to give an alternative proof of the fact established in Task a that $\tau_1+\tau_2$ is a stopping time when τ_1 and τ_2 are stopping times. **(1 credit)**

To complete Task c you might want to take off from the fact that, with obvious notation,

$$\{\tau_1+\tau_2>t\} = \bigcap_{n=1}^{\infty}\{(\tau_1)_n+(\tau_2)_n>t\} \quad \text{for } t\geq 0.$$

Recall that the hitting time $\inf\{t\geq 0 : X(t)\geq x\}$ of any level $x\in\mathbb{R}$ is a stopping time for a continuous adapted stochastic process $\{X(t)\}_{t\geq 0}$. However, in the next task we show that $\inf\{t\geq 0 : X(t)>x\}$ need not be a stopping time.

(d) Consider a random walk $X_n=\sum_{i=1}^n\xi_i$ for $n\in\mathbb{N}$, where $\{\xi_i\}_{i=1}^{\infty}$ are indpendent random variables that are Rademacher distributed $\mathbf{P}\{\xi_i=1\}=\mathbf{P}\{\xi_i=-1\}=\frac{1}{2}$ for $i\in\mathbb{N}$. Construct a continuous time continuous adapted stochastic process $\{X(t)\}_{t\geq 0}$ by connecting all values of the discrete time process $\{X_n\}_{n\in\mathbb{N}}$ with straight lines and taking the filtration to be the filtration generated by X itself. Show that $\inf\{t\geq 0 : X(t)>n\}$ is not a stopping time for any of the levels $n=1,2,\ldots$. **(1 credit)**

Although it can happen that $\inf\{t\geq 0 : X(t)>x\}$ is not a stopping time for a continuous adapted process $\{X(t)\}_{t\geq 0}$ according to Task d, it can also happen that $\inf\{t\geq 0 : X(t)>x\}$ is a stopping time for each $x\in\mathbb{R}$ for a non-continuous adapted process $\{X(t)\}_{t\geq 0}$, as the next and final task illustrates.

(e) Show that $\inf\{t\geq 0 : X(t)>x\}$ is a stopping time for $x\in\mathbb{R}$ with respect to the filtration generated by the process itself for a Poisson process $\{X(t)\}_{t\geq 0}$. **(1 credit)**

Good Luck!

14 The optional stopping theorem

14.1 Optional stopping

The technique to stop a process $\{X(t)\}_{t\in[0,T]}$ at a stopping time τ to obtain a stopped process $\{X(t\wedge\tau)\}_{t\in[0,T]}$ we have used several times. It turns out that this technique together with the Doob-Meyer decomposition are the crucial ingredients for the build up of the whole theory. The reason for the importance of stopping is the same as when we have used it before, namely that it permits proofs to be carried out under additional integrability assumptions etc. that are not valid for the unstopped process.

Stopping times are also called *optional times*, and the key result for the use of the above mentioned stopping technique is the so called *optional stopping theorem*, which says that a continuous martingale remains a martingale when stopped at a stopping time (optional time). Previously we have only stopped Itô integral type of martingales $\int_0^t X\,dB$ for $t\in[0,T]$, where $X\in E_T$, so that by Theorem 6.3

$$\int_0^{t\wedge\tau} X\,dB = \int_0^t I_{[0,\tau]}X\,dB \quad \text{for } t\in[0,T].$$

As $X\in E_T$ implies $I_{[0,\tau]}X\in E_T$, the statement of the optional stopping theorem that $\{\int_0^{t\wedge\tau} X\,dB\}_{t\in[0,T]}$ is a martingale is immediate in this case. But now when we shall develop martingale theory in greater generality than that of Itô integrals, we have to prove the optional stopping theorem in its full generality.

Lemma 14.1. (Localization of conditional expectations) *Let $\mathcal{G}$ and $\mathcal{H}$ be two σ-algebras contained in $\mathcal{F}$, and let $A\in\mathcal{G}\cap H$ be such that $A\cap\mathcal{G}=A\cap\mathcal{H}$. For integrable random variables X and Y such that $\mathbf{P}\{I_A X = I_A Y\}=1$ it then holds that $I_A\mathbf{E}\{X\,|\,\mathcal{G}\}=I_A\mathbf{E}\{Y\,|\,\mathcal{H}\}$ with probability 1.*

Proof. As the random variables $I_A\mathbf{E}\{X\,|\,\mathcal{G}\}$ and $I_A\mathbf{E}\{Y\,|\,\mathcal{H}\}$ are both $\mathcal{G}$-measurable as well as $\mathcal{H}$-measurable (see Exercise 147 below), we have that

$$\begin{aligned}
&\int_{\{I_A\mathbf{E}\{X|\mathcal{G}\}>I_A\mathbf{E}\{Y|\mathcal{H}\}\}} (I_A\mathbf{E}\{X\,|\,\mathcal{G}\}-I_A\mathbf{E}\{Y\,|\,\mathcal{H}\})\,d\mathbf{P}\\
&\quad=\int_{\{I_A\mathbf{E}\{X|\mathcal{G}\}>I_A\mathbf{E}\{Y|\mathcal{H}\}\}} (\mathbf{E}\{I_AX\,|\,\mathcal{G}\}-\mathbf{E}\{I_AY\,|\,\mathcal{H}\})\,d\mathbf{P}\\
&\quad=\int_{\{I_A\mathbf{E}\{X|\mathcal{G}\}>I_A\mathbf{E}\{Y|\mathcal{H}\}\}} (I_AX-I_AY)\,d\mathbf{P}\\
&\quad=0,
\end{aligned}\tag{14.1}$$

so that $\mathbf{P}\big\{I_A\mathbf{E}\{X\,|\,\mathcal{G}\}>I_A\mathbf{E}\{Y\,|\,\mathcal{H}\}\big\}=0$ (see Exercise 148 below). By the symmetric argument we get $\mathbf{P}\big\{I_A\mathbf{E}\{Y\,|\,\mathcal{H}\}>I_A\mathbf{E}\{X\,|\,\mathcal{G}\}\big\}=0$. Hence we have $I_A\mathbf{E}\{X\,|\,\mathcal{G}\}$ $=I_A\mathbf{E}\{Y\,|\,\mathcal{H}\}$ with probability 1. □

Exercise 147. Prove that the random variables $I_A\,\mathbf{E}\{X\,|\,\mathcal{G}\}$ and $I_A\,\mathbf{E}\{Y\,|\,\mathcal{H}\}$ that feature in the proof of Lemma 14.1 are both $\mathcal{G}$-measurable and $\mathcal{H}$-measurable.

Exercise 148. Explain why (14.1) gives $\mathbf{P}\big\{I_A\,\mathbf{E}\{X\,|\,\mathcal{G}\} > I_A\,\mathbf{E}\{Y\,|\,\mathcal{H}\}\big\} = 0$.

Definition 14.2. *For a stopping time τ we define*

$$\mathcal{F}_\tau = \big\{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \ \textit{for all}\ t \geq 0\big\}.$$

Proposition 14.3. *For a stopping time τ the family $\mathcal{F}_\tau$ is a σ-algebra and τ is $\mathcal{F}_\tau$-measurable. Further, $\mathcal{F}_\tau \cap \{\tau = t\} = \mathcal{F}_t \cap \{\tau = t\}$ for $t \geq 0$. For two stopping times σ and τ we have $\mathcal{F}_\sigma \cap \{\sigma \leq \tau\} \subseteq \mathcal{F}_{\sigma\wedge\tau} = \mathcal{F}_\sigma \cap \mathcal{F}_\tau$. In particular $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$ if $\sigma \leq \tau$.*

Exercise 149. Prove Proposition 14.3.

Theorem 14.4. (Optional sampling) *For a martingale (non-negative submartingale) [non-positive supermartingale] $\{M(t)\}_{t\in[0,T]}$ and discrete stopping times $\sigma, \tau \leq T$, we have that $M(\tau)$ is integrable and $\mathbf{E}\{M(\tau)\,|\,\mathcal{F}_\sigma\} = M(\sigma\wedge\tau)$ $(\geq M(\sigma\wedge\tau))$ $[\leq M(\sigma\wedge\tau)]$.*

Proof. We prove the theorem for martingales. It will be evident from that proof how to do the proof for non-negative submartingale and non-positive supermartingales, see also Exercise 151 below. Now let $\{t_k\}_{k=1}^{\infty} \subseteq [0,T]$ be the possible values of σ and τ. The integrability of $M(\tau)$ follows readily in a by now familiar fashion as

$$\begin{aligned}
\mathbf{E}\{|M(\tau)|\} &= \sum_{k=1}^{\infty} \mathbf{E}\big\{I_{\{\tau=t_k\}}\,|M(t_k)|\big\}\\
&\leq \sum_{k=1}^{\infty} \mathbf{E}\big\{I_{\{\tau=t_k\}}\,\mathbf{E}\{|M(T)|\,|\mathcal{F}_{t_k}\}\big\}\\
&= \sum_{k=1}^{\infty} \mathbf{E}\big\{I_{\{\tau=t_k\}}\,|M(T)|\big\}\\
&= \mathbf{E}\{|M(T)|\},
\end{aligned}$$

since $|M|$ is a submartingale. Further, by Lemma 14.1 together with Proposition 14.3 (which gives the first identity below), we have

$$I_{\{\tau=t_k\}}\,\mathbf{E}\{M(T)\,|\,\mathcal{F}_\tau\} = I_{\{\tau=t_k\}}\,\mathbf{E}\{M(T)\,|\,\mathcal{F}_{t_k}\} = I_{\{\tau=t_k\}}\,M(t_k) = I_{\{\tau=t_k\}}\,M(\tau),$$

so that $\mathbf{E}\{M(T)\,|\,\mathcal{F}_\tau\} = M(\tau)$. From this in turn, by Exercise 150 below, we get using Proposition 14.3

$$\begin{aligned} I_{\{\sigma\le\tau\}}\,\mathbf{E}\{M(\tau)\,|\,\mathcal{F}_\sigma\} &= I_{\{\sigma\le\tau\}}\,\mathbf{E}\{\mathbf{E}\{M(T)\,|\,\mathcal{F}_\tau\}\,|\,\mathcal{F}_\sigma\} \\ &= I_{\{\sigma\le\tau\}}\,\mathbf{E}\{M(T)\,|\,\mathcal{F}_\sigma\} \\ &= I_{\{\sigma\le\tau\}}\,M(\sigma). \end{aligned} \tag{14.2}$$

On the other hand, using first Lemma 14.1 alone, and then Lemma 14.1 together with Proposition 14.3, we get

$$\begin{aligned} I_{\{\tau\le\sigma\}}\,\mathbf{E}\{M(\tau)\,|\,\mathcal{F}_\sigma\} &= \sum_{k=1}^{\infty}\sum_{\{\ell\,:\,t_\ell\le t_k\}} I_{\{\sigma=t_k\}}\,I_{\{\tau=t_\ell\}}\,\mathbf{E}\{M(\tau)\,|\,\mathcal{F}_\sigma\} \\ &= \sum_{k=1}^{\infty}\sum_{\{\ell\,:\,t_\ell\le t_k\}} I_{\{\sigma=t_k\}}\,I_{\{\tau=t_\ell\}}\,\mathbf{E}\{M(t_\ell)\,|\,\mathcal{F}_\sigma\} \\ &= \sum_{k=1}^{\infty}\sum_{\{\ell\,:\,t_\ell\le t_k\}} I_{\{\sigma=t_k\}}\,I_{\{\tau=t_\ell\}}\,\mathbf{E}\{M(t_\ell)\,|\,\mathcal{F}_{t_k}\} \\ &= \sum_{k=1}^{\infty}\sum_{\{\ell\,:\,t_\ell\le t_k\}} I_{\{\sigma=t_k\}}\,I_{\{\tau=t_\ell\}}\,M(t_\ell) \\ &= I_{\{\tau\le\sigma\}}\,M(\tau). \end{aligned} \tag{14.3}$$

Putting (14.2) and (14.3) together, we arrive at the statement of the theorem. □

Exercise 150. With the notation of the proof of Theorem 14.4, show that $I_{\{\sigma\le\tau\}}\,\mathbf{E}\{\mathbf{E}\{M(T)\,|\,\mathcal{F}_\tau\}\,|\,\mathcal{F}_\sigma\} = I_{\{\sigma\le\tau\}}\,\mathbf{E}\{M(T)\,|\,\mathcal{F}_\sigma\}$, e.g., by using the definition of conditional expectations together with Proposition 14.3.

Exercise 151. Modify the proof of Theorem 14.4 to work for non-negative submartingales. Explain how this gives the result for non-positive supermartingales.

Theorem 14.5. (Optional stopping) *For a continuous martingale/non-negative submartingale/non-positive supermartingal $\{M(t)\}_{t\in[0,T]}$ and a stopping time τ, the stopped process $\{M(t\wedge\tau)\}_{t\in[0,T]}$ is a continuous martingale/non-negative submartingale/non-positive supermartingal.*

Proof. We prove the theorem for martingales. It will be evident from that proof how to prove it for non-negative submartingale and non-positive supermartingales, see Exercise 153 below. Note that the continuity of the stopped process is trivial. We use the discrete approximation τ_n of τ from the proof of Theorem 6.3, see (6.2). Recall that τ_n is a stopping time such that $\tau_n\downarrow\tau$ as $n\to\infty$. Now the sequence $\{M(t\wedge\tau_n)\}_{n\in\mathbb{N}}$ is uniformly integrable by the following modification of the proof of Theorem 13.10:

$$
\begin{aligned}
&\sup_{n\in\mathbb{N}} \mathbf{E}\big\{I_{\{|M(t\wedge\tau_n)|>y\}}\,|M(t\wedge\tau_n)|\big\}\\
&\qquad\qquad = \sup_{n\in\mathbb{N}} \sum_{k=1}^{\infty} \mathbf{E}\big\{I_{\{\tau_n=2^{-n}k\}} I_{\{|M(t\wedge 2^{-n}k)|>y\}}\,|M(t\wedge 2^{-n}k)|\big\}\\
&\qquad\qquad = \sup_{n\in\mathbb{N}} \sum_{k=1}^{\infty} \mathbf{E}\big\{I_{\{\tau_n=2^{-n}k\}} I_{\{|M(t\wedge 2^{-n}k)|>y\}}\,|\mathbf{E}\{M(t)|\mathcal{F}_{2^{-n}k}\}|\big\}\\
&\qquad\qquad \leq \sup_{n\in\mathbb{N}} \sum_{k=1}^{\infty} \mathbf{E}\big\{I_{\{\tau_n=2^{-n}k\}} I_{\{|M(t\wedge 2^{-n}k)|>y\}}\,\mathbf{E}\{|M(t)||\mathcal{F}_{2^{-n}k}\}\big\}\\
&\qquad\qquad = \sup_{n\in\mathbb{N}} \sum_{k=1}^{\infty} \mathbf{E}\big\{I_{\{\tau_n=2^{-n}k\}} I_{\{|M(t\wedge 2^{-n}k)|>y\}}\,|M(t)|\big\}\\
&\qquad\qquad = \sup_{n\in\mathbb{N}} \mathbf{E}\big\{I_{\{|M(t\wedge\tau_n)|>y\}}\,|M(t)|\big\}\\
&\qquad\qquad = \sup_{n\in\mathbb{N}} \mathbf{E}\big\{I_{\{\sup_{s\in[0,T]}|M(s)|>y\}}\,|M(t)|\big\}\\
&\qquad\qquad \to 0 \quad \text{as } y\to\infty
\end{aligned}
$$

by absolute continuity, since $|M(t)|$ is integrable and $\mathbf{P}\{\sup_{s\in[0,T]}|M(s)|>y\}\to 0$ as $y\to\infty$ by the Doob-Kolmogorov inequality. And so we have $M(t\wedge\tau_n)\to M(t)$ in $\mathbb{L}^1$ as $n\to\infty$ by the continuity of M together with Theorem 13.7. Hence it is sufficient to prove that $\{M(t\wedge\tau_n)\}_{t\in[0,T]}$ is a martingale for each $n\in\mathbb{N}$, see Exercise 152 below.

Noting that s and $t\wedge\tau_n$ are bounded discrete stopping times for every choice of $0\leq s\leq t\leq T$, we get that $\{M(t\wedge\tau_n)\}_{t\in[0,T]}$ is a martingal directly from the optional sampling Theorem 14.4, as

$$\mathbf{E}\{M(t\wedge\tau_n)\,|\,\mathcal{F}_s\} = M(t\wedge\tau_n\wedge s) = M(s\wedge\tau_n) \quad \text{for } 0\leq s\leq t\leq T. \qquad \square$$

Exercise 152. Show that the martingale property for $\{M(t\wedge\tau_n)\}_{t\in[0,T]}$ for each $n\in\mathbb{N}$ together with the convergence $M(t\wedge\tau_n)\to M(t)$ in $\mathbb{L}^1$ as $n\to\infty$ imply the martingale property for $\{M(t)\}_{t\in[0,T]}$.

Exercise 153. Modify the proof of Theorem 14.5 to work for non-negative sub-martingales. Explain how this gives the result for non-positive supermartingales.

Exercise 154. (Wald's identity) Prove that for a continuous martingale $\{M(t)\}_{t\in[0,T]}$ and a stopping time $\tau\leq T$ we have $\mathbf{E}\{M(\tau)\}=\mathbf{E}\{M(0)\}$.

Exercise 155. Prove that for a martingale $\{M(t)\}_{t\geq 0}$ and a stopping time $\tau<\infty$ it does not necessarily hold that $\mathbf{E}\{M(\tau)\}=\mathbf{E}\{M(0)\}$.

We will see a crucially important application of the optional stopping theorem in the next section to prove Theorem 13.17. Here we present a first quick application:

Theorem 14.6. *If there for a continuous process $\{M(t)\}_{t\in[0,T]}$ exists a localizing sequence of stopping times $\{\tau_n\}_{n=1}^{\infty}$ such that $\{M(t\wedge\tau_n)\}_{t\in[0,T]}$ is martingale/non-negative submartingale/non-positive supermartingale for each $n\in\mathbb{N}$, then there exists another localizing sequence of stopping times $\{\tau_n'\}_{n=1}^{\infty}$ such that $\{M(t\wedge\tau_n')\}_{t\in[0,T]}$ is a martingale/non-negative submartingale/non-positive supermartingale bounded by the constant n for each $n\in\mathbb{N}$.*

Proof. Set $\tau_n' = \tau_n\wedge\sigma_n$, where $\sigma_n = \inf\{t\in[0,T] : |M(t)|\geq n\}$. Then the continuity of M together with the optional stopping theorem show that $M(t\wedge\tau_n') = M((t\wedge\tau_n)\wedge\sigma_n)$ is a martingale/non-negative submartingale/non-positive supermartingale bounded by the constant n. $\square$

Example 14.7. For a continuous process $\{M(t)\}_{t\in[0,T]}$ Theorem 13.10 says that the requirement that $\{M(t\wedge\tau_n)\}_{t\in[0,T]}$ is uniformly integrable in Definition 13.9 is void when the latter process is a martingale/non-negative submartingale/non-positive supermartingale, as such processes are always uniformly integrable. We get a knew proof of this voidness using Theorem 14.6, as that result says that we can find another localizing sequence $\{\tau_n'\}_{n=1}^{\infty}$ of stopping times such that $\{M(t\wedge\tau_n')\}_{t\in[0,T]}$ is a bounded by n martingale/non-negative submartingale/non-positive supermartingale, and thus uniformly integrable by Theorem 13.12.

14.2 Proof of Theorem 13.17

First assume that $|M(t)|, A(t)\leq N$ for $t\in[0,T]$, for some constant $N>0$, where A is the compensator of M. (We will relax this assumption later.) Consider a partition $0=t_0<t_1<\ldots<t_n=t$ of the interval $[0,t]\subseteq[0,T]$. Recall from (13.2) that

$$\mathbf{E}\bigg\{M(t_n)^2-M(t_k)^2-\sum_{j=k+1}^{n}(M(t_j)-M(t_{j-1}))^2\,\bigg|\,\mathcal{F}_{t_k}\bigg\}=0,$$

for $k=0,\ldots,n-1$, so that

$$\mathbf{E}\bigg\{\sum_{j=k+1}^{n}(M(t_j)-M(t_{j-1}))^2\,\bigg|\,\mathcal{F}_{t_k}\bigg\}=\mathbf{E}\{M(t_n)^2-M(t_k)^2|\mathcal{F}_{t_k}\}\leq N^2 \tag{14.4}$$

for $k=0,\ldots,n-1$. From repeated applications of (14.4) in turn we get

$$\begin{aligned}&\mathbf{E}\bigg\{\sum_{k=1}^{n-1}\sum_{j=k+1}^{n}(M(t_k)-M(t_{k-1}))^2(M(t_j)-M(t_{j-1}))^2\bigg\}\\ &=\sum_{k=1}^{n-1}\mathbf{E}\bigg\{(M(t_k)-M(t_{k-1}))^2\,\mathbf{E}\bigg\{\sum_{j=k+1}^{n}(M(t_j)-M(t_{j-1}))^2\,\bigg|\,\mathcal{F}_{t_k}\bigg\}\bigg\}\end{aligned}$$

$$\begin{aligned}
&\leq \sum_{k=1}^{n-1} \mathbf{E}\Big\{(M(t_k)-M(t_{k-1}))^2\Big\}\, N^2 \\
&= \mathbf{E}\Big\{\sum_{k=1}^{n-1}(M(t_k)-M(t_{k-1}))^2 \,\Big|\, \mathcal{F}_{t_0}\Big\}\, N^2 \\
&\leq N^4.
\end{aligned} \tag{14.5}$$

By another application of (14.4) we get

$$\mathbf{E}\Big\{\sum_{k=1}^{n}(M(t_k)-M(t_{k-1}))^4\Big\} \leq 4N^2\,\mathbf{E}\Big\{\sum_{k=1}^{n}(M(t_k)-M(t_{k-1}))^2 \,\Big|\, \mathcal{F}_{t_0}\Big\} \leq 4N^4. \tag{14.6}$$

Putting (14.5) and (14.6) together we get

$$\begin{aligned}
&\mathbf{E}\Big\{\Big(\sum_{k=1}^{n}(M(t_k)-M(t_{k-1}))^2\Big)^2\Big\} \\
&\quad= \mathbf{E}\Big\{\sum_{k=1}^{n}(M(t_k)-M(t_{k-1}))^4\Big\} \\
&\qquad\quad + 2\,\mathbf{E}\Big\{\sum_{k=1}^{n-1}\sum_{j=k+1}^{n}(M(t_k)-M(t_{k-1}))^2\,(M(t_j)-M(t_{j-1}))^2\Big\} \\
&\quad\leq 6N^4.
\end{aligned} \tag{14.7}$$

From this in turn together with Hölder's inequality we get

$$\begin{aligned}
&\mathbf{E}\Big\{\sum_{k=1}^{n}(M(t_k)-M(t_{k-1}))^4\Big\} \\
&\leq \mathbf{E}\Big\{\Big(\sup_{r,s\in[0,T]:|r-s|\leq \max_{1\leq j\leq n} t_j-t_{j-1}} |M(r)-M(s)|\Big)^2 \sum_{k=1}^{n}(M(t_k)-M(t_{k-1}))^2\Big\} \\
&\leq \sqrt{\mathbf{E}\Big\{\Big(\sup_{r,s\in[0,T]:|r-s|\leq \max_{1\leq j\leq n} t_j-t_{j-1}} |M(r)-M(s)|\Big)^4\Big\}\,\mathbf{E}\Big\{\Big(\sum_{k=1}^{n}(M(t_k)-M(t_{k-1}))^2\Big)^2\Big\}} \\
&\to 0 \quad \text{as} \quad \max_{1\leq j\leq n} t_j-t_{j-1} \downarrow 0,
\end{aligned} \tag{14.8}$$

since the first expectation inside the square root goes to zero by continuity of M and bounded convergence, while the second expectation inside the square root is bounded by $6N^4$. Now we may finish off the proof (for M and A bounded by N) as follows:

$$\begin{aligned}
&\mathbf{E}\Big\{\Big(\sum_{k=1}^{n}(M(t_k)-M(t_{k-1}))^2 - A(t)\Big)^2\Big\} \\
&= \mathbf{E}\Big\{\Big(\sum_{k=1}^{n}(M(t_k)-M(t_{k-1}))^2 - (A(t_k)-A(t_{k-1}))\Big)^2\Big\} \\
&= \mathbf{E}\Big\{\sum_{k=1}^{n}\big((M(t_k)-M(t_{k-1}))^2 - (A(t_k)-A(t_{k-1}))\big)^2\Big\}
\end{aligned} \tag{14.9}$$

$$\le 2\,\mathbf{E}\Big\{\sum_{k=1}^n (M(t_k)-M(t_{k-1}))^4\Big\} + 2\,\mathbf{E}\Big\{\sum_{k=1}^n (A(t_k)-A(t_{k-1}))^2\Big\}$$

$$\le 2\,\mathbf{E}\Big\{\sum_{k=1}^n (M(t_k)-M(t_{k-1}))^4\Big\} + 2\,\mathbf{E}\Big\{\Big(\sup_{r,s\in[0,T]:|r-s|\le \max_{1\le j\le n} t_j-t_{j-1}} |A(r)-A(s)|\Big) A(t)\Big\},$$

where the first term on the right-hand side goes to zero as $\max_{1\le j\le n} t_j - t_{j-1}\downarrow 0$ by (14.8), while the second term goes to zero by continuity of A together with bounded convergence. [See Exercise 156 below on the second equality in (14.9).]

Of course, to adress the general case when $|M|$ and A need not be bounded by a constant N, we stop the processes at their first contact with the level N, that is, at

$$\tau_N = \inf\{t\in[0,T]: |M(t)|\ge N\}\wedge \inf\{t\in[0,T]: A(t)\ge N\} \quad \text{for } N\in\mathbb{N}.$$

Here τ_N is a stopping time since M and A are continuous and adapted. Now $\{M(t\wedge \tau_N)\}_{t\in[0,T]}$ is a bounded by N martingale by the optional stopping theorem, while $\{M(t\wedge\tau_N)^2 - A(t\wedge\tau_N)\}_{t\in[0,T]}$ is a bounded by N^2 martingale. As the first part of the proof only uses that M and M^2-A are bounded martingales with A bounded, we can use the first part of the proof to conclude that $\{M(t\wedge\tau_N)\}_{t\in[0,T]}$ has quadratic variation $\{A(t\wedge\tau_N)\}_{t\in[0,T]}$. Given a constant $\varepsilon>0$ we now have

$$\limsup_{\max_{1\le j\le n} t_j-t_{j-1}\downarrow 0} \mathbf{P}\Big\{\Big|\sum_{k=1}^n (M(t_k)-M(t_{k-1}))^2 - A(t)\Big| > \varepsilon\Big\}$$

$$\le \mathbf{P}\{\tau_N < T\} + \limsup_{\max_{1\le j\le n} t_j-t_{j-1}\downarrow 0} \mathbf{P}\Big\{\Big|\sum_{k=1}^n (M(t_k\wedge\tau_N)-M(t_{k-1}\wedge\tau_N))^2 - A(t\wedge\tau_N)\Big| > \varepsilon\Big\}$$

$$= \mathbf{P}\{\tau_N < T\} \quad \text{for } N\in\mathbb{N}.$$

As the lim sup on the left-hand side does not depend on N, and as $\mathbf{P}\{\tau_N < T\}\to 0$ as $N\to\infty$ by continuity of M and A, it follows that the lim sup on the left-hand is zero. This in turn is the claim of the theorem to prove, that is, M has qudratic variation process A in the sense of convergence in probability. □

Exercise 156. Prove the second equality in (14.9).

Exercise 157. (Burkholder-Davis-Gundy inequality) Prove that for a square-integrable continuous martingale $\{M(t)\}_{t\in[0,T]}$ with $M(0)=0$, we have $\mathbf{E}\{[M](T)\}\le \mathbf{E}\{\sup_{t\in[0,T]} M(t)^2\}\le 4\,\mathbf{E}\{[M](T)\}$.

15 Stochastic integration with respect to continuous martingales

Here we construct the Itô integral process $\{\int_0^t X\,dM\}_{t\in[0,T]}$ with respect to a continuous local martingale M. We first carry out the construction for square-integrable continuous martingales M, in which case it is the same in essence as that for the integral with respect to BM. The only difference is that we must replace the quadratic variation process $[B](t)=t$ of BM with that of M. The intergal is then extended by means of stopping methods to continuous local martingales.

It would have been more economical to construct the Itô integral for martigales already from the beginning and skip the intermediate step with the integral for BM. But then we would have had to start the course with several lectures on pure martingale theory, so that we would not have been able to treat existence of strong solutions within the framework of the first part of the course.

As usual, we assume the usual conditions. (We do not really need the BM that comes with those conditions now for a good while, albeit it doesn't hurt to have it.)

15.1 Itô integrals for the space S_T

In the same way as for BM, the construction of the integral will be in three steps S, E and P. Here the first step is exactly as before (cf. Definition 4.1):

Definition 15.1. *The* Itô integral process $\{\int_0^t X\,dM\}_{t\in[0,T]}$ *of a simple process* $X\in S_T$ *given by equation* (3.6) *in Definition 3.11 with respect to a continuous square-integrable martingale* $\{M(t)\}_{t\in[0,T]}$ *is defined by* $\int_0^0 X\,dM=0$ *and*

$$\int_0^t X\,dM=\sum_{i=1}^{m} X_{t_{i-1}}\,(M(t_i)-M(t_{i-1}))+X_{t_m}\,(M(t)-M(t_m))\quad \textit{for}\ t\in(t_m,t_{m+1}],$$

for $m=0,\ldots,n-1$. *Further, we define*

$$\int_s^t X\,dM=\int_0^t X\,dM-\int_0^s X\,dM\quad \textit{for}\ s,t\in[0,T].$$

(Consistency) Recall from Exercise 41 that Definition 15.1 is consistent in the sense that if $X\in S_T$ has two representations

$$X(t)=X(0)I_{\{0\}}(t)+\sum_{i=1}^{m} X_{t_{i-1}}I_{(t_{i-1},t_i]}(t)=X(0)I_{\{0\}}(t)+\sum_{i=1}^{n} X'_{t'_{i-1}}I_{(t'_{i-1},t'_i]}(t)$$

for $t\in[0,T]$, then by means of introducing a third grid that contains all times of the grids $0=t_0<t_1<\ldots<t_m=T$ and $0=t'_0<t'_1<\ldots<t'_n=T$, we see that the Itô integral processes $\{\int_0^t X\,dM\}_{t\in[0,T]}$ for the two representations of X coincide.

Theorem 15.2. (PROPERTIES OF ITÔ INTEGRALS FOR THE SPACE S_T) *Let $\{M(t)\}_{t\in[0,T]}$ be a continuous square-integrable martingale.*

(CONTINUITY) *For $X\in S_T$ the process $\{\int_0^t X\,dM\}_{t\in[0,T]}$ is continuous.*

(NADA) *For $X\in S_T$ we have*

$$\int_r^t X\,dM = -\int_t^r X\,dM \quad \text{and} \quad \int_0^t I_{[r,s]}X\,dM = \int_r^s X\,dM \quad \text{for } 0\le r\le s\le t\le T.$$

(LINEARITY) *For $X,Y\in S_T$ and constants $a,b\in\mathbb{R}$ we have*

$$\int_0^t (aX+bY)\,dM = a\int_0^t X\,dM + b\int_0^t Y\,dM \quad \text{for } t\in[0,T].$$

(ADAPTEDNESS) *For $X\in S_T$ the process $\{\int_0^t X\,dM\}_{t\in[0,T]}$ is adapted.*

(MARTINGALE) *For $X\in S_T$, $\{\int_0^t X\,dM\}_{t\in[0,T]}$ is a square-integrable martingale.*

(ZERO-MEAN) *For $X\in S_T$ the process $\{\int_0^t X\,dM\}_{t\in[0,T]}$ has zero mean.*

(ISOMETRY) *For $X,Y\in S_T$ we have*

$$\mathbf{E}\Big\{\Big(\int_0^t X\,dM\Big)\Big(\int_0^t Y\,dM\Big)\Big\} = \mathbf{E}\Big\{\int_0^t XY\,d[M]\Big\} \quad \text{for } t\in[0,T].$$

Proof. The continuity, nada, linearity, adaptedness and zero-mean properties are verified exactly as in Exercises 42-46 which establish the corresponding properties for Itô integrals with respect to BM – just replace every occurance of B with M.

As for the martingale property, in the same fashion we need only replace B with M at all occurances in the proof of Theorem 4.2. In particular, given $0=t_0<t_1<\ldots<s=t_j<\ldots<t_n=T$, the key identity (4.1) of that proof changes to

$$\begin{aligned}\mathbf{E}\Big\{\sum_{i=j+1}^{m+1} X_{t_{i-1}}(M(t_i)-M(t_{i-1}))\,\Big|\,\mathcal{F}_s\Big\} &= \sum_{i=j+1}^{m+1}\mathbf{E}\Big\{X_{t_{i-1}}\,\mathbf{E}\{M(t_i)-M(t_{i-1})\big|\mathcal{F}_{t_{i-1}}\}\,\Big|\,\mathcal{F}_s\Big\}\\ &= 0.\end{aligned}$$

Moving over to the isometry property, recall from Theorem 13.17 and (13.1) that a continuous square-integrable martingale $\{M(t)\}_{t\in[0,T]}$ has a well-defined continuous integrable quadratic variation process $\{[M](t)\}_{t\in[0,T]}$ that satisfies (15.1) below. Therefore the only thing that is required for the proof of isometry is again to replace B with M at every occurance in the proof of the corresponding result Theorem 4.3 for stochastic integrals for S_T with respect to BM. In particular, the indentity $\mathbf{E}\{B(t_i)-B(t_{i-1})\}=0$ of that proof is replaced with $\mathbf{E}\{M(t_i)-M(t_{i-1}\}=0$, while the identity $\mathbf{E}\{(B(t_i)-B(t_{i-1}))^2\}=t_i-t_{i-1}$ is replaced with

$$\mathbf{E}\big\{(M(t_i)-M(t_{i-1}))^2\big\} = [M](t_i)-[M](t_{i-1}). \qquad \square \qquad (15.1)$$

Exercise 158. Prove in full detail one of the properties martingale or isometry in Theorem 15.2.

15.2 Itô integrals for the space $E(M)_T$

The next step E must pay attention to what is the quadratic variation:

Definition 15.3. *Given a constant $T > 0$ and a continuous square-integrable martingale $\{M(t)\}_{t\in[0,T]}$ we say that a stochastic process $\{X(t)\}_{t\geq 0}$ belongs to the class $E(M)_T$, if X is measurable and adapted with*

$$\mathbf{E}\bigg\{\int_0^T X^2\, d[M]\bigg\} < \infty.$$

Exercise 159. In Definition 15.3, show that $\int_0^T X^2\, d[M]$ is a random variable.

We have the following version of Theorem 4.4 for the space $E(M)_T$:

Theorem 15.4. *For $X \in E(M)_T$ there exists a sequence $\{X_n\}_{n=1}^{\infty} \subseteq S_T$ such that*

$$\lim_{n\to\infty} \mathbf{E}\bigg\{\int_0^T (X_n - X)^2\, d[M]\bigg\} = 0. \tag{15.2}$$

Proof. Copy-and-paste to here the proof of Theorem 4.4. Then replace every occurance of dr (integral of a function or stochastic process of the argument r with respect to Lebesgue measure) that features in that proof with $d[M](r)$ (integral with respect to the quadratic variation process of M). Using that $\{[M](t)\}_{t\in[0,T]}$ is continuous increasing and integrable the proof is complete by inspection. [The integral for the Y process $Y(t) = \int_0^t X^{(N)}(s)\, ds$ and the integral for the Z_n process in (5.8) in the last paragraph of the proof shall remain ds integrals and not be changed.] □

Definition and Theorem 15.5. *For a continuous square-integrable martingale $\{M(t)\}_{t\in[0,T]}$ and an $X \in E(M)_T$, the* Itô integral process *$\{\int_0^t X\, dM\}_{t\in[0,T]}$ is well-defined as the unique up to version and continuous with probability 1 stochastic process that is given as the limit in the sense of convergence in mean-square by*

$$\int_0^t X\, dM \leftarrow \int_0^t X_n\, dM \quad \text{as } n\to\infty \text{ for } t \in [0,T], \tag{15.3}$$

where $\{X_n\}_{n=1}^{\infty} \subseteq S_T$ satisfies (15.2). Further, we define

$$\int_s^t X\, dM = \int_0^t X\, dM - \int_0^s X\, dM \quad \text{for } s,t \in [0,T].$$

Proof. Copy-and-paste to here the proof of Definition and Theorem 4.5. Then replace all occurances of B and dr with M and $d[M](r)$, respectively. Also, replace references to (4.2) and (4.3) with references to (15.2) and (15.3), respectively. □

The properties of Itô integrals for the space E_T carry over to the space $E(M)_T$:

Theorem 15.6. (PROPERTIES OF ITÔ INTEGRALS FOR THE SPACE $E(M)_T$) *Let $\{M(t)\}_{t\in[0,T]}$ be a continuous square-integrable martingale.*

(NADA) *For $X \in E(M)_T$ we have*

$$\int_r^t X\,dM = -\int_t^r X\,dM \quad \textit{and} \quad \int_0^t I_{[r,s]} X\,dM = \int_r^s X\,dM \quad \textit{for } 0 \le r \le s \le t \le T.$$

(LINEARITY) *For $X, Y \in E(M)_T$ and constants $a, b \in \mathbb{R}$ we have*

$$\int_0^t (aX + bY)\,dM = a\int_0^t X\,dM + b\int_0^t Y\,dM \quad \textit{for } t \in [0,T].$$

(ADAPTEDNESS) *For $X \in E(M)_T$ the process $\{\int_0^t X\,dM\}_{t\in[0,T]}$ is adapted.*

(MARTINGALE) *For $X \in E(M)_T$, $\{\int_0^t X\,dM\}_{t\in[0,T]}$ is a square-integrable martingale.*

(ZERO-MEAN) *For $X \in E(M)_T$ the process $\{\int_0^t X\,dM\}_{t\in[0,T]}$ has zero mean.*

(ISOMETRY) *For $X, Y \in E(M)_T$ we have*

$$\mathbf{E}\Big\{\Big(\int_0^t X\,dM\Big)\Big(\int_0^t Y\,dM\Big)\Big\} = \mathbf{E}\Big\{\int_0^t XY\,d[M]\Big\} \quad \textit{for } t \in [0,T].$$

(CONVERGENCE) *If $X \in E(M)_T$ and $\{X_n\}_{n=1}^\infty \subseteq E(M)_T$ satisfy*

$$\lim_{n\to\infty} \mathbf{E}\Big\{\int_0^T \big(X_n(r) - X(r)\big)^2\,d[M](r)\Big\} = 0,$$

then we have in the sense of convergence in mean-square

$$\int_0^t X_n\,dM \to \int_0^t X\,dM \quad \textit{as } n \to \infty \textit{ for } t \in [0,T].$$

(ZIPP) *For $X \in E(M)_T$ and an $\mathcal{F}_s$-measurable random variable Y is that is bounded by a (non-random) constant we have $\int_0^t I_{[s,t]}(r)\,Y X(r)\,dM(r) = Y\int_s^t X(r)\,dM(r)$ for $0 \le s \le t \le T$.*

(STOPPING) *For $X \in E(M)_T$ and a stopping time τ we have $\{I_{[0,\tau]}(t)X(t)\}_{t\in[0,T]} \in E(M)_T$ and $\int_0^{t\wedge\tau} X\,dM = \int_0^t I_{[0,\tau]} X\,dM$ for $t \in [0,T]$.*

Proof. The proofs of all these properties for the space E_T in Exercises 49-52, Exercises 54-57, and Theorem 6.3, respectively, carry over to the space $E(M)_T$ with only obvious modifications: Just replace B with M everywhere and use Theorem 15.2 whenever a property for the integral on the space S_T is requested. □

Exercise 160. Prove in full detail one of the properties isometry, martingale or stopping in Theorem 15.6.

15.3 Quadratic variation for continuous local martingales

We show by stopping that continuous local martingales have quadratic variations.

Theorem 15.7. *A continuous local martingale $\{M(t)\}_{t\in[0,T]}$ has a well-defined continous with probability 1 quadratic variation process given by $[M](t)=[M]_m(t)$ for $t\in[0,T]$ on the event $\{\tau_m\geq T\}$, where $\{[M]_m(t)\}_{t\in[0,T]}$ is the quadratic variation of $\{M(t\wedge\tau_m)\}_{t\in[0,T]}$ and $\{\tau_m\}_{m=1}^{\infty}$ is a localizing sequence of stopping times such that $\{M(t\wedge\tau_m)\}_{t\in[0,T]}$ is a square-integrable continuous martingale for $m\in\mathbb{N}$.*

Proof. Theorem 14.6 gives the existence of a localizing sequence $\{\tau_m\}_{m=1}^{\infty}$ such that $\{M(t\wedge\tau_m)\}_{t\in[0,T]}$ is square-integrable for $m\in\mathbb{N}$. The quadratic variations $[M]_m$ and $[M]_n$ of $M(\,\cdot\wedge\tau_m)$ and $M(\,\cdot\wedge\tau_n)$, respectively, make $M(\,\cdot\wedge\tau_m)^2-[M]_m$ and $M(\,\cdot\wedge\tau_n)^2-[M]_n$ continuous martingales, and are uniquely determined by that property. As $\{M(t\wedge\tau_m)\}_{t\in[0,T]}$ and $\{M(t\wedge\tau_n)\}_{t\in[0,T]}$ agree on the event $\{\tau_m\wedge\tau_n\geq T\}$, it follows that $\{[M]_m(t)\}_{t\in[0,T]}$ and $\{[M]_n(t)\}_{t\in[0,T]}$ agree on that event. Hence the definition $[M](t)=[M]_m(t)$ for $t\in[0,T]$ on the event $\{\tau_m\geq T\}$ is consistent (non-multivalued). Here $[M]$ is continuous on the almost sure event $\bigcup_{m=1}^{\infty}\{\tau_m\geq T\}$, as each of the processes $[M]_m$ are continuous. To see that $[M]$ is the quadratic variation of M we consider partitions $0=t_0<t_1<\ldots<t_n=t$ of the interval $[0,t]\subseteq[0,T]$ and note that

$$\limsup_{\max_{1\leq j\leq n}t_j-t_{j-1}\downarrow 0}\mathbf{P}\bigg\{\bigg|\sum_{k=1}^{n}(M(t_k)-M(t_{k-1}))^2-[M](t)\bigg|>\varepsilon\bigg\}$$
$$\leq\limsup_{\max_{1\leq j\leq n}t_j-t_{j-1}\downarrow 0}\mathbf{P}\bigg\{\bigg|\sum_{k=1}^{n}(M(t_k\wedge\tau_m)-M(t_{k-1}\wedge\tau_m))^2-[M]_m(t)\bigg|>\varepsilon\bigg\}+\mathbf{P}\{\tau_m<T\}$$
$$=0+\mathbf{P}\{\tau_m<T\}\quad\text{for each } m\in\mathbb{N},\quad\text{for }\varepsilon>0.$$

It follows that the lim sup on the left-hand is zero for each $\varepsilon>0$, as it does not depend on m and the probability on the right-hand side goes to 0 as $m\to\infty$. □

Exercise 161. With the notation of Theorem 15.7, show that $[M]_m(t)\to[M](t)$ in the sense of convergence in probability for $t\in[0,T]$ as $m\to\infty$.

Corollary 15.8. *A continuous local martingale $\{M(t)\}_{t\in[0,T]}$ has a well-defined continuous quadratic variation process $\{[M](t)\}_{t\in[0,T]}$ which is the unique non-decreasing (adapted) process that makes $\{M(t)^2-[M](t)\}_{t\in[0,T]}$ a continuous local martingale.*

Proof. With the notation of the proof of Theorem 15.7 we have

$$\begin{aligned} M^2(t\wedge\tau_n) - [M](t\wedge\tau_n) &= (M^2(\,\cdot\,\wedge\tau_n))(t\wedge\tau_n) - [M]_m(t\wedge\tau_n) \quad \text{for some } m\geq n \\ &= (M^2(\,\cdot\,\wedge\tau_n))(t\wedge\tau_n) - [M(\,\cdot\,\wedge\tau_m)](t\wedge\tau_n) \\ &= (M^2(\,\cdot\,\wedge\tau_n))(t\wedge\tau_n) - [M(\,\cdot\,\wedge\tau_n)](t\wedge\tau_n) \quad \text{for } t\in[0,T] \\ &= \text{martingale}. \end{aligned}$$

As this martingale is uniformly integrable by Theorem 13.10, it follows that $\{M(t)^2 - [M](t)\}_{t\in[0,T]}$ is a continuous local martingale.

Let $\{A(t)\}_{t\in[0,T]}$ be a non-decreasing process such that $\{M(t)^2 - A(t)\}_{t\in[0,T]}$ is a continuous local martingale. Then A is continuous and by the optional stopping theorem together with Theorem 14.6 we may find a localizing sequence of stopping times $\{\tau_n\}_{n=1}^{\infty}$ such that $\{M(t\wedge\tau_n)^2 - [M](t\wedge\tau_n)\}_{t\in[0,T]}$ and $\{M(t\wedge\tau_n)^2 - A(t\wedge\tau_n)\}_{t\in[0,T]}$ are continuous martingales bounded by n for each $n\in\mathbb{N}$. Hence $\{[M](t\wedge\tau_n) - A(t\wedge\tau_n)\}_{t\in[0,T]}$ is a bounded by $2n$ continuous martingale with finite variation. As such it has zero quadratic variation, so that it is zero by Exercise 157. Hence $\{[M](t\wedge\tau_n)\}_{t\in[0,T]}$ and $\{A(t\wedge\tau_n)\}_{t\in[0,T]}$ agree for $n\in\mathbb{N}$, so that $[M]$ and A agree by sending $n\to\infty$. □

Corollary 15.9. *A pair of continuous local martingales $\{M(t)\}_{t\in[0,T]}$ and $\{N(t)\}_{t\in[0,T]}$ has a well-defined continuous quadratic covariation process $\{[M,N](t)\}_{t\in[0,T]}$ which has the property that $\{M(t)N(t) - [M,N](t)\}_{t\in[0,T]}$ is a continuous local martingale.*

Exercise 162. Prove Corollary 15.9.

Exercise 163. Let $\{M(t)\}_{t\in[0,T]}$ be a continuous local martingale and τ a stopping time. Show that $[M(\,\cdot\,\wedge\tau)](t) = [M(\,\cdot\,\wedge\tau)](t\wedge\tau) = [M](t\wedge\tau)$ for $t\in[0,T]$.

15.4 Itô integrals for the space $P(M)_T$

The following definition should not come as a suprise.

Definition 15.10. *Given a constant $T>0$ and a continuous local martingale $\{M(t)\}_{t\in[0,T]}$ we say that a stochastic process $\{X(t)\}_{t\geq 0}$ belongs to the class $P(M)_T$ of* predictable processes on $[0,T]$, *if X is measurable and adapted with*

$$\mathbf{P}\bigg\{\int_0^T X^2\, d[M] < \infty\bigg\} = 1.$$

We have the following versions of Theorems 5.1 and 5.2:

Theorem 15.11. *For a continuous square-integrable martingale $\{M(t)\}_{t\in[0,T]}$ and an $X \in P(M)_T$, we have in the sense of convergence in probability*

$$\lim_{n\to\infty} \int_0^T (X_n - X)^2\, d[M] = 0 \quad \textit{for some sequence } \{X_n\}_{n=1}^\infty \subseteq E(M)_T. \tag{15.4}$$

Theorem 15.12. *For a continuous square-integrable martingale $\{M(t)\}_{t\in[0,T]}$, an $X \in E(M)_T$ and a constant $C > 0$, we have*

$$\mathbf{P}\bigg\{\sup_{t\in[0,T]}\bigg|\int_0^t X\,dM\bigg| > \lambda\bigg\} \le \frac{C}{\lambda^2} + \mathbf{P}\bigg\{\int_0^T X(r)^2\, d[M](r) \ge C\bigg\} \quad \textit{for } \lambda > 0. \tag{15.5}$$

Proof of Theorem 15.11. Copy-and-paste to here the proof of Theorem 5.1. Replace every occurance of dr in that proof with $d[M](r)$. Also replace references to the spaces E_T and P_T with references to the spaces $E(M)_T$ and $P(M)_T$, respectively. Noting that the process $\{\int_0^t X^2\, d[M]\}_{t\in[0,T]}$ is adapted by a modification of Lemma 6.4, see Exercise 164 below, as well as continuous, as $[M]$ is continuous, see Exercise 164 below, the theorem follows by inspection. □

Exercise 164. Show that the process $\{\int_0^t X^2\, d[M]\}_{t\in[0,T]}$ is adapted and continuous for $\{M(t)\}_{t\in[0,T]}$ a continuous square-integrable martingale and $X \in P(M)_T$.

Exercise 165. Prove Theorem 15.12.

Definition and Theorem 15.13. *For a continuous square-integrable martingale $\{M(t)\}_{t\in[0,T]}$ and an $X \in P(M)_T$, the* Itô integral process $\{\int_0^t X\, dM\}_{t\in[0,T]}$ *is well-defined as the unique up to version and continuous with probability 1 stochastic process that is given as the limit in the sense of convergence in probability by*

$$\int_0^t X\,dM \leftarrow \int_0^t X_n\,dM \quad \textit{as } n\to\infty \textit{ for } t \in [0,T], \tag{15.6}$$

where $\{X_n\}_{n=1}^\infty \subseteq E(M)_T$ satisfies (15.4). *Further, we define*

$$\int_s^t X\,dM = \int_0^t X\,dM - \int_0^s X\,dM \quad \textit{for } s,t \in [0,T].$$

Proof. Copy-and-paste to here the proof of Definition and Theorem 5.3. Replace all occurances of B and dr in that proof with M and $d[M](r)$, respectively. Also replace references to equations (5.10), (5.11) and (5.12) with references to equations (15.4), (15.5) and (15.6), respectively. By inspection this proves the theorem. □

For future use we prove the following property new to us:

Theorem 15.14. (ZORRO) *For a continuous square-integrable martingale* $\{M(t)\}_{t\in[0,T]}$*, an* $X\in P(M)_T$ *and a stopping time* τ*, we have* $X\in P(M(\cdot\wedge\tau))_T$ *and* $I_{[0,\tau]}X \in P(M)_T$ *as well as* $\int_0^t X\,dM(\,\cdot\wedge\tau)=\int_0^t I_{[0,\tau]}X\,dM$ *for* $t\in[0,T]$.

Proof. By inspection of Definition 15.1 we readily see that

$$\left\{\int_0^t X\,dM(\,\cdot\wedge\tau)\right\}_{t\in[0,T]}=\left\{\int_0^{t\wedge\tau} X\,dM\right\}_{t\in[0,T]}\quad\text{for } X\in S_T. \tag{15.7}$$

For an $X\in P(M)_T$ we get $X\in P(M(\,\cdot\wedge\tau))_T$ and $I_{[0,\tau]}X\in P(M)_T$ from

$$\int_0^T X^2\,d[M]\geq\int_0^{T\wedge\tau}X^2\,d[M]=\int_0^T I_{[0,\tau]}X^2\,d[M]=\int_0^T X^2\,d[M(\,\cdot\wedge\tau)]. \tag{15.8}$$

Further, if $\{X_n\}_{n=1}^\infty\subseteq S_T$ satisfies (15.2) for an $X\in E(M)_T$, then (15.8) shows that

$$\mathbf{E}\left\{\int_0^T\left(I_{[0,\tau]}X_n-I_{[0,\tau]}X\right)^2d[M]\right\}=\mathbf{E}\left\{\int_0^T(X_n-X)^2\,d[M(\,\cdot\wedge\tau)]\right\}\to 0\quad\text{as } n\to\infty.$$

From this and Definition and Theorem 15.5 together with (15.7) and the stopping and convergence properties for $E(M)_T$ we get the theorem for $X\in E(M)_T$ as

$$\int_0^t X\,dM(\,\cdot\wedge\tau)\leftarrow\int_0^t X_n\,dM(\,\cdot\wedge\tau)=\int_0^{t\wedge\tau}X_n\,dM=\int_0^t I_{[0,\tau]}X_n\,dM\to\int_0^t I_{[0,\tau]}X\,dM$$

in the sense of convergence in mean-square as $n\to\infty$ for $t\in[0,T]$. From (15.8) we futher see that if $\{X_n\}_{n=1}^\infty\subseteq E(M)_T$ satisfies (15.4) for an $X\in P(M)_T$, then we have

$$\int_0^T\left(I_{[0,\tau]}X_n-I_{[0,\tau]}X\right)^2d[M]=\int_0^T(X_n-X)^2\,d[M(\,\cdot\wedge\tau)]\to 0\quad\text{as } n\to\infty$$

with convergence in probability. From this and Definition and Theorem 15.13 together with the theorem for $X\in E(M)_T$ we get the theorem for $X\in P(M)_T$ as

$$\int_0^t X\,dM(\,\cdot\wedge\tau)\leftarrow\int_0^t X_n\,dM(\,\cdot\wedge\tau)=\int_0^t I_{[0,\tau]}X_n\,dM\to\int_0^t I_{[0,\tau]}X\,dM$$

in the sense of convergence in probability as $n\to\infty$ for $t\in[0,T]$. □

Besides the new zorro property, the other properties of Itô integrals for the space $P(M)_T$ are the expected ones:

Theorem 15.15. (PROPERTIES OF ITÔ INTEGRALS FOR THE SPACE $P(M)_T$) *Let $\{M(t)\}_{t\in[0,T]}$ be a continuous square-integrable martingale.*

(CONVERGENCE) *If $X \in P(M)_T$ and $\{X_n\}_{n=1}^{\infty} \subseteq P(M)_T$ satisfy*

$$\int_0^T \big(X_n(r) - X(r)\big)^2 \, d[M](r) \to 0 \quad \text{as } n \to \infty$$

in the sense of convergence in probability, then we have

$$\int_0^t X_n \, dM \to \int_0^t X \, dM \quad \text{in probability as } n \to \infty \text{ for } t \in [0,T].$$

(NADA) *For $X \in P(M)_T$ we have*

$$\int_r^t X \, dM = -\int_t^r X \, dM \quad \text{and} \quad \int_0^t I_{[r,s]} X \, dM = \int_r^s X \, dM \quad \text{for } 0 \le r \le s \le t \le T.$$

(LINEARITY) *For $X, Y \in P(M)_T$ and constants $a, b \in \mathbb{R}$ we have*

$$\int_0^t (aX + bY) \, dM = a \int_0^t X \, dM + b \int_0^t Y \, dM \quad \text{for } t \in [0,T].$$

(ADAPTEDNESS) *For $X \in P(M)_T$ the process $\{\int_0^t X \, dM\}_{t\in[0,T]}$ is adapted.*

(LOCAL MARTINGALE) *For $X \in P(M)_T$, $\{\int_0^t X \, dM\}_{t\in[0,T]}$ is a local martingale.*

(ZIPP) *For $X \in P(M)_T$ and an $\mathcal{F}_s$-measurable random variable Y is that is bounded by a (non-random) constant we have $\int_0^t I_{[s,t]}(r) Y X(r) \, dM(r) = Y \int_s^t X(r) \, dM(r)$ for $0 \le s \le t \le T$.*

(STOPPING) *For $X \in P(M)_T$ and a stopping time τ we have $I_{[0,\tau]} X \in P(M)_T$ and $\int_0^{t\wedge\tau} X \, dM = \int_0^t I_{[0,\tau]} X \, dM$ for $t \in [0,T]$.*

Proof. The nada, linearity, adaptedness and zipp properties follow more or less immediately from that $P(M)_T$ integrals are convergence in probability limits of $E(M)_T$ integrals, see also Exercises 74-77. The proof of the convergence property is done exactly as in the proof of Theorem 5.4 for the space P_T – just replace every occurence of B and dr with M and $d[M](r)$, respectively. The stopping property is proved by a straightforward modification of the proof of Theorem 6.3 for the space E_T, making use of the newly established convergence, nada and zipp properties for the space $P(M)_T$, see also Exercise 78. Finally, the local martingale property is proved by a straightforward modification of the proof of Theorem 13.13 for the space P_T, making use of the newly established stopping property. □

Exercise 166. Prove in full detail one of the properties convergence, stopping or local martingale in Theorem 15.15.

The time has come to extend the integral to its final local martingale generality:

Definition and Theorem 15.16. *For a continuous local martingale* $\{M(t)\}_{t\in[0,T]}$ *and an* $X\in P(M)_T$ *the* Itô integral process $\{\int_0^t X\,dM\}_{t\in[0,T]}$ *is well-defined as the unique up to version and continuous with probability 1 stochastic process given by*

$$\int_0^t X\,dM = \int_0^t X\,dM(\,\cdot\wedge\tau_m) \quad \textit{for } t\in[0,T] \textit{ on the event } \{\tau_m\geq T\},$$

where $\{\tau_m\}_{m=1}^{\infty}$ *is a localizing sequence of stopping times such that* $\{M(t\wedge\tau_m)\}_{t\in[0,T]}$ *is a square-integrable continuous martingale for* $m\in\mathbb{N}$.

Proof. Theorem 14.6 gives the existence of a localizing sequence $\{\tau_m\}_{m=1}^{\infty}$ such that $\{M(t\wedge\tau_m)\}_{t\in[0,T]}$ is square-integrable for $m\in\mathbb{N}$. By application of the zorro and stopping properties in Theorems 15.14 and 15.15 it follows that

$$\int_0^{t\wedge\tau_n} X\,dM(\,\cdot\wedge\tau_m) = \int_0^t X\,dM(\,\cdot\wedge\tau_m\wedge\tau_n) = \int_0^{t\wedge\tau_m} X\,dM(\,\cdot\wedge\tau_n). \tag{15.9}$$

Hence $\int_0^t X\,dM(\,\cdot\wedge\tau_m) = \int_0^t X\,dM(\,\cdot\wedge\tau_n)$ for $t\in[0,T]$ on the event $\{\tau_m\wedge\tau_n\geq T\}$, so that the definition of $\int_0^t X\,dM$ is consistent (non-multivalued).

Let $\{\tau'_m\}_{m=1}^{\infty}$ be another localizing sequence making $\{M(t\wedge\tau'_m)\}_{t\in[0,T]}$ a square-integrable martingale for $m\in\mathbb{N}$. We must show that the integral given by $\oint_0^t X\,dM = \int_0^t X\,dM(\,\cdot\wedge\tau'_m)$ for $t\in[0,T]$ on the event $\{\tau'_m\geq T\}$ agrees with the integral $\{\int_0^t X\,dM\{_{t\in[0,T]}$. However, changing τ_n to τ'_m in (15.9), we see that $\int_0^t X\,dM = \oint_0^t X\,dM$ for $t\in[0,T]$ on the event $\{\tau_m\wedge\tau'_m\geq T\}$. Sending $m\to\infty$ this gives $\{\int_0^t X\,dM\}_{t\in[0,T]} = \{\oint_0^t X\,dM\}_{t\in[0,T]}$. □

Exercise 167. With the hypothesis and notation of Definition and Theorem 15.16, show that $\int_0^t X\,dM(\,\cdot\wedge\tau_m)\to\int_0^t X\,dM$ in probability as $m\to\infty$ uniformly for $X\in P(M)_T$ and $t\in[0,T]$ in the sense that

$$\lim_{m\to\infty}\mathbf{P}\bigg\{\sup_{t\in[0,T],\,X\in P(M)_T}\bigg|\int_0^t X\,dM(\,\cdot\wedge\tau_m)-\int_0^t X\,dM\bigg|>\varepsilon\bigg\}=0 \quad \text{for each } \varepsilon>0.$$

Theorem 15.17. (Properties of Itô integrals for the space $P(M)_T$) *The properties zorro, convergence, nada, linearity, adaptedness, local martingale, zipp and stopping in Theorems 15.14 and 15.15 all carry over without any changes from Itô integrals with respect to continuous square-integrable integrals to Itô integrals with respect to continuous local martingales.*

Proof. The properties zorro, convergence, nada, linearity, adaptedness, zipp and stopping are all more or less immediate from Exercise 167 together with Theorems 15.14 and 15.15, see also Exercises 168 and 169 below. Moreover, the property local martingale is immediate from the stopping property together with inspection of Definition and Theorem 15.16. □

Exercise 168. Prove the convergence property in Theorem 15.17.

Exercise 169. Prove the stopping property in Theorem 15.17.

16 Continuous local martingales and martingale problems

16.1 Approximation of Itô integrals

Exercise 170. Show that for a continuous square-integrable martingale $\{M(t)\}_{t\in[0,T]}$ and an $X \in P(M)_T$ we have

$$\int_0^T (X_n - X)^2\, d[M] \to 0 \quad \text{as } n \to \infty \quad \text{for some sequence } \{X_n\}_{n=1}^{\infty} \subseteq S_T$$

in the sense of convergence in probability. Conclusions?

Exercise 171. Show that for a continuous local martingale $\{M(t)\}_{t\in[0,T]}$ and an $X \in P(M)_T$ we have

$$\int_0^T (X_n - X)^2\, d[M] \to 0 \quad \text{as } n \to \infty \quad \text{for some sequence } \{X_n\}_{n=1}^{\infty} \subseteq S_T$$

in the sense of convergence in probability. Conclusions?

Unsurprisingly, we need the following version of Theorem 6.6:

Theorem 16.1. *For a continuous local martingale $\{M(t)\}_{t\in[0,T]}$, an $X \in P(M)_T$ and a constant $C > 0$, we have*

$$\mathbf{P}\bigg\{\sup_{t\in[0,T]}\bigg|\int_0^t X\, dM\bigg| > \lambda\bigg\} \le \frac{C}{\lambda^2} + \mathbf{P}\bigg\{\int_0^T X^2\, d[M] \ge C\bigg\} \quad \text{for } \lambda > 0.$$

Proof. Let $\{\tau_m\}_{m=1}^{\infty}$ be a localizing sequence of stopping times such that $\{M(t\wedge\tau_m)\}_{t\in[0,T]}$ is a square-integrable continuous martingale for $m \in \mathbb{N}$. Define

$$\tau_C = T \wedge \inf\bigg\{t \in [0,T] : \int_0^t X^2\, d[M] \ge C\bigg\}$$

[cf. (6.3)], and note that $X_C \equiv X I_{[0,\tau_C]} \in E(M(\,\cdot\wedge\tau_m))_T$ with

$$\int_0^T X_C^2\, d[M](\,\cdot\wedge\tau_m) = \int_0^{T\wedge\tau_m} X_C^2\, d[M] \le \int_0^T X_C^2\, d[M] \le C.$$

Hence Exercise 167 together with Theorem 15.12 and the stopping property give

$$\begin{aligned}
&\mathbf{P}\bigg\{\sup_{t\in[0,T]}\bigg|\int_0^t X\, dM\bigg| > \lambda\bigg\}\\
&\qquad \le \limsup_{m\to\infty} \mathbf{P}\bigg\{\sup_{t\in[0,T]}\bigg|\int_0^t X\, dM - \int_0^t X\, dM(\,\cdot\wedge\tau_m)\bigg| > \varepsilon\bigg\}\\
&\qquad\qquad + \limsup_{m\to\infty} \mathbf{P}\bigg\{\sup_{t\in[0,T]}\bigg|\int_0^t X\, dM(\,\cdot\wedge\tau_m)\bigg| > \lambda - \varepsilon\bigg\}\\
&\qquad \le 0 + \limsup_{m\to\infty} \mathbf{P}\bigg\{\sup_{t\in[0,T]}\bigg|\int_0^t X_C\, dM(\,\cdot\wedge\tau_m)\bigg| > \lambda - \varepsilon\bigg\} + \mathbf{P}\{\tau_C < T\}
\end{aligned}$$

$$\le 0 + \frac{C}{(\lambda-\varepsilon)^2} + 0 + \mathbf{P}\left\{\int_0^T X^2\, d[M] \ge C\right\} \quad \text{for each } \varepsilon > 0. \qquad \square$$

Exercise 172. Show that for a continuous local martingale $\{M(t)\}_{t\in[0,T]}$ and a continuous and adapted process $\{X(t)\}_{t\in[0,T]}$ it holds that

$$\mathbf{P}\left\{\sup_{t\in[0,T]}\left|\int_0^t X\, dM\right| > \lambda\right\} \le \frac{C}{\lambda^2} + \mathbf{P}\left\{\int_0^T X^2\, d[M] > C\right\} \quad \text{for } C,\lambda > 0.$$

The following version of Theorem 6.7 is crucial to derive Itô formulas. It follows from Theorem 16.1 in the same fashion as Theorem 6.7 follows from Theorem 16.2.

Theorem 16.2. *A continuous adapted process $\{X(t)\}_{t\in[0,T]}$ belongs to $P(M)_T$ and*

$$\sup_{t\in[0,T]}\left|\int_0^t X\, dM - \int_0^t \sum_{i=1}^n X(t_{i-1}) I_{(t_{i-1},t_i]}\, dM\right| \to 0 \quad \textit{in probability}$$

for partitions $0 = t_0 < t_1 < \ldots < t_n = T$ of $[0,T]$ such that $\max_{1\le i\le n} t_i - t_{i-1} \downarrow 0$.

Exercise 173. Derive Theorem 16.2 from Theorem 16.1.

16.2 Quadratic variation of Itô integrals

Recall that we had to use two-and-a-half or so pages to derive the formula for the qudratic variation of and Itô integral with respect to BM in Theorem 8.6. Now we will see that things go much easier when we have the Doob-Meyer representation.

Theorem 16.3. *For a continuous local martingale $\{M(t)\}_{t\in[0,T]}$ and an $X \in P(M)_T$, we have*

$$\left[\int_0^{(\cdot)} X\, dM\right](t) = \int_0^t X^2\, d[M] \quad \textit{for } t \in [0,T].$$

Proof. For $X \in S_T$ the claim of the theorem follows in exactly the same fashion as it does in the proof of Theorem 8.6 – just make the appropriate replacements of B and dr with M and $d[M](r)$, respectively.

Now assume that M is square-integrable and that $X \in E(M)_T$. Pick a sequence $\{X_n\}_{n=1}^\infty \in S_T$ such that (15.2) holds. By the Doob-Meyer representation, in order to prove the claim of the theorem, it is sufficient to prove that

$$\mathbf{E}\left\{\left(\int_0^t X\, dM\right)^2 - \left(\int_0^s X\, dM\right)^2 - \left(\int_0^t X^2\, d[M] - \int_0^s X^2\, d[M]\right) \,\middle|\, \mathcal{F}_s\right\} = 0$$

for $0 \le s \le t \le T$. However, by (15.2) together with Definition and Theorem 15.5, the above conditional expectation equals the limit (in the sense of convergence in $\mathbb{L}^1$, recall Exercise 53) as $n\to\infty$ of

$$\mathbf{E}\bigg\{\bigg(\int_0^t X_n\,dM\bigg)^2-\bigg(\int_0^s X_n\,dM\bigg)^2-\bigg(\int_0^t X_n^2\,d[M]-\int_0^s X_n^2\,d[M]\bigg)\,\bigg|\,\mathcal{F}_s\bigg\}.$$

Now, this conditional expectation in turn is always zero by the first part of the proof for the case $X\in S_T$ together with the Doob-Meyer representation.

Return to the general setting when $\{M(t)\}_{t\in[0,T]}$ is a continuous local martingale and $X\in P(M)_T$. Pick a localizing sequence $\{\tau_m\}_{m=1}^\infty$ such that $\{M(t\wedge\tau_m)\}_{t\in[0,T]}$ is square-integrable and $\int_0^T I_{[0,\tau_m]}X^2\,d[M]\leq m$ for $m\in\mathbb{N}$, see Exercise 174 below. Using what we have proved already for M square-integrable and $X\in E(M)_T$, we then get using Exercise 163 and the stopping and zorro properties (see Theorem 15.17)

$$\begin{aligned}\bigg[\int_0^{(\cdot)} X\,dM\bigg](t\wedge\tau_m) &= \bigg[\int_0^{(\cdot\wedge\tau_m)} X\,dM\bigg](t)\\ &= \bigg[\int_0^{(\cdot\wedge\tau_m\wedge\tau_m)} X\,dM\bigg](t)\\ &= \bigg[\int_0^{(\cdot)} I_{[0,\tau_m]}X\,dM(\,\cdot\wedge\tau_m)\bigg](t)\\ &= \int_0^t I_{[0,\tau_m]}X^2\,d[M(\,\cdot\wedge\tau_m)]\\ &= \int_0^{t\wedge\tau_m} X^2\,d[M]\quad\text{for } t\in[0,T],\end{aligned}\tag{16.1}$$

see Exercise 175 below. Hence the theorem follows from sending $m\to\infty$. □

Exercise 174. Explain the details of how to find a localizing sequence of stopping times $\{\tau_m\}_{m=1}^\infty$ that has the two properties claimed in the proof of Theorem 16.3.

Exercise 175. Explain all steps in (16.1) in full detail.

16.3 Burkholder-Davis-Gundy inequality

The inequality in Exercise 157 is a special case of the following famous result, the proof of which is very complicated (except in the case $p=1$ treated in Exercise 157):

Theorem 16.4. (Burkholder-Davis-Gundy inequality[11]) *There exist functions $c,C:(0,\infty)\to(0,\infty)$ such that for every continuous local martingale $\{M(t)\}_{t\in[0,T]}$, for every stopping time $\tau\leq T$, and for every constant $p\in(0,\infty)$, we have*

$$c(p)\,\mathbf{E}\big\{[M](\tau)^{p/2}\big\}\leq\mathbf{E}\Big\{\sup_{t\in[0,\tau]}|M(t)|^p\Big\}\leq C(p)\,\mathbf{E}\big\{[M](\tau)^{p/2}\big\}.$$

[11] See e.g., Revuz and Yor: "Continuous Martingales and Brownian Motion", Section IV.4. As we do not make any essential use of this result, it is no real loss that we do not prove it.

The power of the Burkholder-Davis-Gundy inequality is illustrated by the next two corollaries that give a sharp constancy criteria for local martingales and an improved sufficient condition for Itô integrals to be martingales, respectively.

Corollary 16.5. *A continuous local martingale* $\{M(t)\}_{t\in[0,T]}$ *is constant over an interval* $[s,t]\subseteq[0,T]$ *if and only if* $[M](s)=[M](t)$.

Exercise 176. Derive Corollary 16.5 from Theorem 16.4.

It is the Burkholder-Davis-Gundy inequality for $p=2$ together with Corollary 16.5 that are important to build up the theory. Luckily, we can derive these results directly from Exercise 157, see Exercise 177 below. Thus the omission of a proof of the Burkholder-Davis-Gundy inequality does not create any holes in our theory building.

Exercise 177. Derive the Burkholder-Davis-Gundy inequality for $p=2$ as well as Corollary 16.5 from Exercise 157 without using Theorem 16.4.

Corollary 16.6. *An Itô integral process* $\{\int_0^t X\,dM\}_{t\in[0,T]}$ *for an* $X\in P(M)_T$ *with respect to a continuous local martingale* $\{M(t)\}_{t\in[0,T]}$ *is a martingale if*

$$\mathbf{E}\bigg\{\bigg(\int_0^T X^2\,d[M]\bigg)^{1/2}\bigg\}<\infty.$$

In particular, the Itô integral process $\{\int_0^t X\,dB\}_{t\in[0,T]}$ *for* $X\in P_T$ *is a martingale if*

$$\mathbf{E}\bigg\{\bigg(\int_0^T X(r)^2\,dr\bigg)^{1/2}\bigg\}<\infty.$$

Exercise 178. Derive Corollary 16.6 from Theorems 13.12 and 16.3-16.4.

Exercise 179. Explain what is the improvement in Corollary 16.6 of the sufficient condition for Itô integrals to be martingales as compared with our previous knowledge of results of this type.

Putting together Corollaries 15.9 and 16.5 we get the following appealing result:

Corollary 16.7. *A pair of continuous local martingales* $\{M(t)\}_{t\in[0,T]}$ *and* $\{N(t)\}_{t\in[0,T]}$ *has a well-defined continuous quadratic covariation process* $\{[M,N](t)\}_{t\in[0,T]}$ *which is the unique adapted finite variation process with* $[M,N](0)=0$ *that makes* $\{M(t)N(t)-[M,N](t)\}_{t\in[0,T]}$ *a continuous local martingale.*

Exercise 180. Prove Corollary 16.7.

Putting together Corollaries 13.19 and 16.5 we further obtain the following important (it shall turn out) sharpening of Corollary 13.19:

Corollary 16.8. *A pair of continuous square-integrable martingales $\{M(t)\}_{t\in[0,T]}$ and $\{N(t)\}_{t\in[0,T]}$ has a well-defined continuous, adapted and integrable quadratic co-variation process $\{[M,N](t)\}_{t\in[0,T]}$ (in the sense of convergence in probability) which is uniquely determined by the property that $\{M(t)N(t)-[M,N](t)\}_{t\in[0,T]}$ is a continuous martingale.*

Exercise 181. Prove Corollary (16.8).

Corollary 16.9. *For two continuous local martingales $\{M(t)\}_{t\in[0,T]}$ and $\{N(t)\}_{t\in[0,T]}$ an two processes $X\in P(M)_T$ and $Y\in P(N)_T$, we have*

$$\left[\int_0^{(\cdot)} X\,dM, \int_0^{(\cdot)} Y\,dN\right](t) = \int_0^t XY\,d[M,N] \quad \text{for } t\in[0,T]. \tag{16.2}$$

Proof. For $X,Y\in S_T$ (16.2) follows in exactly the same fashion as in the proofs of Theorems 8.6 and 16.3. Next assume that M and N are square-integrable and that $X\in E(M)_T$ and $Y\in E(N)_T$. Pick sequences $\{X_n\}_{n=1}^\infty, \{Y_n\}_{n=1}^\infty \in S_T$ such that

$$\lim_{n\to\infty} \mathbf{E}\left\{\int_0^T (X_n-X)^2\,d[M]\right\} = \lim_{n\to\infty} \mathbf{E}\left\{\int_0^T (Y_n-Y)^2\,d[N]\right\} = 0. \tag{16.3}$$

By Corollary 16.8, in order to prove (16.2) it is sufficient to prove that

$$\mathbf{E}\left\{\left(\int_0^t X\,dM\right)\left(\int_0^t Y\,dN\right) - \left(\int_0^s X\,dM\right)\left(\int_0^s Y\,dN\right) - \int_s^t XY\,d[M,N] \,\middle|\, \mathcal{F}_s\right\} = 0$$

for $0\le s\le t\le T$. However, by (16.3) and Definition and Theorem 15.5 together with Exercise 182 below, the above conditional expectation equals the limit (in the sense of convergence in $\mathbb{L}^1$, recall Exercise 53) as $n\to\infty$ of

$$\mathbf{E}\left\{\left(\int_0^t X_n\,dM\right)\left(\int_0^t Y_n\,dN\right) - \left(\int_0^s X_n\,dM\right)\left(\int_0^s Y_n\,dN\right) - \int_s^t X_nY_n\,d[M,N] \,\middle|\, \mathcal{F}_s\right\}.$$

Now, this conditional expectation in turn is always zero by the first part of the proof for the case $X,Y\in S_T$ together with Corollary 16.8.

Return to the general setting when $\{M(t)\}_{t\in[0,T]}$ and $\{N(t)\}_{t\in[0,T]}$ are continuous local martingales and $X\in P(M)_T$ and $Y\in P(N)_T$. Pick a localizing sequence $\{\tau_m\}_{m=1}^\infty$ such that $\{M(t\wedge\tau_m)\}_{t\in[0,T]}$ and $\{N(t\wedge\tau_m)\}_{t\in[0,T]}$ are square-integrable and such that $\int_0^T I_{[0,\tau_m]}X^2\,d[M] \le m$ and $\int_0^T I_{[0,\tau_m]}Y^2\,d[N] \le m$ for $m\in\mathbb{N}$, recall Exercise 174. Using what we have proved already for M and N square-integrable with $X\in E(M)_T$ and $Y\in E(N)_T$, we then get from polarization and Exercise 163 together with the stopping and zorro properties for the Itô integral

$$\begin{aligned}
\Big[\int_0^{(\cdot)} X\,dM, \int_0^{(\cdot)} Y\,dN\Big](t\wedge\tau_m)
&= \Big[\int_0^{(\cdot\wedge\tau_m)} X\,dM, \int_0^{(\cdot\wedge\tau_m)} Y\,dN\Big](t)\\
&= \Big[\int_0^{(\cdot\wedge\tau_m\wedge\tau_m)} X\,dM, \int_0^{(\cdot\wedge\tau_m\wedge\tau_m)} Y\,dN\Big](t)\\
&= \Big[\int_0^{(\cdot)} I_{[0,\tau_m]}X\,dM(\,\cdot\wedge\tau_m), \int_0^{(\cdot)} I_{[0,\tau_m]}Y\,dN(\,\cdot\wedge\tau_m)\Big](t)\\
&= \int_0^t I_{[0,\tau_m]}XY\,d[M(\,\cdot\wedge\tau_m), N(\,\cdot\wedge\tau_m)]\\
&= \int_0^{t\wedge\tau_m} XY\,d[M,N] \quad \text{for } t\in[0,T],
\end{aligned}$$

see also Exercise 175. Hence the theorem follows from sending $m\to\infty$. □

Exercise 182. Prove that (16.3) implies that

$$\int_0^t X_nY_n\,d[M,N] \to \int_0^t XY\,d[M,N] \quad \text{in } \mathbb{L}^1 \text{ as } n\to\infty \text{ for } t\in[0,T].$$

16.4 Stochastic differentials and Itô formula

Equipped with Itô integrals for continuous local martingales we may consider stochastic differentials and Itô processes driven by such:

Definition 16.10. *Let $\{M(t)\}_{t\in[0,T]}$ be a continuous local martingale. If $\{\mu(t)\}_{t\in[0,T]}$ is an adapted measurable stochastic process such that*

$$\mathbf{P}\Big\{\int_0^T |\mu(r)|\,dr < \infty\Big\} = 1,$$

if $\sigma\in P(M)_T$, and if $X(0)$ is an $\mathcal{F}_0$-measurable random variable, then we call

$$X(t) = X(0) + \int_0^t \mu(r)\,dr + \int_0^t \sigma\,dM, \quad t\in[0,T],$$

an Itô process *and the corresponding* stochastic differential *is given by*

$$dX(t) = \mu(t)\,dt + \sigma(t)\,dM(t).$$

Note the difference between the concepts of Itô integral process and Itô process!

Corollary 16.11. *The quadratic variation of an Itô process* $\{X(t)\}_{t\in[0,T]}$ *with stochastic differential* $dX(t) = \mu(t)\,dt + \sigma(t)\,dM(t)$ *is given (with probability 1) by*

$$[X]([s,t]) = \int_s^t \sigma^2\,d[M] \quad \textit{for } [s,t] \subseteq [0,T].$$

Exercise 183. Prove Corollary 16.11.

Exercise 184. Show that an Itô process has zero quadratic variation if and only if it has finite variation.

Corollary 16.12. *The quadratic covariation between two Itô processes* $\{X(t)\}_{t\in[0,T]}$ *and* $\{Y(t)\}_{t\in[0,T]}$ *with stochastic differentials* $dX(t) = \mu_X(t)\,dt + \sigma_X(t)\,dM(t)$ *and* $dY(t) = \mu_Y(t)\,dt + \sigma_Y(t)\,dM(t)$*, respectively, is given (with probability 1) by*

$$[X,Y]([s,t]) = \int_s^t \sigma_X\sigma_Y\,d[M] \quad \textit{for } [s,t] \subseteq [0,T].$$

Exercise 185. Prove Corollary 16.12.

Corollary 16.13. *If* $\mathbf{P}\{X(t)=Y(t) \ \textit{for } t\in[0,T]\} = 1$ *for two Itô processes* $\{X(t)\}_{t\in[0,T]}$ *and* $\{Y(t)\}_{t\in[0,T]}$ *with stochastic differentials* $dX(t) = \mu_X(t)\,dt + \sigma_X(t)\,dM(t)$ *and* $dY(t) = \mu_Y(t)\,dt + \sigma_Y(t)\,dM(t)$*, respectively, then we have*

$$\mathbf{P}\{\sigma_X(t)=\sigma_Y(t) \ \textit{a.e. } d[M] \ \textit{for } t\in[0,T]\} = \mathbf{P}\{\mu_X(t)=\mu_Y(t) \ \textit{a.e. for } t\in[0,T]\} = 1.$$

Exercise 186. Prove Corollary 16.13.

Definition 16.14. *If* $dX(t) = \mu(t)\,dt + \sigma(t)\,dM(t)$ *is a stochastic differential and* $\{Y(t)\}_{t\in[0,T]}$ *an adapted measurable process such that*

$$\mathbf{P}\bigg\{\int_0^T |Y(r)|\,|\mu(r)|\,dr < \infty\bigg\} = 1$$

and $Y\sigma \in P(M)_T$*, then we define the Itô process* $\{\int_0^t X\,dY\}_{t\in[0,T]}$ *by*

$$\int_0^t Y\,dX = \int_0^t Y(r)\mu(r)\,dr + \int_0^t Y\sigma\,dM \quad \textit{for } t\in[0,T].$$

In order to make sure that the above definition of $\int_0^t Y\,dX$ is consistent, in the sense of not being multi-valued, we must check that if two Itô processes agree

$$\mathbf{P}\Big\{\int_0^t \mu_1(r)\,dr + \int_0^t \sigma_1\,dM = \int_0^t \mu_2(r)\,dr + \int_0^t \sigma_2\,dM \text{ for } t\in[0,T]\Big\} = 1, \quad (16.4)$$

then

$$\mathbf{P}\Big\{\int_0^t Y(r)\mu_1(r)\,dr + \int_0^t Y\sigma_1\,dM = \int_0^t Y(r)\mu_2(r)\,dr + \int_0^t Y\sigma_2\,dM \text{ for } t\in[0,T]\Big\} = 1. \quad (16.5)$$

However, this follows from Corollary 16.13, see Exercise 187 below.

Exercise 187. Show that (16.4) implies (16.5).

Exercise 188. Explain why the Itô process $\int_0^t Y\,dX$ is well-defined when X is an Itô process and Y is a continuous adapted process. Conclude that the Itô process $\int_0^t Y\,dX$ is well-defined when X and Y are both Itô processes.

Unsurprisingly, we need the following generalization of Theorem 16.2:

Theorem 16.15. *For an Itô process $\{X(t)\}_{t\in[0,T]}$ and a continuous adapted process $\{Y(t)\}_{t\in[0,T]}$, we have*

$$\sup_{t\in[0,T]}\bigg|\int_0^t Y\,dX - \int_0^t \sum_{i=1}^n Y(t_{i-1}) I_{(t_{i-1},t_i]}\,dX\bigg| \to 0 \quad \textit{in probability}$$

for partitions $0=t_0<t_1<\ldots<t_n=T$ *of* $[0,T]$ *such that* $\max_{1\leq i\leq n} t_i - t_{i-1}\downarrow 0$.

Exercise 189. Prove Theorem 16.15.

The proof of the Itô formula in Theorem 11.1 carries over with only obvious modifications, see Exercise 190 below, to the continuous local martingale setting:

Theorem 16.16. (Itô formula) *For an Itô process $\{X(t)\}_{t\in[0,T]}$ all values of which belong to an open interval $I\subseteq\mathbb{R}$ with probability 1 and a function $f\in C^{1,2}([0,T]\times I)$, we have*

$$df(t,X(t)) = f_t'(t,X(t))\,dt + f_x'(t,X(t))\,dX(t) + \frac{1}{2}f_{xx}''(t,X(t))\,d[X](t) \quad \textit{for } t\in[0,T].$$

Exercise 190. Explain how we can be so sure that the above Itô formula Theorem 16.16 is true without any worries.

16.5 Paul Lévy's characterization of BM

We are now prepared to extend Paul Lévy's characterization of BM Theorem 10.1 to its full continuous local martingale generality:

Theorem 16.17. (PAUL LÉVY'S CHARACTERIZATION OF BM) *A continuous local martingale $\{M(t)\}_{t\in[0,T]}$ is BM if and only if it has quadratic variation process $[M](t)=t$ for $t\in[0,T]$.*

Proof. $\boxed{\Leftarrow}$ Assume that $[M](t)=t$ for $t\in[0,T]$ and consider the processes

$$Y_1(t)=\int_0^t Z_1(r)\,dM(r)\quad\text{and}\quad Y_2(t)=\int_0^t Z_2(r)\,dM(r)\quad\text{for } t\in[0,T],$$

where $Z_1(t)=\cos(\theta M(t))\,\mathrm{e}^{\frac{1}{2}\theta^2 t}$ and $Z_2(t)=\sin(\theta M(t))\,\mathrm{e}^{\frac{1}{2}\theta^2 t}$ for $t\in[0,T]$, and where $\theta\in\mathbb{R}$ is a constant. As the processes $\{Z_1(t)\}_{t\in[0,T]}$ and $\{Z_2(t)\}_{t\in[0,T]}$ are bounded by the constant $\mathrm{e}^{\frac{1}{2}\theta^2 T}$ they belong to $E(M)_T$, so that Y_1 and Y_2 are martingales, see Exercise 191 below. Further, Itô's formula Theorem 16.16 shows that

$$\begin{aligned} dZ_1(t)&=\frac{1}{2}\theta^2 Z_1(t)\,dt-\theta\,Z_2(t)\,dM(t)-\frac{1}{2}\theta^2 Z_1(t)\,d[M](t)=-\theta\,dY_2(t),\\ dZ_2(t)&=\frac{1}{2}\theta^2 Z_2(t)\,dt+\theta\,Z_1(t)\,dM(t)-\frac{1}{2}\theta^2 Z_2(t)\,d[M](t)=\quad\theta\,dY_1(t). \end{aligned}$$

Hence Z_1 and Z_2 are martingales (as Y_1 and Y_2 are martingales). It follows that

$$\mathbf{E}\{\mathrm{e}^{i\theta M(t)+\frac{1}{2}\theta^2 t}\big|\mathcal{F}_s\}=\mathbf{E}\{Z_1(t)\big|\mathcal{F}_s\}+i\,\mathbf{E}\{Z_2(t)\big|\mathcal{F}_s\}=Z_1(s)+i\,Z_2(s)=\mathrm{e}^{i\theta M(s)+\frac{1}{2}\theta^2 s}$$

for $0\le s\le t\le T$, which in turn (as M is adapted) by rearrangement gives

$$\mathbf{E}\{\mathrm{e}^{i\theta(M(t)-M(s))}\big|\mathcal{F}_s\}=\mathrm{e}^{-\frac{1}{2}\theta^2(t-s)}\quad\text{and}\quad\mathbf{E}\{\mathrm{e}^{i\theta(M(t)-M(s))}\}=\mathrm{e}^{-\frac{1}{2}\theta^2(t-s)}$$

for $0\le s\le t\le T$. From this we may finish off the proof of the theorem in exactly the same manner as we used (10.1) to finish off the proof of Theorem 10.1. □

Exercise 191. Explain why $Z_1,Z_2\in E(M)_T$ in the proof of Theorem 16.17.

16.6 Local martingale problems[12]

The following approach to weak solutions to SDE was invented by Daniel Strock and S.R.S. Varadhan[13]. The latter is famous for his exceptional ability to rephrase any (that is, any!) problem presented to him as a problem regarding martingales. Unsurprisingly then, he did so too to find weak solutions to SDE.

[12] Here the moment of truth comes at last folks!

[13] See their very demanding book "Multidimensional Diffusion Processes".

Definition 16.18. *The* generator *of the SDE* (9.1) *is the differential operator* $\mathcal{A}_t$ *given by*

$$(\mathcal{A}_t f)(x) = \mu(t,x)\, f'(x) + \frac{\sigma(t,x)^2}{2}\, f''(x) \quad \textit{for } f \in C^2(\mathbb{R}). \tag{16.6}$$

Definition 16.19. *A continuous and adapted stochastic process* $\{X(t)\}_{t\in[0,T]}$ *is a solution to the* local martingale problem *associated with the generator* $\mathcal{A}_t$ *in* (16.6) *if for each* $f \in C^2(\mathbb{R})$ *the following stochastic process is a continuous local martingale*

$$\left\{ f(X(t)) - f(X(0)) - \int_0^t (\mathcal{A}_r f)(X(r))\, dr \right\}_{t\in[0,T]}. \tag{16.7}$$

Theorem 16.20. *A continuous and adapted stochastic process* $\{X(t)\}_{t\in[0,T]}$ *is a weak solution to the SDE* (9.1) *if and only if* $X(0) =_D X_0$ *and* X *is a solution to the local martingale problem associated with the generator* $\mathcal{A}_t$ *in* (16.6).

Proof. $\boxed{\Leftarrow}$ Let $\{X(t)\}_{t\in[0,T]}$ solve the local martingale problem associated with the differential operator $\mathcal{A}_t$ in (16.6). Taking $f(x)=x$ in (16.7) we see that

$$M(t) = X(t) - X(0) - \int_0^t \mu(r, X(r))\, dr \quad \text{for } t \in [0,T] \tag{16.8}$$

is a continuous local martingale. Hence the Itô formula Theorem 16.16 gives

$$\begin{aligned} & df(X(t)) - (\mathcal{A}_t f)(X(t))\, dt \\ &= f'(X(t))\, dX(t) + \frac{1}{2}\, f''(X(t))\, d[X](t) - (\mathcal{A}_t f)(X(t))\, dt \\ &= f'(X(t))\, dM(t) + \frac{1}{2}\, f''(X(t))\, d[M](t) - \frac{\sigma(t,X(t))^2}{2}\, f''(X(t))\, dt \quad \text{for } t \in [0,T], \end{aligned} \tag{16.9}$$

for $f \in C^2(\mathbb{R})$. As X solves the local martingale problem and the process in (16.8) is a continuous local martingale, we may conclude from (16.9) in turn that

$$\left\{ \int_0^t \frac{1}{2}\, f''(X)\, d[M] - \int_0^t \frac{\sigma(r,X(r))^2}{2}\, f''(X(r))\, dr \right\}_{t\in[0,T]}$$

is also a continuous local martingale. However, as this process is also a finite variation process it has zero quadratic variation – an immediate corollary to Theorem 8.3 – so that this continuous local martingale is constant, by Corollary 16.5, and therefore zero, as it takes of at zero at time zero. Taking $f(x)=x^2$ this shows that

$$\{[M](t)\}_{t\in[0,T]} = \left\{ \int_0^t \sigma(r, X(r))^2\, dr \right\}_{t\in[0,T]}.$$

Therefore the following process is a well-defined Itô integral process

$$\{W(t)\}_{t\in[0,T]} = \left\{\int_0^t I_{\{\sigma(r,X(r))\neq 0\}}\,\frac{1}{\sigma(r,X(r))}\,dM(r) + \int_0^t I_{\{\sigma(r,X(r))=0\}}\,dB(r)\right\}_{t\in[0,T]}.$$

As the quadratic variation of this continuous local martingale is $[W](t)=t$, it follows from Paul Lévy's characterization of BM Theorem 16.17 that W is BM.

Putting all our findings together wee see that

$$dM(t) = dX(t) - \mu(t,X(t))\,dt = \sigma(t,X(t))\,dW(t) \quad \text{for } t\in[0,T]. \qquad \square$$

Exercise 192. Show how the implication to the right in Theorem 16.20 follows from the Itô formula for Itô processes in Theorem 7.6.

17 Existence of weak solutions to SDE

17.1 Martingale problems

In order to find a solution to a local martingale problem it turns out to be convenient to rephrase the problem as a so called *martingale problem*:

Corollary 17.1. *If* $\{X(t)\}_{t\in[0,T]}$ *is a solution to the SDE* (9.1) *with the generator* $\mathcal{A}_t$ *in* (16.6) *where the coefficient* σ *is locally bounded*[14], *then for each* $f \in C_0^2(\mathbb{R})$[15] *the following stochastic process is a continuous martingale*

$$\left\{ f(X(t)) - f(X(0)) - \int_0^t (\mathcal{A}_r f)(X(r))\, dr \right\}_{t\in[0,T]}. \tag{17.1}$$

Proof. Since the process (17.1) is a continuous local martingale by Theorem 16.20, that is bounded by the assumptions on f and σ, see Exercise 193 below, it is a continuous martingale by Theorem 13.12. □

Exercise 193. Prove that the process (17.1) is bounded under the hypothesis of Corollary 17.1.

Exercise 194. Given an alternative proof of Corollary 17.1 based on Itô's formula together with the Burkholder-Davis-Gundy inequality for $p = 2$ (recall Exercise 177) and Theorem 13.12.

Definition 17.2. *A continuous and adapted stochastic process* $\{X(t)\}_{t\in[0,T]}$ *is a solution to the* martingale problem *associated with the generator* $\mathcal{A}_t$ *in* (16.6) *if for each* $f \in C_0^2(\mathbb{R})$ *the process* (17.1) *is a continuous martingale.*

Theorem 17.3. *A continuous and adapted stochastic process* $\{X(t)\}_{t\in[0,T]}$ *is a weak solution to the SDE* (9.1) *with a locally bounded* σ *coefficient if and only if* $X(0) =_D X_0$ *and* X *is a solution to the martingale problem associated with the generator* $\mathcal{A}_t$ *in* (16.6).

Proof. The implication to the right follows from Corollary 17.1. For the implication to the left, let X be a continuous and adapted process that solves the martingale problem associated with $\mathcal{A}_t$. By Theorem 16.20 it is sufficient to show that X solves the local martingale problem associated with $\mathcal{A}_t$. To that end, given an $f \in C^2(\mathbb{R})$

[14] The coefficient function $\sigma : [0,T]\times\mathbb{R}$ is bounded on compact subsets of $[0,T]\times\mathbb{R}$.

[15] The class of two times continuously differentiable functions $f : \mathbb{R} \to \mathbb{R}$ with compact support.

and an $k\in\mathbb{N}$, pick an $f_k\in C_0^2(\mathbb{R})$ that agree with f on the interval $[-k,k]$. Then

$$\{M_k(t)\}_{t\in[0,T]} = \bigg\{ f_k(X(t)) - f_k(X(0)) - \int_0^t (\mathcal{A}_r f_k)(X(r))\,dr \bigg\}_{t\in[0,T]}$$

is a continuous martingale. Define a stopping time $\tau_k=\inf\{t\geq 0: |X(t)|\geq k\}$. By the continuity of X we have $\tau_k\uparrow\infty$ as $k\to\infty$. Hence it is sufficient to show that $\{M(t\wedge\tau_k)\}_{t\in[0,T]}$ is a martingale for each $k\in\mathbb{N}$, where M is the process given by (17.1) (which clearly is continuous). However, by the optional stopping theorem

$$\begin{aligned}
&\{M(t\wedge\tau_k)\}_{t\in[0,T]}\\
&= \bigg\{ f(X(t\wedge\tau_k)) - f(X(0)) - \int_0^{t\wedge\tau_k} (\mathcal{A}_r f)(X(r))\,dr \bigg\}_{t\in[0,T]}\\
&= \bigg\{ f_k(X(t\wedge\tau_k)) + I_{\{\tau_k=0\}}[f(X(0)) - f_k(X(0))] - f(X(0)) - \int_0^{t\wedge\tau_k} (\mathcal{A}_r f_k)(X(r))\,dr \bigg\}_{t\in[0,T]}\\
&= \big\{ M_k(t\wedge\tau_k) - I_{\{\tau_k>0\}}[f(X(0)) - f_k(X(0))]\big\}_{t\in[0,T]}\\
&= \{M_k(t\wedge\tau_k)\}_{t\in[0,T]}
\end{aligned}$$

is a martingale when M_k is a continuous martingale. □

17.2 Existence of solutions to martingale problems

For strong solutions to SDE, uniqueness criteria (e.g., Theorems 10.5 and 11.3) typically are less demanding in terms of their hypothesises than existence criteria (e.g., Theorem 12.3). For weak solutions it turns out to be the other way around, so that weak uniqueness is more demanding than weak existence.

Our weak existence proof uses weak convergence of probability measures for an Euler iteration scheme (recall Exercise 113). For that purpose we cite the following standard result from a standard graduate course in weak convergence[16]:

Lemma 17.4. *Let $\{_1X(t)\}_{t\in[0,T]}$, $\{_2X(t)\}_{t\in[0,T]}$, ... be continuous stochastic processes such that*

$$\lim_{\lambda\to\infty}\sup_{k\geq 1}\mathbf{P}\{|_kX(0)|>\lambda\}=0, \tag{17.2}$$

and such that there exist constants $C,\alpha,\beta>0$ such that

$$\mathbf{E}\big\{|_kX(t)-{_kX(s)}|^\alpha\big\} < C\,|t-s|^{1+\beta} \quad \text{for } s,t\in[0,T] \text{ and } k\in\mathbb{N}. \tag{17.3}$$

There exist a continuous stochastic process $\{X(t)\}_{t\in[0,T]}$ and a sequence $\{k_j\}_{j=1}^\infty\subseteq\mathbb{N}$ with $k_j\uparrow\infty$ as $j\to\infty$ such that for each bounded continuous function $F:C([0,T])$[17] $\to\mathbb{R}$ we have $F(_{k_j}X)\to F(X)$ as $j\to\infty$ in the sense of convergence in distribution.

[16]See e.g., Karatzas and Shreve: "Brownian Motion and Stochastic Calculus", Section 2.4.B.

[17]The space of continuous functions $f:[0,T]\to\mathbb{R}$ equipped with the norm $\|f\|=\sup_{t\in[0,T]}|f(t)|$.

As this is a course on SDE, the two main results of the course so far are Theorem 10.5 (see also Theorem 11.3) and Theorem 12.3 concerning uniqueness and existence of strong solutions to SDE, respectively. We are now prepared to state and prove the third main result of the course, which concerns existence of weak solutions to SDE. We prove the existence by solving the corresponding martingale problem.

Theorem 17.5. (STROOCK-VARADHAN) *Consider the generator $\mathcal{A}_t$ in* (16.6) *where the coefficients $\mu, \sigma : [0,T]\times\mathbb{R} \to \mathbb{R}$ are bounded and continuous. For each random variable X_0 the martingale problem associated with $\mathcal{A}_t$ has a solution $\{X(t)\}_{t\in[0,T]}$ such that $X(0) =_D X_0$.*

Proof. Given a $k \in \mathbb{N}$, define a process $\{{}_kX(t)\}_{t\in[0,T]}$ recursively by ${}_kX(0) = X_0$ and

$${}_kX(t) = {}_kX(\tfrac{\ell}{k}) + \mu(\tfrac{\ell}{k}, {}_kX(\tfrac{\ell}{k}))\,(t-\tfrac{\ell}{k}) + \sigma(\tfrac{\ell}{k}, {}_kX(\tfrac{\ell}{k}))\,(B(t)-B(\tfrac{\ell}{k})) \quad \text{for } t \in (\tfrac{\ell}{k}, \tfrac{\ell+1}{k}\wedge T],$$

for $\ell = 1, 2, \ldots, \lfloor kT \rfloor$. Notice that, writing $\mu_k(0) = \sigma_k(0) = 0$ and

$$\mu_k(t) = \mu(\tfrac{\ell}{k}, {}_kX(\tfrac{\ell}{k})) \quad \text{and} \quad \sigma_k(t) = \sigma(\tfrac{\ell}{k}, {}_kX(\tfrac{\ell}{k})) \quad \text{for } t \in (\tfrac{\ell}{k}, \tfrac{\ell+1}{k}\wedge T]$$

for $\ell = 1, 2, \ldots, \lfloor kT \rfloor$, the process ${}_kX$ solves the non-diffusion type SDE

$${}_kX(t) = {}_kX(0) + \int_0^t \mu_k(r)\,dr + \int_0^t \sigma_k\,dB \quad \text{for } t \in [0,T], \quad {}_kX(0) = X_0. \tag{17.4}$$

In order to apply Lemma 17.4 to the processes $\{{}_1X(t)\}_{t\in[0,T]}$, $\{{}_2X(t)\}_{t\in[0,T]}$, ... we note that they are continuous martingales by (17.4) and that (17.2) holds since ${}_kX(0) = X_0$ for $k \in \mathbb{N}$. Further, (17.3) holds with $\alpha = 4$ and $\beta = 1$ since (17.4) together with the Burkholder-Davis-Gundy inequality for $p = 4$ give

$$\begin{aligned}
&\mathbf{E}\{({}_kX(t) - {}_kX(s))^4\} \\
&\quad = \mathbf{E}\bigg\{\bigg(\int_s^t \mu_k(r)\,dr + \int_s^t \sigma_k\,dB\bigg)^4\bigg\} \\
&\quad \le 8\,\mathbf{E}\bigg\{\bigg(\int_s^t \mu_k(r)\,dr\bigg)^4\bigg\} + 8\,\mathbf{E}\bigg\{\bigg(\int_s^t \sigma_k\,dB\bigg)^4\bigg\} \\
&\quad \le 8\,\mathbf{E}\bigg\{\bigg(\int_s^t 1\,dr\bigg)^2\bigg(\int_s^t \mu_k(r)^2\,dr\bigg)^2\bigg\} + 8\,C(4)\,\mathbf{E}\bigg\{\bigg(\int_s^t \sigma_k(r)^2\,dr\bigg)^2\bigg\} \\
&\quad \le 8\,(T^2 + C(4))\,(t-s)^2 \sup_{(r,x)\in[0,T]\times\mathbb{R}} \big(\mu(r,x)^4 + \sigma(r,x)^4\big)
\end{aligned}$$

for $s, t \in [0,T]$, see also Exercise 195 below. Hence Lemma 17.4 shows that there exists a continuous process $\{X(t)\}_{t\in[0,T]}$ and a sequence of integers $\{k_j\}_{j=1}^{\infty}$ with $k_j \uparrow \infty$ as $j \to \infty$ such that

$$\bigg(f({}_{k_j}X(t)) - f({}_{k_j}X(s)) - \int_s^t (\mathcal{A}_r f)({}_{k_j}X(r))\,dr\bigg)\, g\big(\{{}_{k_j}X(r)\}_{r\in[0,s]}\big) \tag{17.5}$$

converges in distribution as $j \to \infty$ to

$$\left(f(X(t)) - f(X(s)) - \int_s^t (\mathcal{A}_r f)(X(r))\,dr\right) g\big(\{X(r)\}_{r\in[0,s]}\big) \tag{17.6}$$

for $f \in C_0^2(\mathbb{R})$, $0 \le s \le t \le T$ and any bounded continuous function $g : C([0,s]) \to \mathbb{R}$. In fact, by continuity and boundedness of μ and σ also

$$\left(f({}_{k_j}X(t)) - f({}_{k_j}X(s)) - \int_s^t ({}_{k_j}\mathcal{A}_r f)({}_{k_j}X(r))\,dr\right) g\big(\{{}_{k_j}X(r)\}_{r\in[0,s]}\big) \tag{17.7}$$

converges in distribution to the limit (17.6) for $f \in C_0^2(\mathbb{R})$, $0 \le s \le t \le T$ and any bounded continuous $g : C([0,s]) \to \mathbb{R}$, where

$$({}_k\mathcal{A}_t f)(x) = \mu_k(t) f'(x) + \frac{\sigma_k(t)^2}{2}\, f''(x) \quad \text{for } f \in C_0^2(\mathbb{R})$$

is the generator of the SDE (17.4). This is so because the absolute value of the difference between the random variables in (17.5) and (17.7) is bounded by

$$2|t-s| \sup_{(u,y),(v,z)\in[0,T]\times\operatorname{supp}(f),\, |u-v|\le\frac{1}{k_j},\, |y-z|\le \sup_{r,\hat{r}\in[0,T],|r-\hat{r}|\le\frac{1}{k_j}} |{}_{k_j}X(r)-{}_{k_j}X(\hat{r})|} \Bigg(|\mu(u,y)-\mu(v,z)| + \frac{|\sigma(u,y)^2-\sigma(v,x)^2|}{2}\Bigg) \sup_{x\in\mathbb{R}} \big(|f'(x)|+|f''(x)|\big) \sup_{h\in C([0,s])} |g(h)|, \tag{17.8}$$

which converges to zero in distribution as $j \to \infty$, see Exercise 196 below.

It is sufficient to prove that

$$\{M(t)\}_{t\in[0,T]} = \left\{ f(X(t)) - f(X(0)) - \int_0^t (\mathcal{A}_r f)(X(r))\,dr \right\}_{t\in[0,T]}$$

is a martingale with respect to the filtration $\{\mathcal{F}_t^X\}_{t\in[0,T]} = \{\sigma(X(r) : r \in [0,t])\}_{t\in[0,T]}$ generated by X for any $f \in C_0^2(\mathbb{R})$, which is to say that

$$\mathbf{E}\big\{M(t) - M(s)\,\big|\,\mathcal{F}_s^X\big\} = \mathbf{E}\left\{ f(X(t)) - f(X(s)) - \int_s^t (\mathcal{A}_r f)(X(r))\,dr \,\middle|\, \mathcal{F}_s^X \right\} = 0$$

for $0 \le s \le t \le T$, for $f \in C_0^2(\mathbb{R})$. This in turn is the same thing as

$$\mathbf{E}\left\{ \left(f(X(t)) - f(X(s)) - \int_s^t (\mathcal{A}_r f)(X(r))\,dr \right) I_\Lambda \right\} = 0 \quad \text{for } \Lambda \in \mathcal{F}_s^X, \tag{17.9}$$

for $0 \le s \le t \le T$ and $f \in C_0^2(\mathbb{R})$. By standard approximation methods (17.9) holds if

$$\mathbf{E}\left\{ \left(f(X(t)) - f(X(s)) - \int_s^t (\mathcal{A}_r f)(X(r))\,dr \right) g\big(\{X(r)\}_{r\in[0,s]}\big) \right\} = 0 \tag{17.10}$$

for any bounded continuous $g : C([0,s]) \to \mathbb{R}$, for $0 \le s \le t \le T$ and $f \in C_0^2(\mathbb{R})$, see Exercise 197 below. As the sequence (indexed by j) of random variables in (17.7) is bounded by a deterministic constant it is uniformly integrable. Hence the expected value of the random variables in (17.7) converges to the expected value of the limit random variable in (17.6). However, as (17.4) implies that

$$\Big\{f({}_kX(t))-f({}_kX(0))-\int_0^t({}_k\mathcal{A}_rf)({}_kX(r))\,dr\Big\}_{t\in[0,T]} \tag{17.11}$$

is a martingale with respect to the filtration generated by B for $f\in C_0^2(\mathbb{R})$, see Exercise 198 below, and as ${}_kX$ also is adapted to that filtration, it follows that (recall Exercise 197)

$$\mathbf{E}\Big\{\Big(f({}_kX(t))-f({}_kX(s))-\int_s^t({}_k\mathcal{A}_rf)({}_kX(r))\,dr\Big)g\big(\{{}_kX(r)\}_{r\in[0,s]}\big)\Big\}=0$$

for any bounded continuous $g:C([0,s])\to\mathbb{R}$, for $0\le s\le t\le T$ and $f\in C_0^2(\mathbb{R})$. □

It is unsatisfactory that the Burkholder-Davis-Gundy inequality for the power $p=4$ is used in the proof of Theorem 17.5 as we have not provided a proof of that inequality. This problem is circumvented by the following exercise:

Exercise 195. With the notation of the proof of Theorem 17.5, show by means of direct calculations (not involving the Burkholder-Davis-Gundy inequality) that

$$\mathbf{E}\Big\{\Big(\int_s^t\sigma_k\,dB\Big)^4\Big\}\le 4\,(t-s)^2\sup_{(r,x)\in[0,T]\times\mathbb{R}}\sigma(r,x)^4.$$

Exercise 196. With the notation of the proof of Theorem 17.5, show that the random variable in (17.8) converges to zero in distribution as $j\to\infty$.

Exercise 197. Show that (17.9) holds if (17.10) holds.

Exercise 198. Show that the process in (17.11) is a martingale.

Here is a tricky one:

Exercise 199. Why can we not adopt truncation type of techniques as in the proof of Theorem 12.3 to relax the condition in Theorem 17.5 that the coefficients μ and σ are bounded to local boundedness?

The following result was originally proved by methods based on genious instead of using martingale problem techniques:

Corollary 17.6. (SKOROHOD) *Let the SDE* (9.1) *have bounded and continuous coefficients* $\mu,\sigma:[0,T]\times\mathbb{R}\to\mathbb{R}$. *The SDE has a weak solution for each initial value* X_0.

Exercise 200. Prove Corollary 17.6.

We close this lecture by remarking that it is Lemma 17.4 that is the weak spot of the proof of Theorem 17.5, as there exist must sharper criteria for convergence in distribution of stochastic processes than that lemma. So in order to improve on Theorem 17.5 one should look for better such convergence criteria to base the proof on.

18 Uniqueness of weak solutions to SDE

18.1 Regular conditional probabilities

In order to adress uniqueness of weak solutions to SDE we make crucial use of the following concept of conditional probability that comes with more structure as compared with what is available by the usual direct application of the Radon-Nikodym theorem. The proof of this result is rather long and demanding[18].

Lemma 18.1. (Regular conditional probability) *Let Z be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ taking values in a complete separable metric space S with Borel sets $\mathcal{B}(S)$. Given a σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ there exists a stochastic process $\{Q(B)\}_{B \in \mathcal{B}(S)}$ called a* regular conditional probability *such that $Q(\omega, \cdot)$ is a probability measure on $(S, \mathcal{B}(S))$ for each $\omega \in \Omega$ that satisfies*

$$\mathbf{P}\{Z \in B \,|\, \mathcal{G}\} =^{19} Q(B) \quad \textit{for } B \in \mathcal{B}(S). \tag{18.1}$$

One appealing consequence of the existence of regular conditional probabilities is the following result about conditional characteristic functions.

Corollary 18.2. *With the notation of Lemma 18.1, let Z be an $\mathbb{R}$-valued random variable and suppose that*

$$\mathbf{E}\{e^{i\theta Z} | \mathcal{G}\} = \varphi(\theta) \quad \textit{for } \theta \in \mathbb{R}, \tag{18.2}$$

for some stochastic process $\{\varphi(\theta)\}_{\theta \in \mathbb{R}}$ such that $\varphi(\omega, \cdot)$ is a characteristic function for some probability measure $P^{(\omega)}$ on $\mathbb{R}$ for each $\omega \in \Omega$. Then we have

$$\mathbf{P}\big\{\mathbf{P}\{Z \in B \,|\, \mathcal{G}\} = P^{(\cdot)}(B) \;\textit{ for all } B \in \mathcal{B}(\mathbb{R})\big\} = 1. \tag{18.3}$$

Proof. Use Lemma 18.1 to find a regular conditional probability Q such that

$$\mathbf{P}\{Z \in B \,|\, \mathcal{G}\} = Q(B) \quad \text{for } B \in \mathcal{B}(\mathbb{R}). \tag{18.4}$$

From (18.2) together with (18.4) we readily conclude that

$$\int_{\mathbb{R}} e^{i\theta z}\, dQ(z) = \int_{\mathbb{R}} e^{i\theta z}\, dP^{(\cdot)}(z) \quad \text{for } \theta \in \mathbb{R}, \tag{18.5}$$

[18] See e.g., Karatas and Shreve: "Brownian Motion and Stochastic Calculus", pp. 84-85 together with Parthasarathy: "Probability Measures on Metric Spaces", Chapter V.

[19] Recall that conditional expectations and probabilities are unique and well-defined in the sense of equality almost surely only.

see Exercise 202 below. As the equality in (18.5) is almost sure for each choice of an $\theta\in\mathbb{R}$ we may conclude that

$$\int_{\mathbb{R}} e^{i\theta z}\,dQ(\omega,z)=\int_{\mathbb{R}} e^{i\theta z}\,dP^{(\omega)}(z)\quad\text{simultaneously for all }\ \theta\in\mathbb{Q}, \tag{18.6}$$

for ω on an event with probability 1. As the functions on both sides of the equality (18.6) are continuous in the θ argument for each $\omega\in\Omega$, it follows that

$$\int_{\mathbb{R}} e^{i\theta z}\,dQ(\omega,z)=\int_{\mathbb{R}} e^{i\theta z}\,dP^{(\omega)}(z)\quad\text{simultaneously for all }\ \theta\in\mathbb{R}, \tag{18.7}$$

for ω on an event with probability 1, see Exercise 203 below. From (18.7) in turn we conclude that the probability measures $Q(\omega,\cdot)$ and $P^{(\omega)}(\cdot)$ on $\mathbb{R}$ agree [simultaneously for all Borel subsets $\mathcal{B}(\mathbb{R})$ of $\mathbb{R}$] on an event with probability 1. In view of (18.4), this in turn is the same thing as (18.3). □.

Exercise 201. Give two reasons that we cannot conclude (18.3) immediately from (18.2).

Exercise 202. Show that (18.2) and (18.4) imply (18.5).

Exercise 203. Show that (18.6) implies (18.7).

18.2 Uniqueness of solutions to martingale problems

We are now prepared to state and prove the fourth main result of the course, which concerns uniqueness of weak solutions to SDE.

Theorem 18.3. (Stroock-Varadhan) *Consider the generator $\mathcal{A}_t$ in (16.6) which is supposed to have a locally bounded σ coefficient. Assume that given any $f\in C_0^\infty(\mathbb{R})$*[20]*, $s\in[0,T)$ and $t\in(0,T-s]$ the so called* Cauchy problem

$$\frac{\partial g(r,x)}{\partial r}+(\mathcal{A}_{r+s}g)(x)=0\quad\textit{for}\ \ (r,x)\in[0,t]\times\mathbb{R},\quad g(t,\cdot)=f, \tag{18.8}$$

has a solution $g\in C_B([0,t]\times\mathbb{R})\cap C^{1,2}([0,t]\times\mathbb{R})$. Given an $\mathbb{R}$-valued random variable X_0 a solution $\{X(t)\}_{t\in[0,T]}$ to the martingale problem associated with $\mathcal{A}_t$ such that $X(0)=_D X_0$ has uniquely determined fidi's.

Proof. Given an $s\in[0,T)$, let $\{Y(t)\}_{t\in[0,T-s]}$ solve the martingale problem associated with $\mathcal{A}_{\cdot+s}$. By Theorem 17.3, $\{Y(r)\}_{r\in[0,t]}$ then solves the SDE

$$dY(r)=\mu(r+s,Y(r))\,dr+\sigma(r+s,Y(r))\,dB(r)\quad\text{for }\ r\in[0,t],\ \ \text{for }\ t\in[0,T-s].$$

[20] The class of infinitely many times differentiable functions $f:\mathbb{R}\to\mathbb{R}$ with compact support.

Now, given an $f\in C_0^\infty(\mathbb{R})$ and a $t\in(0,T-s]$, let $g_{s,t}^{(f)}\in C_B([0,t]\times\mathbb{R})\cap C^{1,2}([0,t]\times\mathbb{R})$ solve the Cauchy problem (18.8). By Itô's formula Theorem 16.16 it follows that

$$\begin{aligned}&\{g_{s,t}^{(f)}(r,Y(r))-g_{s,t}^{(f)}(0,Y(0))\}_{r\in[0,t]}\\ &=\bigg\{g_{s,t}^{(f)}(r,Y(r))-g_{s,t}^{(f)}(0,Y(0))-\int_0^r\big(\partial_1 g_{s,t}^{(f)}(\tau,Y(\tau))+(\mathcal{A}_{\tau+s}g)(\tau,Y(\tau))\big)\,d\tau\bigg\}_{r\in[0,t]}\end{aligned}\tag{18.9}$$

is a continuous local martingale, see Exercise 204 below. Since this process is bounded it must in fact be a martingale. This in turn gives

$$\mathbf{E}\{f(Y(t))\}-\mathbf{E}\{g_{s,t}^{(f)}(0,Y(0))\}=\mathbf{E}\{g_{s,t}^{(f)}(t,Y(t))\}-\mathbf{E}\{g_{s,t}^{(f)}(0,Y(0))\}=0 \tag{18.10}$$

for $s\in[0,T)$, $t\in(0,T-s]$ and $f\in C_0^\infty(\mathbb{R})$.

We have to show that any pair of solutions $\{X_1(t)\}_{t\in[0,T]}$ and $\{X_2(t)\}_{t\in[0,T]}$ to the martingale problem associated with $\mathcal{A}_t$ such that $X_1(0)=_D X_2(0)=_D X_0$ have common fidi's. By Exercise 205 below, this holds if

$$\mathbf{E}\bigg\{\prod_{i=1}^n f_i(X_1(t_i))\bigg\}=\mathbf{E}\bigg\{\prod_{i=1}^n f_i(X_2(t_i))\bigg\} \tag{18.11}$$

for $n\in\mathbb{N}$, $0\leq t_1<\ldots<t_n\leq T$ and $f_1,\ldots,f_n\in C_0^\infty(\mathbb{R})$. To prove (18.11), note that (18.11) holds for $n=1$ by application of (18.10) with $s=0$, as that equation gives

$$\mathbf{E}\{f(X_j(t))\}=\mathbf{E}\{g_{0,t}^{(f)}(0,X_j(0))\}=\mathbf{E}\{g_{0,t}^{(f)}(0,X_0)\}\quad\text{for } j=1,2,$$

for $t\in(0,T]$ and $f\in C_0^\infty(\mathbb{R})$. Now assume that we have proved (18.11) for $n=k$ and consider the case $n=k+1$. Choose a regular conditional probability Q_j such that

$$\mathbf{P}\big\{A\,\big|\,X_j(t_1),\ldots,X_j(t_k)\big\}=Q_j(A)\quad\text{for } A\in\sigma(X_j)^{21}, \tag{18.12}$$

for $j=1,2$, see Exercise 206 below. Then the process $\{Z_j(t)\}_{t\in[0,T-t_k]}=\{X_j(t+t_k)\}_{t\in[0,T-t_k]}$ solves the martingale problem associated with $\mathcal{A}_{\cdot+t_k}$ for the filtration $\{\mathcal{F}_{t+t_k}\}_{t\in[0,T-t_k]}$ and the probability measure Q_j for $j=1,2$, because

$$\begin{aligned}&\int_\Lambda\bigg(f(Z_j(t))-f(Z_j(s))-\int_s^t(\mathcal{A}_{r+t_k}f)(Z_j(r))\,dr\bigg)\,dQ_j\\ &=\mathbf{E}\bigg\{I_\Lambda\bigg(f(X_j(t+t_k))-f(X_j(s+t_k))-\int_{s+t_k}^{t+t_k}(\mathcal{A}_rf)(X_j(r))\,dr\bigg)\,\bigg|\,X_j(t_1),\ldots,X_j(t_k)\bigg\}\\ &=\mathbf{E}\bigg\{I_\Lambda\mathbf{E}\bigg\{f(X_j(t+t_k))-f(X_j(s+t_k))-\int_{s+t_k}^{t+t_k}(\mathcal{A}_rf)(X_j(r))\,dr\,\bigg|\,\mathcal{F}_{s+t_k}\bigg\}\\ &\qquad\qquad\bigg|\,X_j(t_1),\ldots,X_j(t_k)\bigg\}\\ &=0\quad\text{for } \Lambda\in\mathcal{F}_{s+t_k},\ 0\leq s<t\leq T-t_k \text{ and } f\in C_0^2(\mathbb{R})\end{aligned}$$

[21] The σ-algebra $\sigma(X_j(t):t\in[0,T])$ generated by the process $\{X_j(t)\}_{t\in[0,T]}$.

since X_j solves the martingale problem for $\mathcal{A}_t$. Hence (18.10) gives

$$\begin{aligned}
\mathbf{E}\Big\{\prod_{i=1}^{k+1} f_i(X_j(t_i))\Big\} &= \mathbf{E}\Big\{\mathbf{E}\{f_{k+1}(X_j(t_{k+1}))\big|X_j(t_1),\ldots,X_j(t_k)\}\prod_{i=1}^{k} f_i(X_j(t_i))\Big\}\\
&= \mathbf{E}\Big\{\Big(\int_\Omega f_{k+1}(Z_j(t_{k+1}-t_k))\,dQ_j\Big)\prod_{i=1}^{k} f_i(X_j(t_i))\Big\}\\
&= \mathbf{E}\Big\{\mathbf{E}^{Q_j}\{g^{(f_{k+1})}_{t_k,t_{k+1}-t_k}(0,Z_j(0))\}\prod_{i=1}^{k} f_i(X_j(t_i))\Big\}\\
&= \mathbf{E}\Big\{\mathbf{E}\{g^{(f_{k+1})}_{t_k,t_{k+1}-t_k}(0,X_j(t_k))\big|X_j(t_1),\ldots,X_j(t_k)\}\prod_{i=1}^{k} f_i(X_j(t_i))\Big\}\\
&= \mathbf{E}\Big\{g^{(f_{k+1})}_{t_k,t_{k+1}-t_k}(0,X_j(t_k))\prod_{i=1}^{k} f_i(X_j(t_i))\Big\} \quad \text{for } j=1,2.
\end{aligned}$$

Here the right-hand side does not depend on j by the assumption that (18.11) holds for $n=k$, as that assumption implies $(X_1(t_1),\ldots,X_1(t_k)) =_D (X_2(t_1),\ldots,X_2(t_k))$ by Exercise 205 below. $\Box$.

Exercise 204. With the notation of the proof of Theorem 18.3, show that the process $\{g^{(f)}_{s,t}(r,Y(r)) - g^{(f)}_{s,t}(0,Y(0))\}_{r\in[0,t]}$ in (18.9) is a local martingale.

Exercise 205. Prove that for two $\mathbb{R}^n$-valued random variables Y and Z we have

$$Y =_D Z \quad \Leftrightarrow \quad \mathbf{E}\Big\{\prod_{i=1}^{n} f_i(Y_i)\Big\} = \mathbf{E}\Big\{\prod_{i=1}^{n} f_i(Z_i)\Big\} \quad \text{for any } f_1,\ldots,f_n\in C_0^\infty(\mathbb{R}).$$

Exercise 206. Explain in detail how (18.12) follows from Lemma 18.1.

Remark 18.4. The Cauchy problem (18.8) has a solution if, for example[22], the coefficients μ and σ for generator $\mathcal{A}_t$ in (16.6) are bounded with σ bounded away from zero[23] and satisfy a *global Hölder condition*, which is to say that there exist constants $K,\alpha>0$ such that

$$|\mu(s,x)-\mu(t,y)|+|\sigma(s,x)-\sigma(t,y)| \leq K\,|(s,x)-(t,y)|^\alpha \quad \text{for } (s,x),(t,y)\in[0,T]\times\mathbb{R}.$$

Corollary 18.5. *Under the hypothesis of Theorem 18.3 the SDE* (9.1) *displays uniqueness for weak solutions.*

Exercise 207. Prove Corollary 18.5.

[22] See e.g., Stroock and Varadhan: "Multidimensional Diffusion Processes", Theorem 3.2.1.

[23] We have $|\sigma(t,x)|\geq\varepsilon$ for $(t,x)\in[0,T]\times\mathbb{R}$ for some constant $\varepsilon>0$.

18.3 Feynman-Kac formula

According to Theorem 18.3 existence of solutions to a Cauchy problem associated with the generator of an SDE implies weak uniqueness of solutions to the SDE. In Theorem 18.6 below we prove the second part of a remarkable duality, namely that existence of weak solutions to the SDE implies uniqueness of solutions to the Cauchy problem[24].

Theorem 18.6. (FEYNMAN-KAC FORMULA) *Consider the generator* $\mathcal{A}_t$ *in* (16.6) *where the coefficients* μ *and* σ *satisfy the global linear growth condition in Definition 12.2. Assume that the SDE*

$$dX(s) = \mu(s, X(s))\, ds + \sigma(s, X(s))\, dB(s) \quad \text{for } s \in [t,T], \quad X(t) = x, \tag{18.13}$$

has a weak solution $\{X_t^x(s)\}_{s\in[t,T]}$ *for each* $x \in \mathbb{R}$ *and each* $t \in [0,T)$. *Consider the Cauchy problem*

$$\frac{\partial u(t,x)}{\partial t} + (\mathcal{A}_t u)(t,x) + k(t,x)u(t,x) = g(t,x) \quad \text{for } (t,x) \in [0,T]\times\mathbb{R}, \quad u(T,\cdot) = f, \tag{18.14}$$

where $k : [0,T]\times\mathbb{R} \to (-\infty, 0]$ *is measurable, while* $f \in C(\mathbb{R})$, $g \in C([0,T]\times\mathbb{R})$ *and the solution* $u \in C^{1,2}([0,T]\times\mathbb{R})$ *satisfy the polynomial growth condition*

$$|f(x)| + |g(t,x)| + |u(t,x)| \leq C\,(1 + |x|^{2\alpha}) \quad \text{for } (t,x) \in [0,T]\times\mathbb{R}, \tag{18.15}$$

for some constants $C > 0$ *and* $\alpha \geq 1$. *Any such solution* u *to* (18.14) *is given by*

$$u(t,x) = \mathbf{E}\bigg\{ f(X_t^x(T)) \exp\bigg[\int_t^T k(r, X_t^x(r))\, dr\bigg] - \int_t^T g(s, X_t^x(s)) \exp\bigg[\int_t^s k(r, X_t^x(r))\, dr\bigg] ds \bigg\}$$

for $(t,x) \in [0,T]\times\mathbb{R}$. *In particular, any such solution* u *to* (18.14) *is unique.*

Proof. By Itô's formula Theorem 16.16 together with (18.13) and (18.14), we have

$$\begin{aligned}
& d\bigg(u(s, X_t^x(s)) \exp\bigg[\int_t^s k(r, X_t^x(r))\, dr\bigg]\bigg) \\
&= \Big(\partial_1 u\, ds + \partial_2 u\, dX_t^x(s) + \frac{1}{2}\,\partial_{22}^2 u\, d[X_t^x](s) + u\,k\,ds\Big) \exp\big[\ldots\big] \\
&= \Big(\partial_1 u\, ds + \partial_2 u\,\big(\mu\, ds + \sigma\, dB(s)\big) + \frac{1}{2}\,\partial_{22}^2 u\,\sigma^2\, ds + u\,k\,ds\Big) \exp\big[\ldots\big] \\
&= \Big(\partial_2 u(s, X_t^x(s))\sigma(s, X_t^x(s))\, dB(s) + g(s, X_t^x(s))\, ds\Big) \exp\bigg[\int_t^s k(r, X_t^x(r))\, dr\bigg]
\end{aligned} \tag{18.16}$$

[24]Many prominent probabilists hold Theorem 18.6 as one of their absolute favorite results in probability theory, simply because it shows how one of the argubly most important problems in pure mathematics can be solved by means of probabilistic methods[25].

[25]That is, we are not the less talented younger brothers and sisters to the pure mathematicians anymore!

for $s\in[t,T]$. Introducing the stopping time

$$\tau_n=\inf\{s\in[t,T]:|X_t^x(s)|\geq n\},$$

it follows that the integral of the left-hand side of (18.16) over the interval $[t,T\wedge\tau_n]$

$$\begin{aligned}
u(T\wedge\tau_n,X_t^x(T\wedge\tau_n))\,&\exp\bigg[\int_t^{T\wedge\tau_n}k(r,X_t^x(r))\,dr\bigg]\\
&-u(t,X_t^x(t))\,\exp\bigg[\int_t^{t}k(r,X_t^x(r))\,dr\bigg]\\
=f(X_t^x(T))\,&\exp\bigg[\int_t^{T}k(r,X_t^x(r))\,dr\bigg]1_{\{\tau_n\geq T\}}\\
&+u(\tau_n,X_t^x(\tau_n))\,\exp\bigg[\int_t^{\tau_n}k(r,X_t^x(r))\,dr\bigg]1_{\{\tau_n<T\}}\\
&-u(t,x)
\end{aligned}$$

is equal to the integral of the right-hand side of (18.16) over the interval $[t,T\wedge\tau_n]$

$$\begin{aligned}
&\int_t^{T\wedge\tau_n}g(s,X_t^x(s))\,\exp\bigg[\int_t^{s}k(r,X_t^x(r))\,dr\bigg]\,ds\\
&\quad+\int_t^{T\wedge\tau_n}\partial_2u(s,X_t^x(s))\,\exp\bigg[\int_t^{s}k(r,X_t^x(r))\,dr\bigg]\sigma(s,X_t^x(s))\,dB(s).
\end{aligned}$$

Here it is obvious (as k is non-positive) that

$$\mathbf{E}\bigg\{\int_t^{T\wedge\tau_n}\partial_2u(s,X_t^x(s))\,\exp\bigg[\int_t^{s}k(r,X_t^x(r))\,dr\bigg]\sigma(s,X_t^x(s))\,dB(s)\bigg\}=0,\qquad(18.17)$$

see Exercise 209 below. Hence it follows from rearrangement that

$$\begin{aligned}
u(t,x)=\mathbf{E}\bigg\{&f(X_t^x(T))\,\exp\bigg[\int_t^{T}k(r,X_t^x(r))\,dr\bigg]1_{\{\tau_n\geq T\}}\bigg\}\\
&+\mathbf{E}\bigg\{u(\tau_n,X_t^x(\tau_n))\,\exp\bigg[\int_t^{\tau_n}k(r,X_t^x(r))\,dr\bigg]1_{\{\tau_n<T\}}\bigg\}\\
&-\mathbf{E}\bigg\{\int_t^{T\wedge\tau_n}g(s,X_t^x(s))\,\exp\bigg[\int_t^{s}k(r,X_t^x(r))\,dr\bigg]\,ds\bigg\}.
\end{aligned}\qquad(18.18)$$

By a modification of the argument employed to show that (the linear growth condition for μ and σ) gives (12.7), we see that

$$\mathbf{E}\Big\{\sup_{s\in[t,T]}X_t^x(s)^{2\alpha}\Big\}\leq D\,(1+|x|^{2\alpha})\,\mathrm{e}^{D(T-t)}\quad\text{for }\;(t,x)\in[0,T]\times\mathbb{R},\qquad(18.19)$$ [26]

where D is a constant that depends on T, the global linear growth coefficient C and $\alpha\geq 1$ only[27], see Exercise 210 below. From (18.19) in turn together with dominated convergence and the assumed polynomial growth conditions on f, u and g, the non-positivity of k and the continuyity of X_t^x (implying that $\tau_n\to\infty$ as $n\to\infty$, we

[26] Unless $\alpha=1$, the proof of (18.19) will require the Burkholder-Davis-Gundy equality with $p=2\alpha$.

[27] You may employ Grönwall's lemma here if you like.

conclude that the right-hand side of (18.18) converges to the right-hand side of the Feynman-Kac formula as $n \to \infty$, see Exercise 211 below. □

Exercise 208. Explain what exactly is meant by a solution to the SDE (18.13) (that takes off at time s rather than time 0).

Exercise 209. Prove (18.17).

Exercise 210. Prove (18.19).

Exercise 211. Prove that the right-hand side of (18.18) converges to the right-hand side of the Feynman-Kac formula as $n \to \infty$.

Exercise 212. Show that the polynomial growth conditions (18.15) on the functions f and/or g can be dropped if f is non-negative and/or g is non-positive.

Exercise 213. Explain how we can deduce the claimed uniqueness of the solution u to the Cauchy problem (18.14) in Theorem 18.6 from the Feynman-Kac formula.

19 Multidimensional SDE

In this final chapter of these lecture notes we show how our results about existence and uniqueness of solutions to SDE extend to a multidimensional setting. It turns out that the new difficulties we encounter are almost exclusivly of a notational character, and that virtually no new ideas or methods are required.

19.1 Multidimensional BM and Itô processes

Definition 19.1. *An $\mathbb{R}^d$-valued stochastic process $\{B(t)\}_{t\geq 0}$ with $B(0)=0$ is a* multidimensional $\mathbb{R}^d$-valued Brownian motion (BM) *if it has the following properties:*

- (Continuity) $[0,\infty)\ni t \curvearrowright B(\omega,t)\in\mathbb{R}^d$ *is continuous for all (or almost all) $\omega\in\Omega$;*
- (Independent increments) $B(t)-B(s)$ *is independent of $\{B(r)\}_{r\in[0,s]}$ for $0\leq s<t$;*
- (Stationary normal increments) $B(t)-B(s)$ *is zero-mean normal distributed in $\mathbb{R}^d$ with covariance matrix $(t-s)\,I$ for $0\leq s<t$, where I is the identity matrix.*

Definition 19.2. *Let $\{B(t)\}_{t\geq 0}$ be an $\mathbb{R}^d$-valued BM that is adapted to an augmented filtration $\{\mathcal{F}_t\}_{t\geq 0}$ on a complete probability space, and that is such that $B(t)-B(s)$ is independent of $\mathcal{F}_s$ for $0\leq s<t$. This we call the* usual conditions.

Henceforth we assume the usual conditions.

Definition 19.3. *Let $\{\sigma(t)\}_{t\in[0,T]}$ be an $\mathbb{R}_{n|d}$-matrix valued stochastic process such that $\sigma_{i,j}\in P_T$ for $i=1,\ldots,n$ and $j=1,\ldots,d$. We define the multidimensional $\mathbb{R}^n$-*valued Itô integral process *$\{\int_0^t \sigma\, dB\}_{t\in[0,T]}$ of σ with respect to an $\mathbb{R}^d$-valued BM by*

$$\left(\int_0^t \sigma\, dB\right)_i = \sum_{j=1}^d \int_0^t \sigma_{i,j}\, dB_j \quad \textit{for } t\in[0,T], \ \textit{ for } i=1,\ldots,n.$$

Exercise 214. Our Definition 19.1 of $\mathbb{R}^d$-valued BM B is in fact more restrictive than the definition that is usually employed in other parts of mathematical statistics, where it is instead required that $B(t)-B(s)$ is zero-mean normal distributed in $\mathbb{R}^d$ with covariance matrix $(t-s)\,V$ for $0\leq s<t$, for some non-negative definite matrix V. Explain why the latter more general definition of $\mathbb{R}^d$-valued BM does not give rise to more different Itô integral processes $\{\int_0^t \sigma\, dB\}_{t\in[0,T]}$.

Of course, equipped with multidimensional Itô integral processes, the next step is to introduce multidimensional Itô processes and stochastic differentials.

Definition 19.4. *Let B be an $\mathbb{R}^d$-valued BM. If $\{\mu(t)\}_{t\in[0,T]}$ is an $\mathbb{R}^n$-valued stochastic process with component processes that are measurable and adapted and satisfy*

$$\mathbf{P}\bigg\{\int_0^T |\mu_i(r)|\,dr < \infty\bigg\} = 1 \quad \textit{for } i=1,\dots,n,$$

if $\{\sigma(t)\}_{t\in[0,T]}$ is an $\mathbb{R}_{n|d}$-matrix valued stochastic process such that $\sigma_{i,j} \in P_T$ for $i=1,\dots,n$ and $j=1,\dots,d$, and if $X(0)$ is an $\mathcal{F}_0$-measurable random variable, then we call

$$\{X(t)\}_{t\in[0,T]} = \bigg\{X(0) + \int_0^t \mu(r)\,dr + \int_0^t \sigma\,dB\bigg\}_{t\in[0,T]}, \tag{19.1}$$

a multidimensional $\mathbb{R}^n$-valued Itô process *and*

$$dX(t) = \mu(t)\,dt + \sigma(t)\,dB(t)$$

the corresponding multidimensional $\mathbb{R}^n$-valued stochastic differential.

Exercise 215. Show that if $\{X(t)\}_{t\in[0,T]}$ and $\{Y(t)\}_{t\in[0,T]}$ are multidimensional Itô processes (with respect to a common filtration and multidimensional BM), then $\{(X(t),Y(t))\}_{t\in[0,T]}$ is also a multidimensional Itô process.

Exercise 216. Show that the quadratic covariation process between two components X_i and X_j of the $\mathbb{R}^n$-valued Itô process $\{X(t)\}_{t\in[0,T]}$ given in (19.1) satisfies

$$\{[X_i,X_j](t)\}_{t\in[0,T]} = \bigg\{\int_0^t \sum_{k=1}^d \sigma_{i,k}(r)\sigma_{j,k}(r)\,dr\bigg\}_{t\in[0,T]} \quad \text{for } i=1,\dots,n.$$

Theorem 19.5. (Itô formula) *For an $\mathbb{R}^n$-valued Itô process $\{X(t)\}_{t\in[0,T]}$ and a function $f\in C^{1,2}([0,T]\times\mathbb{R}^n)$, we have*

$$df(t,X(t)) = f_t'(t,X(t))\,dt + \sum_{i=1}^n f_{x_i}'(t,X(t))\,dX_i(t) + \frac{1}{2}\sum_{i,j=1}^n f_{x_ix_j}''(t,X(t))\,d[X_i,X_j](t).$$

Exercise 217. Prove Theorem 19.5.

Note that in view of Exercise 215 we do not need separate Itô formulas of the type in Theorem 9.12 for functions $f(X(t),Y(t))$ of two multidimensional Itô processes, as such formulas are already contained in the Itô formula Theorem 19.5, see also Exercise 218 below.

Exercise 218. Use Exercise 215 and Theorem 19.5 to derive the Itô formula for a function $f(X(t), Y(t))$ of two multidimensional Itô processes.

We will need versions of Definition 19.4 and Theorem 9.12 for Itô integrals and Itô processes with respect to continuous local martingales:

Definition 19.6. *Let $\sigma_1 \in P(M_1)_T, \ldots, \sigma_n \in P(M_n)_T$ where $\{M_1(t)\}_{t\in[0,T]}, \ldots, \{M_n(t)\}_{t\in[0,T]}$ are continuous local martingales, let $\{\mu_1(t)\}_{t\in[0,T]}, \ldots, \{\mu_n(t)\}_{t\in[0,T]}$ be measurable and adapted processes such that $\int_0^T |\mu_1(t)|\,dt, \ldots, \int_0^T |\mu_n(t)|\,dt < \infty$ with probability 1, and let $X(0)$ be an $\mathcal{F}_0$-measurable random variable. We call the $\mathbb{R}^n$-valued stochastic process $\{X(t)\}_{t\in[0,T]}$ with components given by*

$$\{X_i(t)\}_{t\in[0,T]} = \left\{X_i(0) + \int_0^t \mu_i(r)\,dr + \int_0^t \sigma_i\,dM_i\right\}_{t\in[0,T]} \quad \text{for } i = 1, \ldots, n,$$

a multidimensional $\mathbb{R}^n$-valued Itô process *and*

$$dX(t) = \mu(t)\,dt + \sigma(t)\,dM(t)$$

the corresponding multidimensional $\mathbb{R}^n$-valued stochastic differential.

Theorem 19.5. (Itô formula) *For an $\mathbb{R}^n$-valued Itô process $\{X(t)\}_{t\in[0,T]}$ and a function $f \in C^{1,2}([0,T]\times\mathbb{R}^n)$, we have*

$$df(t, X(t)) = f_t'(t, X(t))\,dt + \sum_{i=1}^n f_{x_i}'(t, X(t))\,dX_i(t) + \frac{1}{2}\sum_{i,j=1}^n f_{x_ix_j}''(t, X(t))\,d[X_i, X_j](t).$$

Exercise 217. Prove Theorem 19.5.

Of course, Corollary 16.9 is essential for applications of Theorem 19.5.

19.2 Multidimensional SDE

Definition 19.7. *A* multidimensional stochastic differential equation (SDE) of diffusion type *is given by*

$$dX(t) = \mu(t, X(t))\,dt + \sigma(t, X(t))\,dB(t) \quad \text{for } t \in [0,T], \quad X(0) = X_0, \tag{19.2}$$

where $\mu : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0,T]\times\mathbb{R}^n \to \mathbb{R}_{n|d}$ are measurable (in each of their components) coefficient functions, *X_0 is an $\mathcal{F}_0$-measurable $\mathbb{R}^n$-valued random variable called the* initial value, *and $\{B(t)\}_{t\geq 0}$ is an $\mathbb{R}^d$-valued BM.*

A solution to the multidimensional SDE (19.2) is an $\mathbb{R}^n$-valued Itô process $\{X(t)\}_{t\in[0,T]}$ such that $X(0)=X_0$ and

$$X(t) = X(0) + \int_0^t \mu(r, X(r))\, dr + \int_0^t \sigma(r, X(r))\, dB(r) \quad \text{for } t \in [0,T]. \qquad (19.3)$$

Of course, for this solution to be well-defined X must be an adapted process that is continuous with probability 1 such that

$$\mathbf{P}\Big\{\bigcap_{i=1}^n \Big\{\int_0^T |\mu_i(t,X(t))|\, dt < \infty\Big\} \cap \bigcap_{i=1}^n \bigcap_{j=1}^d \Big\{\int_0^T \sigma_{i,j}(t,X(t))^2\, dt < \infty\Big\}\Big\} = 1.$$

Definition 19.8. *An $\mathbb{R}^n$-valued process $\{X(t)\}_{t\in[0,T]}$ is a* strong solution *to the SDE* (19.2) *for a given $\mathbb{R}^d$-valued BM B and a given initial value X_0, if* (19.3) *holds for this choice of B with $X(0)=X_0$ with probability 1.*

Definition 19.9. *A $\mathbb{R}^n$-valued process $\{X(t)\}_{t\in[0,T]}$ is a* weak solution *to the SDE* (19.2) *if there exists an $\mathbb{R}^d$-valued BM B such that* (19.3) *holds with $X(0) =_D X_0$.*

Definition 19.10. *Solutions to the SDE* (19.2) *have* strong uniqueness *if whenever $\{X_1(t)\}_{t\in[0,T]}$ and $\{X_2(t)\}_{t\in[0,T]}$ are strong solutions for a common given an $\mathbb{R}^d$-valued BM B and a common given initial value X_0, it holds that*

$$\mathbf{P}\big\{X_1(t) = X_2(t) \text{ for all } t \in [0,T]\big\} = 1.$$

Definition 19.11. *Solutions to the SDE* (19.2) *have* weak uniqueness *if whenever $\{X_1(t)\}_{t\in[0,T]}$ and $\{X_2(t)\}_{t\in[0,T]}$ are weak solutions they have common fidi's.*

19.3 Strong uniqueness

Definition 19.12. *The coefficients $\mu : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0,T]\times\mathbb{R}^n \to \mathbb{R}_{n|d}$ of the multidimensional SDE* (19.2) *are said to satisfy a* local Lipschitz condition *if to each $m \in \mathbb{N}$ there exists a constant $K_m > 0$ such that*

$$\|\mu(t,x)-\mu(t,y)\|^2 + \sum_{i=1}^n \sum_{j=1}^d \big(\sigma_{i,j}(t,x)-\sigma_{i,j}(t,y)\big)^2 \le K_m\, \|x-y\|^2$$

for $t \in [0,T]$ and $x, y \in \mathbb{R}^n$ with $\|x\|, \|y\| \le m$, where $\|\cdot\|$ denotes the Euclidian norm.

Theorem 19.13. *If the coefficients of the multidimensional SDE* (19.2) *satisfy a local Lipschitz condition, then we have strong uniqueness for solutions to the SDE.*

Proof. Consider two strong solutions $\{{}_1X(t)\}_{t\in[0,T]}$ and $\{{}_2X(t)\}_{t\in[0,T]}$ to the SDE (19.2) together with the stopping time

$$\tau_m = \inf\big\{t\in[0,T] : \|{}_1X(t)\|\geq m\big\} \wedge \inf\big\{t\in[0,T] : \|{}_2X(t)\|\geq m\big\}.$$

Then the process ${}_iX^{(m)}(t) = {}_iX(t\wedge\tau_m)$ satisfies $\|{}_iX^{(m)}(t)\|\leq m$ for $t\in[0,T]$ and

$${}_iX^{(m)}(t) = \int_0^t \mu(r,{}_iX(r))\, I_{[0,\tau_m]}(r)\, dr + \int_0^t \sigma(\,\cdot\,,{}_iX)\, I_{[0,\tau_m]}\, dB \quad \text{for } t\in[0,T],$$

for $i=1,2$. Since $\tau_m\uparrow\infty$ as $m\to\infty$ it is sufficent to prove that $\mathbf{P}\{{}_1X^{(m)}(t) = {}_2X^{(m)}(t)\} = 1$ for $t\in[0,T]$ and $m\in\mathbb{N}$, recall Exercise 121. However, by the triangle inequality and the elementary inequality $(x+y)^2\leq 2x^2+2y^2$ together with isometry, the Cauchy-Schwarz inequality and the local Lipschitz condition, we have

$$\begin{aligned}
&\mathbf{E}\{\|{}_1X^{(m)}(t) - {}_2X^{(m)}(t)\|^2\}\\
&\leq 2\,\mathbf{E}\bigg\{\bigg\|\int_0^t \big(\mu(r,{}_1X(r)) - \mu(r,{}_2X(r))\big) I_{[0,\tau_m]}(r)\, dr\bigg\|^2\bigg\}\\
&\qquad + 2\,\mathbf{E}\bigg\{\bigg\|\int_0^t \big(\sigma(\,\cdot\,,{}_1X) - \sigma(\,\cdot\,,{}_2X)\big) I_{[0,\tau_m]}\, dB\bigg\|^2\bigg\}\\
&= 2\,\mathbf{E}\bigg\{\sum_{i=1}^n\bigg(\int_0^t \big(\mu_i(r,{}_1X(r)) - \mu_i(r,{}_2X(r))\big) I_{[0,\tau_m]}(r)\, dr\bigg)^2\bigg\}\\
&\qquad + 2\,\mathbf{E}\bigg\{\sum_{i=1}^n\bigg(\sum_{j=1}^d\int_0^t \big(\sigma_{i,j}(\,\cdot\,,{}_1X) - \sigma_{i,j}(\,\cdot\,,{}_2X)\big) I_{[0,\tau_m]}\, dB_j\bigg)^2\bigg\}\\
&\leq 2\,\mathbf{E}\bigg\{\sum_{i=1}^n\bigg(\int_0^t dr\bigg)\bigg(\int_0^t \big(\mu_i(r,{}_1X(r)) - \mu_i(r,{}_2X(r))\big)^2 I_{[0,\tau_m]}(r)\, dr\bigg)\bigg\} \qquad (19.4)\\
&\qquad + 2d\,\mathbf{E}\bigg\{\sum_{i=1}^n\sum_{j=1}^d\bigg(\int_0^t \big(\sigma_{i,j}(\,\cdot\,,{}_1X) - \sigma_{i,j}(\,\cdot\,,{}_2X)\big) I_{[0,\tau_m]}\, dB_j\bigg)^2\bigg\}\\
&= 2t\int_0^t \mathbf{E}\Big\{\big\|\mu_i(r,{}_1X(r)) - \mu_i(r,{}_2X(r))\big\|^2 I_{[0,\tau_m]}(r)\Big\}\, dr\\
&\qquad + 2d\int_0^t\sum_{i=1}^n\sum_{j=1}^d \mathbf{E}\Big\{\big(\sigma_{i,j}(r,{}_1X(r)) - \sigma_{i,j}(r,{}_2X(r))\big)^2 I_{[0,\tau_m]}(r)\Big\}\, dr\\
&\leq 2\,(T{+}d)\,K_m\int_0^t \mathbf{E}\{\|{}_1X^{(m)}(r) - {}_2X^{(m)}(r)\|^2\}\, dr \quad \text{for } t\in[0,T].
\end{aligned}$$

Hence an application of Grönwall's lemma with $C=0$, $u(t) = 2\,(T{+}d)\,K_m$ and $v(t) = \mathbf{E}\{\|{}_1X^{(m)}(t) - {}_2X^{(m)}(t)\|^2\}$ (which is continuous by continuity of ${}_1X^{(m)}$ and ${}_2X^{(m)}$ together with the bounded convergence theorem) gives $v(t)=0$. □

Exercise 219. Let $\operatorname{diag}(x) = \operatorname{diag}(x_1, \ldots, x_n)$ denote the diagonal $\mathbb{R}_{n|n}$-matrix with diagonal entries $x_1, \ldots, x_n$ for $(x_1, \ldots, x_n) \in \mathbb{R}^n$. Solve the following non-diffusion type multidimensional SDE for the so called *multidimensional stochastic exponential* $\{\mathcal{E}(X)(t)\}_{t\in[0,T]}$ of an $\mathbb{R}^n$-valued Itô process $\{X(t)\}_{t\in[0,T]}$

$$d\mathcal{E}(X)(t) = \operatorname{diag}(\mathcal{E}(X)(t)\, dX(t) \quad \text{for } t \in [0,T], \quad \mathcal{E}(X)(0) = 1.$$

19.4 Strong existence

Definition 19.14. *The coefficients* $\mu : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ *and* $\sigma : [0,T]\times\mathbb{R}^n \to \mathbb{R}_{n|d}$ *of the multidimensional SDE* (19.2) *are said to satisfy a* global linear growth condition *if there exists a constant* $C > 0$ *such that*

$$\|\mu(t,x)\|^2 + \sum_{i=1}^{n}\sum_{j=1}^{d} \sigma_{i,j}(t,x)^2 \leq C\,(1+\|x\|^2) \quad \textit{for } t\in[0,T] \textit{ and } x \in \mathbb{R}^n.$$

Theorem 19.15. *If the coefficients of the multidimensional SDE* (19.2) *satisfy a local Lipschitz condition and a global linear growth condition, then there exists a strong solution to the SDE for every given BM B and any given initial value X_0 that is strongly unique.*

Proof. As we have uniqueness by Theorem 19.13 it is sufficient to prove existence. Further, it is sufficient to prove existence for any square-integrable initial value $\mathbf{E}\{\|X_0\|^2\} < \infty$ by the truncation argument we used in the proof of Theorem 12.3, recall Exercise 131. Now consider a Picard-Lindelöf iteration where $X_0(t) = X_0$ and

$$X_{k+1}(t) = X_0 + \int_0^t \mu(r, X_k(r))\, dr + \int_0^t \sigma(\,\cdot\,, X_k)\, dB \quad \text{for } t \in [0,T], \tag{19.5}$$

for $k \geq 0$. To establish that the process X_{k+1} on the left-hand side of (19.5) is well-defined it is sufficient to show that the process X_k on the right-hand side satisfies

$$\mathbf{E}\Big\{\sup_{t\in[0,T]} \|X_k(t)\|^2\Big\} < \infty, \tag{19.6}$$

because by analogy with (12.4) the global linear growth then gives

$$\mathbf{E}\Big\{\int_0^t \sigma_{i,j}(r, X_k(r))^2\, dr\Big\} \leq \int_0^t C\big(1+\mathbf{E}\{\|X_k(r)^2\|\}\big)\, dr < \infty \quad \text{for } t\in[0,T],$$

so that $\sigma_{i,j}(\,\cdot\,, X_k) \in E_T$ for $i = 1, \ldots, n$ and $j = 1, \ldots, d$, while by analogy with (12.5)

$$\mathbf{E}\Big\{\int_0^t |\mu_i(r, X_k(r))|\, dr\Big\} \leq \sqrt{T}\left[\int_0^t C\big(1+\mathbf{E}\{\|X_k(r)\|^2\}\big)\, dr\right]^{1/2} < \infty \quad \text{for } t\in[0,T],$$

so that $\int_0^T |\mu_i(r, X_k(r)|\, dr < \infty$ with probability 1 for $i = 1, \ldots, n$. Now, (19.6) holds trivially for $k = 0$ since $\mathbf{E}\{\|X_0\|^2\} < \infty$. Further, if (19.6) holds for a certain $k \in \mathbb{N}$, then by analogy with (19.4), (19.5) and the elementary inequality $(x+y+z)^2 \leq 3\,x^2 + 3\,y^2 + 3\,z^2$ together with Doob's maximal inequality and isometry give

$$\begin{aligned}
&\mathbf{E}\Big\{\sup_{s\in[0,t]} \|X_{k+1}(s)\|^2\Big\} \\
&\leq 3\,\mathbf{E}\{\|X_0\|^2\} + 3\,\mathbf{E}\bigg\{\sup_{s\in[0,t]}\bigg\|\int_0^s \mu(r, X_k(r))\,dr\bigg\|^2\bigg\} + 3\,\mathbf{E}\bigg\{\sup_{s\in[0,t]}\bigg\|\int_0^s \sigma(\,\cdot\,, X_k)\,dB\bigg\|^2\bigg\} \\
&\leq 3\,\mathbf{E}\{\|X_0\|^2\} + (3T{+}12d)\int_0^t C\big(1 + \mathbf{E}\{\|X_k(r)\|^2\}\big)\,dr \\
&\leq D + D\int_0^t \mathbf{E}\Big\{\sup_{r\in[0,s]}\|X_k(r)\|^2\Big\}\,ds \quad \text{for } t\in[0,T],
\end{aligned} \tag{19.7}$$

where $D = 3\,\mathbf{E}\{X_0^2\} + (3T{+}12d)\,(TC{+}C)$. Hence (19.6) holds for $k{+}1$, so that (19.6) holds for all $k \in \mathbb{N}$ by induction. In fact, recalling Exercise 132, from (19.7) we get by iteration the following stronger version of (19.6)

$$\mathbf{E}\Big\{\sup_{s\in[0,t]}\|X_k(s)\|^2\Big\} \leq (D\vee 1)\,\mathrm{e}^{D\,t} \quad \text{for } t\in[0,T] \text{ and } k\geq 0. \tag{19.8}$$

Now assume that a global Lipschitz condition holds, that is, a local Lipschitz condition where we may choose the Lipscitz constant K_m not to depend on m. We will relax the global Lipschitz condition to a local Lipschitz condition later. By application of the arguments used to obtain (19.7) [this time using the inequality $(x+y)^2 \leq 2\,x^2 + 2\,y^2$ instead of $(x+y+z)^2 \leq 3\,x^2 + 3\,y^2 + 3\,z^2$], we then readily get

$$\mathbf{E}\bigg\{\sup_{s\in[0,t]}\big\|X_{k+1}(s) - X_k(s)\big\|^2\bigg\} \leq L\int_0^t \mathbf{E}\bigg\{\sup_{r\in[0,s]}\big\|X_k(r) - X_{k-1}(r)\big\|^2\bigg\}\,ds \tag{19.9}$$

for $t \in [0,T]$ and $k \in \mathbb{N}$, where $L = (2T{+}8d)\,K$, recall Exercise 133. Using (19.6) it therefore follows using induction by analogy with (12.9) and Exercise 134 that

$$\mathbf{E}\bigg\{\sup_{s\in[0,t]}\big\|X_{k+1}(s) - X_k(s)\big\|^2\bigg\} \leq M\,\frac{(LT)^k}{k!} \quad \text{for } t\in[0,T] \text{ and } k\in\mathbb{N}, \tag{19.10}$$

for some constant $M < \infty$. Using (19.10) in turn we may conclude by the same arguments that were employed in the proof of Theorem 12.3 that there exists an $\mathbb{R}^n$-valued stochastic process $\{X(t)\}_{t\in[0,T]}$ that is continuous with probability 1 such that $\sup_{t\in[0,T]}\|X_k(t) - X(t)\| \to 0$ as $k\to\infty$ with probability 1, and such that

$$\mathbf{E}\Big\{\sup_{t\in[0,T]}\|X(t)\|^2\Big\} < \infty \quad \text{with} \quad \mathbf{E}\bigg\{\sup_{t\in[0,T]}\big\|X(t) - X_k(t)\big\|^2\bigg\} \leq \frac{2M\,\mathrm{e}^{2LT}}{2^k} \tag{19.11}$$

for $k \in \mathbb{N}$. Recalling (19.5) and (19.6) it follows that the Itô process

$$Y(t) = X_0 + \int_0^t \mu(r, X(r))\,dr + \int_0^t \sigma(\,\cdot\,, X)\,dB \quad \text{for } t\in[0,T]$$

is well-defined. In order to establish the existence of a solution it is therefore sufficient to show that $\mathbf{P}\{X(t)=Y(t) \text{ for } t\in[0,T]\}=1$. However, by an obvious version of (19.9) together with (19.11), we have

$$\mathbf{E}\Big\{\sup_{t\in[0,T]}\|Y(t)-X_{k+1}(t)\|^2\Big\}\leq L\int_0^T\mathbf{E}\Big\{\sup_{s\in[0,t]}\|X(s)-X_k(s)\|^2\Big\}\,dt\leq\frac{2MLT\,\mathrm{e}^{2LT}}{2^k}$$

for $k\in\mathbb{N}$. From this in turn together with (19.11) we readily get $\mathbf{P}\{X=Y\}=1$.

It remains to show how to relax the global Lipschitz condition to a local Lipschitz condition. To that end write $\hat{x}=x/\|x\|$ and consider the truncated coefficients

$$\mu^{(m)}(t,x)=\begin{cases}\mu(t,x) & \text{for } \|x\|\leq m\\ \mu(t,m\hat{x}) & \text{for } \|x\|>m\end{cases} \quad\text{and}\quad \sigma^{(m)}(t,x)=\begin{cases}\sigma(t,x) & \text{for } \|x\|\leq m\\ \sigma(t,m\hat{x}) & \text{for } \|x\|>m\end{cases}.$$

The truncated coefficients $\mu^{(m)}$ and $\sigma^{(m)}$ satisfy a global linear growth condition with the same growth constant C as the non-truncated coefficients, and a global Lipschitz condition with global Lipschitz constant $2K_m$ where K_m is the local Lipschitz constant for the non-truncated coefficients, see Exercise 220 below. We may now finish off the proof of the theorem in the same way as we finished off the proof of Theorem 12.3. □

Exercise 220. Show that the coefficients $\mu^{(m)}$ and $\sigma^{(m)}$ in the proof of Theorem 19.15 satisfy a global linear growth condition with the same growth constant C as the coefficients μ and σ, and that they satisfy a global Lipschitz condition with Lipschitz constant $2K_m$ where K_n is the local Lipschitz constant for μ and σ.

19.5 Paul Lévy's characterization of BM

Theorem 19.16. (Paul Lévy's characterization of BM) *An $\mathbb{R}^n$-valued stochastic process $\{M(t)\}_{t\in[0,T]}$ each component $\{M_i(t)\}_{t\in[0,T]}$, $i=1,\ldots,n$, of which is a continuous local martingale is an $\mathbb{R}^n$-valued BM if and only if it holds that $[M_i,M_j](t)=\delta(i-j)\,t$ for $t\in[0,T]$, for $i,j=1,\ldots,n$.*

Proof. $\boxed{\Leftarrow}$ Assume that $[M_i,M_j](t)=\delta(i-j)\,t$ for $t\in[0,T]$, for $i,j=1,\ldots,n$. Let $\langle\cdot,\cdot\rangle$ denote the inner product on $\mathbb{R}^n$, pick a constant $\theta\in\mathbb{R}^n$ and define

$$Y_1(t)=\int_0^t Z_1\,d\langle\theta,M\rangle \quad\text{and}\quad Y_2(t)=\int_0^t Z_2\,d\langle\theta,M\rangle \quad \text{for } t\in[0,T],$$

where $Z_1(t)=\cos(\langle\theta,M(t)\rangle)\,\mathrm{e}^{\frac{1}{2}\|\theta\|^2t}$ and $Z_2(t)=\sin(\langle\theta,M(t)\rangle)\,\mathrm{e}^{\frac{1}{2}\|\theta\|^2t}$ for $t\in[0,T]$. [Note that $\{\langle\theta,M(t)\rangle\}_{t\in[0,T]}$ is a continuous local maringale.] As $|Z_1(t)|,|Z_2(t)|\leq\mathrm{e}^{\frac{1}{2}\|\theta\|^2T}$ and $d[\langle\theta,M\rangle](t)=\|\theta\|^2dt$ for $t\in[0,T]$, we have $Z_1,Z_2\in E(\langle\theta,M\rangle)_T$, so that Y_1 and Y_2 are martingales. Further, Itô's formula Theorem 16.16 shows that

$$dZ_1(t) = \frac{1}{2}\|\theta\|^2 Z_1(t)\,dt - Z_2(t)\,d\langle\theta, M(t)\rangle - \frac{1}{2}Z_1(t)\,d[\langle\theta, M\rangle](t) = -dY_2(t),$$
$$dZ_2(t) = \frac{1}{2}\|\theta\|^2 Z_2(t)\,dt + Z_1(t)\,d\langle\theta, M(t)\rangle - \frac{1}{2}Z_2(t)\,d[\langle\theta, M\rangle](t) = \quad dY_1(t).$$

Hence Z_1 and Z_2 are martingales (as Y_1 and Y_2 are martingales). It follows that

$$\mathbf{E}\big\{\mathrm{e}^{i\langle\theta,M(t)\rangle+\frac{1}{2}\|\theta\|^2 t}\,\big|\,\mathcal{F}_s\big\} = \mathbf{E}\big\{Z_1(t)\,\big|\,\mathcal{F}_s\big\} + i\,\mathbf{E}\big\{Z_2(t)\,\big|\,\mathcal{F}_s\big\} = \mathrm{e}^{i\langle\theta,M(s)\rangle+\frac{1}{2}\|\theta\|^2 s}$$

for $0 \le s \le t \le T$, which in turn (as M is adapted) by rearrangement gives

$$\mathbf{E}\big\{\mathrm{e}^{i\langle\theta,M(t)-M(s)\rangle}\,\big|\,\mathcal{F}_s\big\} = \mathrm{e}^{-\frac{1}{2}\|\theta\|^2(t-s)} \quad \text{and} \quad \mathbf{E}\big\{\mathrm{e}^{i\langle\theta,M(t)-M(s)\rangle}\big\} = \mathrm{e}^{-\frac{1}{2}\|\theta\|^2(t-s)}$$

for $0 \le s \le t \le T$. From this we may finish off the proof of the theorem in exactly the same manner as we used (10.1) to finish off the proof of Theorem 10.1. □

Exercise 221. Prove the implication to the right in Theorem 19.16.

19.6 Martingale problems

Definition 19.17. *The* diffusion matrix *of the SDE* (19.2) *is the function* $a : [0,T]\times\mathbb{R}^n \to \mathbb{R}_{n|n}$ *given by*

$$a_{i,j}(t,x) = \sum_{k=1}^{d} \sigma_{i,k}(t,x)\sigma_{j,k}(t,x) \quad \textit{for } (t,x) \in [0,T]\times\mathbb{R}^n, \ \textit{for } i,j = 1,\ldots,n.$$

Definition 19.18. *The* generator *of the SDE* (19.2) *is the differential operator*

$$(\mathcal{A}_t f)(x) = \sum_{i=1}^{n} \mu_i(t,x)\, f'_{x_i}(x) + \sum_{i,j=1}^{n} \frac{a_{i,j}(t,x)}{2}\, f''_{x_i,x_j}(x) \quad \textit{for } f \in C^2(\mathbb{R}^n). \qquad (19.12)$$

Definition 19.19. *A continuous and adapted $\mathbb{R}^n$-valued stochastic process $\{X(t)\}_{t\in[0,T]}$ is a solution to the* local martingale problem *associated with the generator $\mathcal{A}_t$ in* (19.12) *if for any $f \in C^2(\mathbb{R}^n)$ the following stochastic process is a continuous local martingale*

$$\Big\{f(X(t)) - f(X(0)) - \int_0^t (\mathcal{A}_r f)(X(r))\,dr\Big\}_{t\in[0,T]}. \qquad (19.13)$$

Theorem 19.20. *A continuous and adapted $\mathbb{R}^n$-valued stochastic process $\{X(t)\}_{t\in[0,T]}$ is a weak solution to the SDE* (19.2) *if and only if $X(0) =_D X_0$ and X is a solution to the local martingale problem associated with the generator $\mathcal{A}_t$ in* (19.12).

Proof. $\boxed{\Leftarrow}$ Let $\{X(t)\}_{t\in[0,T]}$ solve the local martingale problem associated with the differential operator $\mathcal{A}_t$ in (19.12). Taking $f(x)=x_i$ in (19.13) we see that

$$\{M_i(t)\}_{t\in[0,T]} = \left\{X_i(t) - X_i(0) - \int_0^t \mu_i(r, X(r))\,dr\right\}_{t\in[0,T]}, \quad i=1,\ldots,n, \tag{19.14}$$

are continuous local martingales. Hence we may apply the continuous local martingale version of the Itô formula Theorem 19.5 to conclude that

$$\begin{aligned} &df(X(t)) - (\mathcal{A}_t f)(X(t))\,dt \\ &= \sum_{i=1}^n f'_{x_i}(X(t))\,dX_i(t) + \frac{1}{2}\sum_{i,j=1}^n f''_{x_ix_j}(X(t))\,d[X_i,X_j](t) - (\mathcal{A}_t f)(X(t))\,dt \\ &= \sum_{i=1}^n f'_{x_i}(X(t))\,dM_i(t) + \frac{1}{2}\sum_{i,j=1}^n f''_{x_ix_j}(X(t))\big(d[M_i,M_j](t) - a_{i,j}(t,X(t))\,dt\big) \end{aligned}$$

for $t\in[0,T]$, for $f\in C^2(\mathbb{R}^n)$. As X solves the local martingale problem and the processes in (19.14) are continuous local martingales, it follows that

$$\left\{\int_0^t \frac{1}{2}\sum_{i,j=1}^n f''_{x_ix_j}(X(r))\big(d[M_i,M_j](r) - a_{i,j}(r,X(r))\,dr\big)\right\}_{t\in[0,T]} \tag{19.15}$$

is a continuous local martingale. But as this process has finite variation it has zero quadratic variation and is thus zero by Corollary 16.5. From this in turn we get

$$[M_i,M_j](t) = \int_0^t a_{i,j}(r,X(r))\,dr \quad \text{for } t\in[0,T], \text{ for } i,j=1,\ldots,n, \tag{19.16}$$

see Exercise 222 below. Hence it remains to deduce from (19.16) that

$$M(t) = \int_0^t \sigma(\,\cdot\,,X)\,dW \quad \text{for } t\in[0,T], \tag{19.17}$$

for some $\mathbb{R}^d$-valued BM $\{W(t)\}_{t\in[0,T]}$, as (19.14) together with (19.17) give (19.2). We will prove (19.17) in the particular case when $n=d$, as that turns out to give no loss of generality, see Exercise 228 below.

As the diffusion matrix a is symmetric and non-negative definite, see Exercise 223 below, there exists an orthogonal $\mathbb{R}_{n|n}$-matrix valued process $\{q(t)\}_{t\in[0,T]}$ such that

$$q(t)^{\mathrm{tr}} a(t,X(t))\,q(t) = q(t)^{-1} a(t,X(t))\,q(t) = \mathrm{diag}(\lambda_1(t),\ldots,\lambda_n(t)) \quad \text{for } t\in[0,T], \tag{19.18}$$

for some non-negative stochastic processes $\{\lambda_1(t)\}_{t\in[0,T]},\ldots,\{\lambda_n(t)\}_{t\in[0,T]}$. Here the processes q and $\lambda_1,\ldots,\lambda_n$ are measurable and adapted, see Exercise 224 below. Now we may define continuous local martingales

$$\{N_i(t)\}_{t\in[0,T]} = \left\{\sum_{k=1}^n \int_0^t q_{k,i}\,dM_k\right\}_{t\in[0,T]} \quad \text{for } i=1,\ldots,n. \tag{19.19}$$

This is so because $q(t)_{k,i}^2 \le \sum_{k=1}^n q(t)_{k,i}^2 = 1$ for $t \in [0,T]$ and $k,i=1,\ldots,n$, so that

$$\int_0^T q_{k,i}^2\, d[M_k] \le \int_0^T d[M_k] = [M_k](T) < \infty,$$

giving $q_{k,i} \in P(M_k)_T$. From (19.16) and (19.18) together with Corollary 16.9 we get

$$[N_i, N_j](t) = \sum_{k,\ell=1}^n \int_0^t q_{k,i} q_{\ell,j}\, d[M_k, M_\ell] = \delta(i-j) \int_0^t \lambda_i(r)\, dr \quad \text{for } t \in [0,T], \tag{19.20}$$

for $i,j=1,\ldots,n$. For the $\mathbb{R}^n$-valued process with components

$$\{\hat{W}_i(t)\}_{t\in[0,T]} = \left\{ \int_0^t I_{\{\lambda_i(r)>0\}} \frac{1}{\sqrt{\lambda_i(r)}}\, dN_i(r) + \int_0^t I_{\{\lambda_i(r)=0\}}\, dB_i(r) \right\}_{t\in[0,T]} \tag{19.21}$$

for $i=1,\ldots,n$, Paul Lévy's characterization of BM Theorem 19.16 together with (19.20) shows that $\hat{W}$ is an $\mathbb{R}^n$-valued BM, see Exercise 225 below. Further, we have

$$\left\{ \int_0^t \sqrt{\lambda_i}\, d\hat{W}_i \right\}_{t\in[0,T]} = \{N_i(t)\}_{t\in[0,T]} \quad \text{for } i=1,\ldots,n. \tag{19.22}$$

Now note that as $q_{k,i}^2 \le 1$ for $k,i=1,\ldots,n$ it follows from (19.20) that

$$\int_0^T q(t)_{i,k}^2 \lambda_k(t)\, dt \le \int_0^T \lambda_k(t)\, dt < \infty,$$

so that $x_{i,k} \equiv q_{i,k}\sqrt{\lambda_k} \in P_T$ for $k,i=1,\ldots,n$. Hence (19.19) and (19.22) give

$$\sum_{k=1}^n \int_0^t x_{i,k}\, d\hat{W}_k = \sum_{k=1}^n \int_0^t q_{i,k}\, dN_k = \sum_{k,j=1}^n \int_0^t q_{i,k} q_{j,k}\, dM_j = M_i(t) \tag{19.23}$$

for $t \in [0,T]$, for $i=1,\ldots,n$. Further, note that $x = q\,\mathrm{diag}(\sqrt{\lambda_1},\ldots,\sqrt{\lambda_n})$ satisfies

$$x(t)\,x(t)^{\mathrm{tr}} = q(t)\,\mathrm{diag}(\lambda_1(t),\ldots,\lambda_n(t))\,q(t)^{\mathrm{tr}} = a(t,X(t)) = \sigma(t,X(t))\,\sigma(t,X(t))^{\mathrm{tr}} \tag{19.24}$$

for $t \in [0,T]$. It follows that there exists an $\mathbb{R}_{n|n}$-matrix valued measurable and adapted process $\{R(t)\}_{t\in[0,T]}$ such that $R(t)R(t)^{\mathrm{tr}} = I$ and $\sigma(t,X(t)) = x(t)R(t)$ for $t \in [0,T]$, see Exercise 226 below. Hence the following process

$$\{W(t)\}_{t\in[0,T]} = \left\{ \int_0^t R^{\mathrm{tr}}\, d\hat{W} \right\}_{t\in[0,T]} \tag{19.25}$$

is well-defined and is an $\mathbb{R}^n$-valued BM by Paul Lévy's characterization of BM Theorem 19.16, see Exercise 227 below. Recalling (19.23) we now get (19.17) as

$$M(t) = \int_0^t x\, d\hat{W} = \int_0^t xRR^{\mathrm{tr}}\, d\hat{W} = \int_0^t \sigma(r,X(r))\, dW(r) \quad \text{for } t \in [0,T]. \qquad \square$$

Exercise 222. Show that the fact that the process (19.15) is zero implies (19.16).

Exercise 223. Show that the diffusion matrix $a(t,x)$ is symmetric and non-negative definite for each $(t,x)\in[0,T]\times\mathbb{R}^n$.

Exercise 224. Show that the processes q and $\lambda_1,\ldots,\lambda_n$ in (19.18) are measurable and adapted.

Exercise 225. Show that the process $\hat{W}$ in (19.21) is an $\mathbb{R}^n$-valued BM.

Exercise 226. Show that (19.24) implies existence of an $\mathbb{R}_{n|n}$-matrix valued measurable and adapted process $\{R(t)\}_{t\in[0,T]}$ such that $RR^{\mathrm{tr}}=I$ and $\sigma(\,\cdot\,,X)=xR$.

Exercise 227. Show that the process W in (19.25) is an $\mathbb{R}^n$-valued BM.

Exercise 228. With the notation and setting of the proof of Theorem 19.20, explain why it gives no loss of generality to prove (19.17) when $n=d$ only[28].

Exercise 229. Show how the implication to the right in Theorem 19.20 follows from the Itô formula Theorem 19.5.

Corollary 19.21. *If $\{X(t)\}_{t\in[0,T]}$ is a solution to the SDE* (19.2) *with the generator $\mathcal{A}_t$ in* (19.12) *where the coefficient σ is locally bounded, then for each $f\in C_0^2(\mathbb{R}^n)$ the following stochastic process is a continuous martingale*

$$\left\{f(X(t))-f(X(0))-\int_0^t(\mathcal{A}_r f)(X(r))\,dr\right\}_{t\in[0,T]}. \tag{19.26}$$

Exercise 230. Prove Corollary 19.21.

Definition 19.22. *A continuous and adapted $\mathbb{R}^n$-valued stochastic process $\{X(t)\}_{t\in[0,T]}$ is a solution to the* martingale problem *associated with the generator $\mathcal{A}_t$ in* (19.12) *if for each $f\in C_0^2(\mathbb{R}^n)$ the process* (19.26) *is a continuous martingale.*

Theorem 19.23. *A continuous and adapted stochastic process $\{X(t)\}_{t\in[0,T]}$ is a weak solution to the SDE* (19.2) *with a locally bounded σ coefficient if and only if $X(0)=_D X_0$ and X is a solution to the martingale problem associated with the generator $\mathcal{A}_t$ in* (19.12).

[28] See also Karatzas and Shreve: "Brownian Motion and Stochastic Calculus", pages 170-172 and 316-317. In particular note the paragraph "It suffices ..." at the lower part of page 316.

Proof. The implication to the right is Corollary 19.21. For the implication to the left, let X be a continuous and adapted process that solves the martingale problem associated with $\mathcal{A}_t$. By Theorem 19.20 it is sufficient to show that X solves the local martingale problem associated with $\mathcal{A}_t$. To that end, given an $f \in C^2(\mathbb{R}^n)$ and an $k \in \mathbb{N}$, pick an $f_k \in C_0^2(\mathbb{R}^n)$ that agree with f on the ball $\{x \in \mathbb{R}^n : \|x\| \leq k\}$. Then

$$\{M_k(t)\}_{t\in[0,T]} = \left\{ f_k(X(t)) - f_k(X(0)) - \int_0^t (\mathcal{A}_r f_k)(X(r))\, dr \right\}_{t\in[0,T]}$$

is a continuous martingale. Define a stopping time $\tau_k = \inf\{t \geq 0 : \|X(t)\| \geq k\}$. By the continuity of X we have $\tau_k \uparrow \infty$ as $k \to \infty$. Hence it is sufficient to show that $\{M(t\wedge\tau_k)\}_{t\in[0,T]}$ is a martingale for each $k \in \mathbb{N}$, where M is the process in (19.26). However, that is done by exactly the same computation as that in the proof of Theorem 17.3, which shows that $\{M(t\wedge\tau_k)\}_{t\in[0,T]} = \{M_k(t\wedge\tau_k)\}_{t\in[0,T]}$, where the latter process is a martingale by the optional stopping theorem. □

19.7 Weak existence

Lemma 19.24.[29] *Let $\{{}_1X(t)\}_{t\in[0,T]}$, $\{{}_2X(t)\}_{t\in[0,T]}$, ... be continuous $\mathbb{R}^n$-valued stochastic processes such that*

$$\lim_{\lambda\to\infty} \sup_{k\geq 1} \mathbf{P}\{\|{}_kX(0)\| > \lambda\} = 0, \tag{19.27}$$

and such that there exist constants $C, \alpha, \beta > 0$ such that

$$\mathbf{E}\{\|{}_kX(t) - {}_kX(s)\|^\alpha\} < C\,|t-s|^{1+\beta} \quad \text{for } s,t\in[0,T] \text{ and } k\in\mathbb{N}. \tag{19.28}$$

There exist a continuous $\mathbb{R}^n$-valued stochastic process $\{X(t)\}_{t\in[0,T]}$ and a sequence $\{k_j\}_{j=1}^\infty \subseteq \mathbb{N}$ with $k_j \uparrow \infty$ as $j\to\infty$ such that for each bounded continuous function $F : C([0,T])^n \to \mathbb{R}$ we have $F({}_{k_j}X) \to F(X)$ as $j\to\infty$ with convergence in distribution.

Theorem 19.25. (Stroock-Varadhan) *Consider the generator $\mathcal{A}_t$ in (19.12) where the coefficients μ and σ are bounded and continuous. For each random variable X_0 the martingale problem associated with $\mathcal{A}_t$ has a solution $\{X(t)\}_{t\in[0,T]}$ such that $X(0) =_D X_0$.*

Proof. Given a $k\in\mathbb{N}$, define a process $\{{}_kX(t)\}_{t\in[0,T]}$ recursively by ${}_kX(0) = X_0$ and

$${}_kX(t) = {}_kX(\tfrac{\ell}{k}) + \mu(\tfrac{\ell}{k}, {}_kX(\tfrac{\ell}{k}))\,(t-\tfrac{\ell}{k}) + \sigma(\tfrac{\ell}{k}, {}_kX(\tfrac{\ell}{k}))\,(B(t)-B(\tfrac{\ell}{k})) \quad \text{for } t\in(\tfrac{\ell}{k}, \tfrac{\ell+1}{k}\wedge T],$$

for $\ell = 1, 2, \ldots, \lfloor kT \rfloor$. Notice that, writing $\mu_k(0) = \sigma_k(0) = 0$ and

$$\mu_k(t) = \mu(\tfrac{\ell}{k}, {}_kX(\tfrac{\ell}{k})) \quad\text{and}\quad \sigma_k(t) = \sigma(\tfrac{\ell}{k}, {}_kX(\tfrac{\ell}{k})) \quad \text{for } t\in(\tfrac{\ell}{k}, \tfrac{\ell+1}{k}\wedge T]$$

[29]See e.g., Karatzas and Shreve: "Brownian Motion and Stochastic Calculus", Section 2.4.B.

for $\ell=1,2,\ldots,\lfloor kT\rfloor$, the process ${}_kX$ solves the non-diffusion type SDE

$$ {}_kX(t) = {}_kX(0) + \int_0^t \mu_k(r)\,dr + \int_0^t \sigma_k\,dB \quad \text{for } t\in[0,T], \quad {}_kX(0) = X_0. \tag{19.29}$$

In order to apply Lemma 19.24 to the processes $\{{}_1X(t)\}_{t\in[0,T]}$, $\{{}_2X(t)\}_{t\in[0,T]}$, ... we note that they are continuous martingales by (19.29) and that (19.27) holds since ${}_kX(0)=X_0$ for $k\in\mathbb{N}$. Further, (19.28) holds with $\alpha=4$ and $\beta=1$ since (19.29) together with Exercise 231 below and the Burkholder-Davis-Gundy inequality for $p=4$ give (see also Exercise 195)

$$\begin{aligned}
&\mathbf{E}\{\|{}_kX(t)-{}_kX(s)\|^4\}\\
&= \mathbf{E}\bigg\{\bigg\|\int_s^t \mu_k(r)\,dr + \int_s^t \sigma_k\,dB\bigg\|^4\bigg\}\\
&\le 8\,\mathbf{E}\bigg\{\bigg\|\int_s^t \mu_k(r)\,dr\bigg\|^4\bigg\} + 8\,\mathbf{E}\bigg\{\bigg\|\int_s^t \sigma_k\,dB\bigg\|^4\bigg\}\\
&\le 8\,\mathbf{E}\bigg\{\bigg(\int_s^t 1\,dr\bigg)^2\bigg(\int_s^t \|\mu_k(r)\|^2\,dr\bigg)^2\bigg\} + 8\,n\,d^3\sum_{i=1}^n\sum_{j=1}^n C(4)\,\mathbf{E}\bigg\{\bigg(\int_s^t (\sigma_k)_{i,j}(r)^2\,dr\bigg)^2\bigg\}\\
&\le 8\bigg(T^2 \sup_{(r,x)\in[0,T]\times\mathbb{R}} \|\mu(r,x)\|^4 + n\,d^3 C(4)\sum_{i=1}^n\sum_{j=1}^n \sup_{(r,x)\in[0,T]\times\mathbb{R}} \sigma_{i,j}(r,x)^4\bigg)(t-s)^2
\end{aligned}$$

for $s,t\in[0,T]$. Hence Lemma 19.24 shows that there exists a continuous process $\{X(t)\}_{t\in[0,T]}$ and a sequence of integers $\{k_j\}_{j=1}^\infty$ with $k_j\uparrow\infty$ as $j\to\infty$ such that

$$\bigg(f({}_{k_j}X(t)) - f({}_{k_j}X(s)) - \int_s^t (\mathcal{A}_r f)({}_{k_j}X(r))\,dr\bigg)\,g\big(\{{}_{k_j}X(r)\}_{r\in[0,s]}\big) \tag{19.30}$$

converges in distribution as $j\to\infty$ to

$$\bigg(f(X(t)) - f(X(s)) - \int_s^t (\mathcal{A}_r f)(X(r))\,dr\bigg)\,g\big(\{X(r)\}_{r\in[0,s]}\big) \tag{19.31}$$

for $f\in C_0^2(\mathbb{R}^n)$, $0\le s\le t\le T$ and any bounded continuous function $g: C([0,s])^n\to\mathbb{R}$. In fact, by continuity and boundedness of μ and σ also

$$\bigg(f({}_{k_j}X(t)) - f({}_{k_j}X(s)) - \int_s^t ({}_{k_j}\mathcal{A}_r f)({}_{k_j}X(r))\,dr\bigg)\,g\big(\{{}_{k_j}X(r)\}_{r\in[0,s]}\big) \tag{19.32}$$

converges in distribution to the limit (19.31) for $f\in C_0^2(\mathbb{R}^n)$, $0\le s\le t\le T$ and any bounded continuous $g: C([0,s])^n\to\mathbb{R}$, where

$$({}_k\mathcal{A}_t f)(x) = \sum_{i=1}^n \mu_k(t)_i\, f'_{x_i}(x) + \sum_{i,j=1}^n\bigg(\sum_{\ell=1}^d \frac{\sigma_k(t)_{i,\ell}\,\sigma_k(t)_{j,\ell}}{2}\bigg) f''_{x_i x_j}(x) \quad \text{for } f\in C_0^2(\mathbb{R}^n)$$

is the generator of the SDE (19.29), see Exercise 232 below.

It is sufficient to prove that

$$\bigg\{f(X(t)) - f(X(0)) - \int_0^t (\mathcal{A}_r f)(X(r))\,dr\bigg\}_{t\in[0,T]}$$

is a martingale with respect to the filtration $\{\mathcal{F}_t^X\}_{t\in[0,T]}$ generated by X for $f \in C_0^2(\mathbb{R}^n)$. By analogy with the proof of Theorem 17.5, this in turn is so if

$$\mathbf{E}\bigg\{\bigg(f(X(t)) - f(X(s)) - \int_s^t (\mathcal{A}_r f)(X(r))\,dr\bigg) g\big(\{X(r)\}_{r\in[0,s]}\big)\bigg\} = 0 \qquad (19.33)$$

for any bounded continuous $g : C([0,s])^n \to \mathbb{R}$, for $0 \le s \le t \le T$ and $f \in C_0^2(\mathbb{R}^n)$. However, as (19.29) implies that

$$\bigg\{f({}_kX(t)) - f({}_kX(0)) - \int_0^t ({}_k\mathcal{A}_r f)({}_kX(r))\,dr\bigg\}_{t\in[0,T]}$$

is a martingale with respect to the filtration generated by B for $f \in C_0^2(\mathbb{R}^n)$, recall Exercise 198, and as ${}_kX$ also is adapted to that filtration, it follows that

$$\mathbf{E}\bigg\{\bigg(f({}_kX(t)) - f({}_kX(s)) - \int_s^t ({}_k\mathcal{A}_r f)({}_kX(r))\,dr\bigg) g\big(\{{}_kX(r)\}_{r\in[0,s]}\big)\bigg\} = 0 \qquad (19.34)$$

for any bounded continuous $g : C([0,s])^n \to \mathbb{R}$, for $0 \le s \le t \le T$ and $f \in C_0^2(\mathbb{R}^n)$. As the sequence of random variables indexed by j in (19.32) is uniformly integrable and converges weakly to the random variable in (19.31), (19.34) implies (19.33). □

Exercise 231. Show that $(\sum_{i=1}^n (\sum_{j=1}^d x_{i,j})^2)^2 \le n d^3 \sum_{i=1}^n \sum_{j=1}^d x_{i,j}^4$.

Exercise 232. Show that the random variable (19.32) converges in distribution to the random variable (19.31) when (19.30) does.

Corollary 19.26. (SKOROHOD) *If the SDE* (19.2) *has bounded and continuous coefficients μ and σ, then the SDE has a weak solution for each initial value X_0.*

Exercise 233. Prove Corollary 19.26.

19.8 Weak uniqueness

Theorem 19.27. (STROOCK-VARADHAN) *Consider the generator $\mathcal{A}_t$ in* (19.12) *which is supposed to have a locally bounded σ coefficient. Assume that given any $f \in C_0^\infty(\mathbb{R}^n)$, $s \in [0,T)$ and $t \in (0, T-s]$ the so called* Cauchy problem

$$\frac{\partial g(r,x)}{\partial r} + (\mathcal{A}_{r+s} g)(x) = 0 \quad \text{for } (r,x) \in [0,t]\times\mathbb{R}^n, \quad g(t,\cdot) = f, \qquad (19.35)$$

has a solution $g \in C_B([0,t]\times\mathbb{R}^n) \cap C^{1,2}([0,t]\times\mathbb{R}^n)$. Given an $\mathbb{R}^n$-valued random variable X_0 a solution $\{X(t)\}_{t\in[0,T]}$ to the martingale problem associated with $\mathcal{A}_t$ such that $X(0) =_D X_0$ has uniquely determined fidi's.

Exercise 234. Prove Theorem 19.27.

Remark 19.28. The Cauchy problem (19.35) has a solution if, for example[30], the coefficients μ and σ for generator $\mathcal{A}_t$ in (19.12) are bounded and satisfy a global Hölder condition with a *strongly elliptic* diffusion matrix a, that is,

$$\xi^{\text{tr}} a\, \xi \geq K\, \|\xi\|^2 \quad \text{for} \;\; (t,x) \in [0,T] \times \mathbb{R}^n \;\; \text{and} \;\; \xi \in \mathbb{R}^n, \;\; \text{for some constant} \;\; K > 0.$$

Corollary 19.29. *Under the hypothesis of Theorem 19.27 the SDE* (19.2) *displays uniqueness for weak solutions.*

Exercise 235. Prove Corollary 19.29.

[30] See e.g., Stroock and Varadhan: "Multidimensional Diffusion Processes", Theorem 3.2.1.