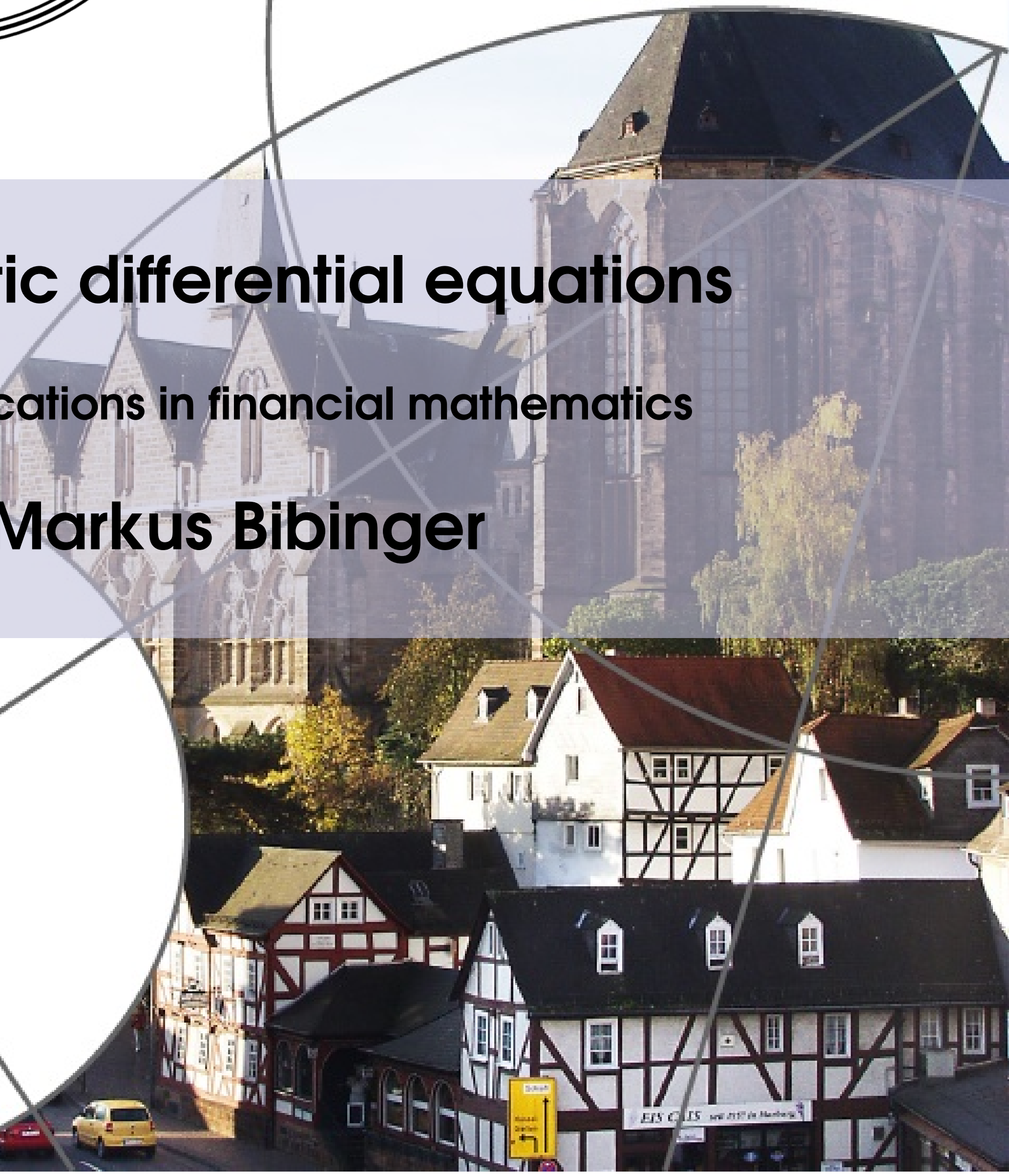
Stochastic differential equations

with applications in financial mathematics

Markus Bibinger



Philipps-Universität Marburg
Fachbereich 12 Mathematik und Informatik
Stochastik, Hans-Meerwein-Straße 6, D-35032 Marburg
Tel. +49 6421/28-25471, Fax +49 6421/28-25469
bibinger@mathematik.uni-marburg.de
https://www.uni-marburg.de/fb12/stoch/personal/bibinger

This version: February 9, 2017

Philipps Universität Marburg

Contents

I Stochastic integration

1. White Noise

Many processes in nature involve random fluctuations which we have to account for in our models. In principle, everything can be random and the probabilistic structure of these random influences can be arbitrarily complicated. As it turns out, the so called "white noise" plays an outstanding role.

Engineers want the white noise process $(\dot{W}(t), t \in \mathbb{R}_+)$ to have the following properties:

- The random variables $\{\dot{W}(t) \,|\, t \in \mathbb{R}_+\}$ are independent.
- $\dot{W}$ is stationary, that is, the distribution of $(\dot{W}(t+t_1), \dot{W}(t+t_2), \dots, \dot{W}(t+t_n))$ does not depend on t.
- The expectation $\mathbb{E}[\dot{W}(t)]$ is zero.

Hence, this process is supposed to model independent and identically distributed shocks with zero mean. Unfortunately, mathematicians can prove that such a real-valued stochastic process cannot have measurable trajectories $t \mapsto \dot{W}(t)$ except for the trivial process $\dot{W}(t) = 0$.

Problem 1.1 If $(t, \omega) \mapsto \dot{W}(t, \omega)$ is jointly measurable with $\mathbb{E}[\dot{W}(t)^2] < \infty$ and $\dot{W}$ has the above stated properties, then for all $t \geq 0$

$$\mathbb{E}\Big[\Big(\int_0^t \dot{W}(s)\,ds\Big)^2\Big] = 0$$

holds and $\dot{W}(t) = 0$ almost surely. Can we relax the hypothesis $\mathbb{E}[\dot{W}(t)^2] < \infty$?

Nevertheless, applications forced people to consider equations like

$$\dot{x}(t) = \alpha x(t) + \dot{W}(t), \quad t \geq 0.$$

The way out of this dilemma is found by looking at the corresponding integrated equation:

$$x(t) = x(0) + \int_0^t \alpha x(s)\,ds + \int_0^t \dot{W}(s)\,ds, \quad t \geq 0.$$

What properties should we thus require for the integral process $W(t) := \int_0^t \dot{W}(s)ds$, $t \geq 0$? A straight-forward deduction (from wrong premises) yields

- $W(0) = 0$.
- The increments $(W(t_1), W(t_2) - W(t_1), W(t_3) - W(t_2), \dots, W(t_n) - W(t_{n-1}))$ are independent for $0 < t_1 \leq t_2 \leq \cdots \leq t_n < \infty$.
- The increments are stationary, that is, $W(t+s) - W(t) \stackrel{\mathcal{L}}{=} W(s)$ holds for all $t, s \geq 0$.
- The expectation $\mathbb{E}[W(t)]$ is zero.
- The trajectories $t \mapsto W(t)$ are continuous.

The last point is due to the fact that integrals over measurable (and integrable) functions are always continuous. It is highly nontrivial to show that – up to indistinguishability and up to the norming $\mathrm{Var}[W(1)] = 1$ – the only stochastic process fulfilling these properties is Brownian motion (also known as Wiener process), see Øksendal [1998]. Recall that Brownian motion is almost surely nowhere differentiable!

Rephrasing the stochastic differential equation, we now look for a stochastic process $(X(t), t \geq 0)$ satisfying

$$X(t) = X(0) + \int_0^t \alpha X(s) ds + W(t), \quad t \geq 0, \tag{1.0.1}$$

where $(W(t), t \geq 0)$ is a standard Brownian motion. The precise formulation involving filtrations will be given later, here we shall focus on finding processes X solving (1.0.1).

The so-called variation of constants approach in ODEs would suggest the solution

$$X(t) = X(0)e^{\alpha t} + \int_0^t e^{\alpha(t-s)} \dot{W}(s)\, ds, \tag{1.0.2}$$

which we give a sense (in fact, that was Wiener's idea) by partial integration:

$$X(t) = X(0)e^{\alpha t} + W(t) + \int_0^t \alpha e^{\alpha(t-s)} W(s)\, ds. \tag{1.0.3}$$

This makes perfect sense now since Brownian motion is (almost surely) continuous and we could even take the Riemann integral. The verification that (1.0.3) defines a solution is straight forward:

$$\begin{aligned}
\int_0^t \alpha X(s)\, ds &= X(0) \int_0^t \alpha e^{\alpha s}\, ds + \alpha \int_0^t W(s)\, ds + \alpha^2 \int_0^t \int_0^s e^{\alpha(s-u)} W(u)\, du\, ds \\
&= X(0)(e^{\alpha t} - 1) + \alpha \int_0^t W(s)\, ds + \alpha^2 \int_0^t W(u) \int_u^t e^{\alpha(s-u)}\, ds\, du \\
&= X(0)(e^{\alpha t} - 1) + \int_0^t \alpha W(u) e^{\alpha(t-u)}\, du \\
&= X(t) - X(0) - W(t).
\end{aligned}$$

Note that the initial value $X(0)$ can be chosen arbitrarily. The expectation $\mu(t) := \mathbb{E}[X(t)] = \mathbb{E}[X(0)]e^{\alpha t}$ exists if $X(0)$ is integrable. Surprisingly this expectation function satisfies the deterministic linear equation, hence it converges to zero for $\alpha < 0$ and explodes for $\alpha > 0$. How about the variation around this mean value? Let us suppose that $X(0)$ is deterministic, $\alpha \neq 0$ and consider the variance function

$$
\begin{aligned}
v(t) &:= \operatorname{Var}[X(t)] = \mathbb{E}\Big[\Big(W(t) + \int_0^t \alpha e^{\alpha(t-s)} W(s)\,ds\Big)^2\Big] \\
&= \mathbb{E}[W(t)^2] + 2\int_0^t \alpha e^{\alpha(t-s)}\,\mathbb{E}[W(t)W(s)]\,ds + \int_0^t\int_0^t \alpha^2 e^{\alpha(2t-u-s)}\,\mathbb{E}[W(s)W(u)]\,du\,ds \\
&= t + 2\int_0^t \alpha e^{\alpha(t-s)} s\,ds + 2\int_0^t\int_s^t \alpha^2 e^{\alpha(2t-u-s)} s\,du\,ds \\
&= t + \int_0^t \Big(2\alpha e^{\alpha(t-s)} s + 2\alpha(e^{2\alpha(t-s)} - e^{\alpha(t-s)})s\Big)\,ds \\
&= \tfrac{1}{2\alpha}\big(e^{2\alpha t} - 1\big).
\end{aligned}
$$

This shows that for $\alpha < 0$ the variance converges to $\frac{1}{2|\alpha|}$ indicating a stationary behaviour, which will be made precise in the sequel. On the other hand, for $\alpha > 0$ we find that the standard deviation $\sqrt{v(t)}$ grows with the same order as $\mu(t)$ for $t \to \infty$ which lets us expect a very erratic behaviour.

In anticipation of the Itô calculus, the preceding calculation can be simplified by regarding (1.0.2) directly. The second moment of $\int_0^t e^{\alpha(t-s)}\,dW(s)$ is immediately seen to be $\int_0^t e^{2\alpha(t-s)}\,ds$, the above value.

2. The Itô integral

2.1 Preliminaries on stochastic processes

From now on we shall always be working on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where a *filtration* $(\mathcal{F}_t)_{t\geq 0}$, that is, a nested family of σ-fields $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $s \leq t$, is defined that satisfies the *usual conditions*:

- $\mathcal{F}_s = \bigcap_{t>s} \mathcal{F}_t$ for all $s \geq 0$ (right-continuity);
- all $A \in \mathcal{F}$ with $\mathbb{P}(A) = 0$ are contained in $\mathcal{F}_0$.

A family $(X(t), t \geq 0)$ of $\mathbb{R}^d$-valued random variables on our probability space is called a *stochastic process* and this process is $(\mathcal{F}_t)$-*adapted* if all $X(t)$ are $\mathcal{F}_t$-measurable. Denoting the Borel σ-field on $[0,\infty)$ by $\mathcal{B}$, this process X is *measurable* if $(t,\omega) \mapsto X(t,\omega)$ is a $\mathcal{B} \otimes \mathcal{F}$-measurable mapping. We say that $(X(t), t \geq 0)$ is *continuous* if the *trajectories* $t \mapsto X(t,\omega)$ are continuous for all $\omega \in \Omega$. One can show that a process is measurable if it is (right-)continuous [Karatzas and Shreve, 1991, Thm. 1.14].

Two processes $X(t)$ and $Y(t)$ are *modifications* of each other if $\mathbb{P}(X_t = Y_t) = 1$ for all $t \geq 0$.
A stochastic process $(X(t), t \geq 0)$ is called càdlàg if the trajectories are right-continuous and left limits $X(t-) = \lim_{s\uparrow t} X(s,\omega)$ exist. We denote with $\Delta X(t) = X(t) - X(t-)$.
Analogously, $(X(t), t \geq 0)$ is called càglàd if the trajectories are left-continuous and right limits $X(t+) = \lim_{s\downarrow t} X(s,\omega)$ exist.

Definition 2.1.1 A (standard one-dimensional) *Brownian motion* with respect to the filtration $(\mathcal{F}_t)$ is a continuous $(\mathcal{F}_t)$-adapted real-valued process $(W(t), t \geq 0)$ such that

- W(0)=0;
- for all $0 \leq s \leq t$: $W(t) - W(s)$ is independent of $\mathcal{F}_s$;
- for all $0 \leq s \leq t$: $W(t) - W(s)$ is $\mathcal{N}(0, t-s)$-distributed.

R Brownian motion can be constructed in different ways, see Karatzas and Shreve [1991], but the proof of the existence of such a process is in any case non-trivial.

We shall often consider a larger filtration $(\mathcal{F}_t)$ than the canonical filtration $(\mathcal{F}_t^W)$ of Brownian motion in order to include random initial conditions. Given a Brownian motion process W' on a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ with the canonical filtration $\mathcal{F}_t' = \sigma(W'(s), s \leq t)$ and the random variable X_0'' on a different space $(\Omega'', \mathcal{F}'', \mathbb{P}'')$, we can construct the product space with $\Omega = \Omega' \times \Omega''$, $\mathcal{F} = \mathcal{F}' \otimes \mathcal{F}''$, $\mathbb{P} = \mathbb{P}' \otimes \mathbb{P}''$ such that $W(t,\omega',\omega'') := W'(t,\omega')$ and $X_0(\omega',\omega'') := X_0''(\omega'')$ are independent and W is an $(\mathcal{F}_t)$-Brownian motion for $\mathcal{F}_t = \sigma(X_0; W(s), s \leq t)$. Note that X_0 is $\mathcal{F}_0$-measurable which always implies that X_0 and W are independent.

Our aim here is to construct the integral $\int_0^t Y(s)\,dW(s)$ with Brownian motion as integrator, and even more general integrators, and a fairly general class of stochastic integrands Y.

Definition 2.1.2 Let V be the class of real-valued stochastic processes $(Y(t), t \geq 0)$ that are adapted, measurable and that satisfy

$$\|Y\|_V := \Big(\int_0^\infty \mathbb{E}\big[Y(t)^2\big]\,dt\Big)^{1/2} < \infty.$$

Let V^* be the class of real-valued stochastic processes $(Y(t), t \geq 0)$ that are adapted, measurable and that satisfy

$$\mathbb{P}\Big(\int_0^\infty Y(t)^2\,dt < \infty\Big) = 1.$$

2.2 The Stieltjes integral

In the lecture on measure and integration theory (stochastics I) we have learned about the one-to-one correspondence between right-continuous functions $F_\mu : \mathbb{R} \to \mathbb{R}$ of bounded variation and finite signed measures via the relation $\mu((a,b]) = F_\mu(b) - F_\mu(a)$. The integral $\int f\,d\mu$ in this case is called the (Lebesgue-) Stieltjes integral $\int f\,dF_\mu$. Denote with $\mathcal{V}$ the set of adapted càdlàg processes of bounded variation. A process on a compact time interval $[0,T]$ is of *bounded variation*, if

$$\sup_\pi \sum_i \big|X(t_{i+1}) - X(t_i)\big| < \infty, \text{ a.s.},$$

for partitions $\pi = \{t_1, \ldots, t_n | t_1 < \ldots < t_n\}$ of $[0,T]$. A process $(X(t))_{t\in\mathbb{R}_+}$ is of finite variation if it is of bounded variation over any compact interval.
For $X(t,\omega) \in \mathcal{V}$ define pathwise for a measurable bounded process H:

$$\int_0^T H(s,\omega)\,dX(s,\omega) = \int_0^T H(s,\omega)\,\mu_X(\omega, ds)$$

with $\mu_X(\omega,[0,t]) = X(t,\omega)$ the associated Lebesgue-Stieltjes measure. If $s \mapsto H(s,\omega)$ is continuous, the above integral exists as Riemann-Stieltjes integral and

$$\sum_i H(\tau_i,\omega)\big(X(t_i,\omega) - X(t_{i-1},\omega)\big) \to \int_0^T H(s,\omega)\,dX(s,\omega) \tag{2.2.1}$$

as $\sup_i |t_i - t_{i-1}| \to 0$ for any $t_{i-1} \leq \tau_i \leq t_i$. The Riemann-Stieltjes integral as the above limit exists for one continuous process and one process of bounded variation. We have the partial integration formula $\int_0^T f\,dF = fF\big|_0^T - \int_0^T F\,df$. Thereby, we obtain with a Wiener process W, for instance,

$$\int_0^t e^t\,dW(t,\omega) = e^t W(t,\omega)\big|_0^T - \int_0^T W(t,\omega)\,de^t = e^T W(T,\omega) - \int_0^T W(t,\omega)e^t\,dt\,.$$

Problem 2.1 The pathwise integral can not be extended to integrators with infinite variation.

■ **Example 2.1** Consider as an example $\int_0^T W(t)\,dW(t)$ with a Wiener process W. Different discrete approximations are

$$\underline{W}(t,\omega) = \sum_{j\geq 0} W(j/2^n,\omega)\mathbf{1}_{[\mathbf{j/2^n,(j+1)/2^n}]}(\mathbf{t}), \tag{2.2.2a}$$

$$\overline{W}(t,\omega) = \sum_{j\geq 0} W((j+1)/2^n,\omega)\mathbf{1}_{[\mathbf{j/2^n,(j+1)/2^n}]}(\mathbf{t}). \tag{2.2.2b}$$

We derive for the two different Riemann-Stieltjes sum approximations

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{j\geq 0} W(j/2^n)\big(W((j+1)/2^n)-W(j/2^n)\big)\Big] \\
&\quad = \sum_{j\geq 0}\mathbb{E}\left[W(j/2^n)\big(W((j+1)/2^n)-W(j/2^n)\big)\right] = 0, \\
&\mathbb{E}\Big[\sum_{j\geq 0} W((j+1)/2^n)\big(W((j+1)/2^n)-W(j/2^n)\big)\Big] \\
&\quad = \sum_{j\geq 0}\mathbb{E}\left[\big(W((j+1)/2^n)-W(j/2^n)\big)^2\right] = T.
\end{aligned}$$

Further, one can show that both variances tend to zero. Thus, we conclude that the natural approximation of the integral by Riemann-Stieltjes sums of the form $\sum_j W(\tau_j)(W(t_j) - W(t_{j-1})$ are, contrarily to the Riemann-Stieltjes integral, not independent of the exact choice of the τ_j. ■

2.3 Quadratic variation of Brownian motion

Definition 2.3.1 We say that a real-valued process $(X(t))_{t\in\mathbb{R}_+}$ has *finite quadratic variation*, if a finite process $\big([X,X](t)\big)_{t\in\mathbb{R}_+}$ exists, such that

$$\sum_{i=1}^{n}\big(X(t_i)-X(t_{i-1})\big)^2 \xrightarrow{p} [X,X](t)\ , t\in\mathbb{R}_+, \tag{2.3.1}$$

as $|\Pi|\to 0$ where Π denotes a partition given by real numbers (t_i) with $t_0 = 0 < t_1 < \ldots < t_n = t$, and $|\Pi| = \max_i(t_{i+1}-t_i)$.
We call $[X,X](t)$ the *quadratic variation (process)* of X. The *quadratic covariation* up to time t between two processes X and Y is given by

$$[X,Y](t) = p - \lim_{|\Pi|\to 0}\sum_{t_i\in\Pi}(X(t_{i+1})-X(t_i))(Y(t_{i+1})-Y(t_i)) \quad \forall t\geq 0.$$

Lemma 2.3.2 For a continuous real-valued process X of bounded variation it holds true that $[X,X](t) = 0$.

Proof.

$$\sum_{i=1}^{n}\big(X(t_i)-X(t_{i-1})\big)^2 \leq \sup_{1\leq i\leq n}|X(t_i)-X(t_{i-1})|\sum_{i=1}^{n}|X(t_i)-X(t_{i-1})| \to 0$$

what implies the claim. ■

Proposition 2.3.3 For a one-dimensional standard Brownian motion, we have $[B,B](t) = t$.

Proof. It is sufficient to prove that

$$\sum_{i=1}^{n} \big(B(t_i) - B(t_{i-1})\big)^2 \xrightarrow{L_2} t\,. \tag{2.3.2}$$

We have already seen that

$$\mathbb{E}\Big[\sum_{i=1}^{n} \big(B(t_i) - B(t_{i-1})\big)^2\Big] = t\,.$$

Set $Z_i = \big(B(t_i) - B(t_{i-1})\big)^2 - (t_i - t_{i-1}), i = 1,\dots,n$. Then, we derive that $\mathbb{E}[Z_i] = 0$ for all $i = 1,\dots,n$, and

$$\begin{aligned}\operatorname{Var}\Big(\sum_{i=1}^{n} Z_i\Big) &= \sum_{i=1}^{n} \operatorname{Var}(Z_i) = \sum_{i=1}^{n} 2(t_i - t_{i-1})^2 \\ &\le 2\sup_i |t_i - t_{i-1}| \sum_{i=1}^{n} (t_i - t_{i-1}) \to 0\,.\end{aligned}$$

■

In fact, it is possible to show that the above convergence holds in the stronger almost sure sense, cf. page 29 of Revuz and Yor [1999]. Thus, paths of Brownian motion are almost surely not of bounded variation. Consequently, we can not define integrals like $\int_0^t W(t)\,dW(t)$ in the Riemann-Stieltjes sense.

■ **Example 2.2** Reconsider a Riemann-Stieltjes type approximating sum for $\int_0^T B(t)\,dB(t)$, with a standard Brownian motion B, again. First, we take left interval endpoints and find that

$$\begin{aligned}\sum_{i=1}^{n} B(t_{i-1})\big(B(t_i) - B(t_{i-1})\big) &= \sum_{i=1}^{n} \frac{B^2(t_i) - B^2(t_{i-1}) - \big(B(t_i) - B(t_{i-1})\big)^2}{2} \\ &= \frac{1}{2}\sum_{i=1}^{n} \big(B^2(t_i) - B^2(t_{i-1})\big) - \frac{1}{2}\sum_{i=1}^{n} \big(B(t_i) - B(t_{i-1})\big)^2 \\ &\xrightarrow{P} \frac{1}{2}\big(B_T^2 - T\big)\,.\end{aligned}$$

Using left endpoints was Itô's choice and the limit defines the specific Itô integral

$$\int_0^T B(t)\,dB(t) = \frac{1}{2}\big((B(T))^2 - T\big)\,. \tag{2.3.3}$$

Choosing other intermediate points τ_i than left endpoints renders different results. For $\tau_i = (1-\alpha)t_{i-1} + \alpha t_i, 0 \le i \le n$, one obtains for $0 \le \alpha \le 1$:

$$\sum_{i=1}^{n} B(\tau_i)\big(B(t_i) - B(t_{i-1})\big) \xrightarrow{P} \frac{(B(T))^2}{2} + \Big(\alpha - \frac{1}{2}\Big)T\,. \tag{2.3.4}$$

Hence, each $\alpha \in [0,1]$ induces a different notion of the integral. There are three important and meaningful cases:

- $\alpha = 0$, this renders the *Itô integral*.
- $\alpha = 1/2$, this leads to the *Fisk-Stratonovich* integral, the second important notion of stochastic integrals.
- $\alpha = 1$ corresponds to the *backward (Itô) integral*.

■

Observe that for the Fisk-Stratonovich integral, which we denote by $\int_0^T B(t) \circ dB(t) = B_T^2/2$, the usual calculus well-known from Riemann integration theory applies. However, in this case $\int_0^T B(t) \circ dB(t)$ is not a martingale. Only Itô's choice renders martingales if we perform stochastic integration with martingales as integrators. Especially for financial applications and interpretation, this is a crucial aspect.

2.4 The integral for simple processes

Definition 2.4.1 A stochastic process $(H(t))_{t\in\mathbb{R}_+}$ is called *simple predictable* if it is of the form

$$H(t,\omega) = \sum_{i=1}^{n} H_{i-1}(\omega)\mathbf{1}_{(T_{i-1},T_i]}(t,\omega),$$

with an increasing finite sequence $(T_i)_{0\le i\le n}$ of stopping times and $\mathcal{F}_{T_{i-1}}$-measurable random variables H_{i-1}.
Equipped with the norm $\|H\| = \sup_{t,\omega}|H(t,\omega)|$, we consider the normed space of simple predictable processes $\mathcal{S}$. Let L^0 be the space of real-valued random variables topologized through convergence in probability, e.g., with the metric (cf. Theorem 3.5 of Dudley [1976])

$$\rho(X,Y) = \mathbb{E}\left[\frac{|X-Y|}{1+|X-Y|}\right].$$

For some càdlàg processes X and $H \in \mathcal{S}$, we define *the stochastic integral $I_X(H)$* as the random variable

$$I_X(H) = \sum_{i=1}^{n} H_{i-1}\big(X(T_i) - X(T_{i-1})\big) \in L^0\,. \tag{2.4.1}$$

This is independent from the particular illustration. $I_X : \mathcal{S} \to L^0$ is a linear functional.

2.5 Extension of the integral

Definition 2.5.1 Some adapted càdlàg process X is called a *total semimartingale (or good integrator)* if $I_X : \mathcal{S} \to L^0$ is continuous, that is,

$$\text{for } (H^n)_{n\in\mathbb{N}} \subseteq \mathcal{S}, H \in \mathcal{S}:\; H^n \to H \;\Rightarrow\; I_X(H^n) \xrightarrow{p} I_X(H)\,.$$

The last definition is equivalent to:

$$\text{For } (H^n)_{n\in\mathbb{N}} \subseteq \mathcal{S}, H \in \mathcal{S}:\; \|H^n\| \to 0 \;\Rightarrow\; \lim_{n\to\infty} \mathbb{P}\left(|I_X(H^n)| \ge \varepsilon\right) = 0 \;\forall \varepsilon > 0\,.$$

Definition 2.5.2 A process X is called a semimartingale if the stopped process $(X(s \wedge t)_{s\in\mathbb{R}_+}$ is a total semimartingale for all $t \in \mathbb{R}_+$.

Theorem 2.5.3 — Jacod's theorem. Let $\mathbb{P}, \mathbb{Q}$ be two probability measures with $\mathbb{Q} \ll \mathbb{P}$. Then, any $\mathbb{P}$-semimartingale is a $\mathbb{Q}$-semimartingale. ∎

Proof. Since convergence in probability under $\mathbb{P}$ implies convergence in probability under $\mathbb{Q}$, the theorem readily follows from Definition 2.5.2. ∎

Proposition 2.5.4 Any process $X \in \mathcal{V}$, that is, an adapted càdlàg process of bounded variation, is a semimartingale.

Proof. Denote with $\mathbf{V}(X)_T$ the variation of X on $[0,T]$. It holds that

$$\left|I_X(H)\right| \leq \|H\| \, \mathbf{V}(X)_T \,.$$

If $\|H^n\| \to 0$, it follows that $\left|I_X(H^n)\right| \xrightarrow{p} 0$. ∎

Proposition 2.5.5 Any L^2-martingale with càdlàg paths is a semimartingale.

Proof. From Doob's optional stopping theorem we know that for a right-continuous martingale $(X(t))_{t\in\mathbb{R}_+}$ and $[0,T]$-valued stopping times τ_1, τ_2 with $\tau_1 \leq \tau_2$, $X(\tau_1) = \mathbb{E}\left[X(\tau_2)|\mathcal{F}_{\tau_1}\right]$, almost surely.
Jensen's inequality yields that $X^2(\tau_1) \leq \mathbb{E}\left[X^2(\tau_2)|\mathcal{F}_{\tau_1}\right]$. Let $H \in \mathcal{S}$. By optional stopping and the martingale property, we conclude

$$\begin{aligned}
\mathbb{E}\left[\left(I_X(H)\right)^2\right] &= \mathbb{E}\left[\left(\sum_{i=1}^{n} H_{i-1}\left(X(T_i) - X(T_{i-1})\right)\right)^2\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n} H_{i-1}^2\left(X(T_i) - X(T_{i-1})\right)^2\right] \\
&\leq \|H\|^2\, \mathbb{E}\left[\sum_{i=1}^{n} \left(X(T_i) - X(T_{i-1})\right)^2\right] = \|H\|^2\, \mathbb{E}[X^2(T_n) - X^2(T_0)]\,.
\end{aligned}$$

Therefore, when $\|H^n\| \to 0$, we have that $I_X(H^n) \to 0$ in $L^2(\mathbb{P})$. This implies convergence of $I_X(H^n)$ to zero in probability. ∎

Corollary 2.5.6 The standard Brownian motion is a semimartingale.

Definition 2.5.7 A sequence of processes $(H^n)_{n\in\mathbb{N}}$ is said to converge *uniformly on compacts in probability (ucp)* to the process H, if for any $t > 0$:

$$\sup_{0\leq s\leq t} \left|H^n(s) - H(s)\right| \xrightarrow{p} 0\,.$$

In the following, we write $\mathbb{D}$ and $\mathbb{L}$ for the space of all adapted càdlàg processes and the space of all adapted càglàd processes, respectively. Restricting to bounded processes, we write $\mathbf{b}\mathbb{D}$ and $\mathbf{b}\mathbb{L}$. Note that ucp-convergence is metrizable and we write $\mathbb{D}_{ucp}$ for the normed space.

Theorem 2.5.8 $\mathbb{D}_{ucp}$ is a Banach space.

Proof. Let X^n be a Cauchy sequence with respect to the metric d in $\mathbb{D}_{ucp}$. Then, for sufficiently large n,m, it holds true that

$$\mathbb{P}\big(\sup_{0\le s\le t}\left|X^n(t)-X^m(t)\right|>\varepsilon\big)\le\varepsilon$$

for any given $\varepsilon>0$. We construct a sequence $(n_k)_{k\in\mathbb{N}}$ recursively with $n_0=0$ and n_k the smallest number, such that for all $n,m\ge n_k$:

$$\mathbb{P}\big(\sup_{0\le s\le t}\left|X^n(t)-X^m(t)\right|>2^{-k}\big)\le 2^{-k}.$$

With Borel-Cantelli, we deduce that

$$\mathbb{P}\underbrace{\Big(\sup_{0\le t\le T}\left|X^{n_k+1}(t)-X^{n_k}(t)\right|>2^{-k}\text{ infinitely often}\Big)}_{=A}=0.$$

For $\omega\in A^{\complement}$ the sequence $(X^{n_k})_{k\in\mathbb{N}}$ is a Cauchy sequence with respect to the sup-norm on $[0,T]$. In particular, for any fix $t\in\mathbb{R}_+$, $(X^{n_k}(t))_{k\in\mathbb{N}}$ forms a real Cauchy sequence. Set

$$X(t,\omega)=\begin{cases}\lim_{k\to\infty}X^{n_k}(t,\omega) & ,\omega\in A^{\complement}\\ 0 & ,\omega\in A\end{cases}\quad\forall\, t\in[0,T]\,.$$

X is adapted and càdlàg by uniform convergence. Moreover,

$$\mathbb{P}\big(\sup_{0\le t\le T}\left|X^{n_k}(t)-X(t)\right|\to 0\big)\ge\mathbb{P}(A^{\complement})=1\,.$$

Convergence with respect to d of a subsequence implies for a Cauchy sequence convergence to the same limit, what completes the proof. ∎

Theorem 2.5.9 $\mathcal{S}$ is dense in $\mathbb{L}_{ucp}$, that is, for all $Y\in\mathbb{L}$ there exists a sequence $H^n\subseteq\mathcal{S}$, such that $H^n\xrightarrow{ucp}Y$.

Proof. Let $Y\in\mathbb{L}$ and $R_n=\inf\{t\,||Y(t)|>n\}$. For the sequence of stopping times R_n, define $Y^n(t)=Y(t\wedge R_n)$, such that $Y^n\in\mathbf{b}\mathbb{L}$. The càglàd property ensures that

$$\mathbb{P}\big(\sup_{0\le t\le T}\left|Y^n(t)-Y(t)\right|\ge\varepsilon\big)\ge\mathbb{P}\big(\sup_{0\le t\le T}|Y(s)|\ge n\big)\to 0\,.$$

Thus $\mathbf{b}\mathbb{L}$ is dense in $\mathbb{L}$ and it remains to prove that $\mathcal{S}$ is dense in $\mathbf{b}\mathbb{L}$. Set $Z(t)=\lim_{u\downarrow t}Y(u)$ with $Z\in\mathbb{D}$ and $T_0^\varepsilon=0,\dots,T_{n+1}^\varepsilon=\inf\{t\,|t>T_n^\varepsilon\text{ and }|Z(t)-Z(T_n^\varepsilon)|>\varepsilon\}$. Since $Z\in\mathbb{D}$, T_n^ε is a sequence of stopping times increasing to infinity.

$$Z^\varepsilon=\sum_n Z(T_n^\varepsilon)\mathbf{1}_{[T_n^\varepsilon,T_{n+1}^\varepsilon]}\xrightarrow{ucp}Z\quad\text{as }\varepsilon\to 0\,.$$

Finally, define the simple processes

$$Y^{n,\varepsilon}=Y_0\mathbf{1}_{\{0\}}+\sum_{i=1}^n Z(T_i^\varepsilon)\mathbf{1}_{[T_i^\varepsilon\wedge n,T_{i+1}^\varepsilon\wedge n]}\xrightarrow{ucp}Y\quad\text{as }\varepsilon\to 0,n\to\infty\,.$$

∎

Definition 2.5.10 For $X \in \mathbb{D}$ and $Y \in \mathcal{S}$, let $J_X : \mathcal{S} \to \mathbb{D}$ be the linear functional

$$J_X(Y) = \sum_{i=1}^{n} Y_i\big(X(T_i \wedge \bullet) - X(T_{i-1} \wedge \bullet)\big) \tag{2.5.1}$$

for $Y = \sum_{i=1}^{n} Y_i \mathbf{1}_{(T_{i-1},T_i]}$ with an increasing finite sequence $(T_i)_{0\le i\le n}$ of stopping times and $\mathcal{F}_{T_i}$-measurable random variables Y_i. The image space is now $\mathbb{D}$ and the process $((J_X(Y))_t)_{t\in\mathbb{R}_+}$ is called *the stochastic integral* of Y with respect to X. A common short notation is $Y \bullet X$.

Theorem 2.5.11 For a semimartingale X, the mapping $J_X : \mathcal{S} \to \mathbb{D}_{ucp}$ is continuous, that is,

$$H^n \xrightarrow{ucp} H \;\Rightarrow J_X(H^n) \xrightarrow{ucp} J_X(H)\,.$$

Proof. By linearity we can without loss of generality consider the case $H = 0$.
Step 1 For $\|H^n\| \to 0$ and H^n uniformly bounded $J_X(H^n) \to 0$ holds.
Let X be a total semimartingale and $\delta > 0$ fix. We construct a sequence of stopping times

$$T_n = \inf\big\{t \ge 0 \mid |(J_X(H^n))_t > \delta\big\}\,.$$

Then, $H^n \mathbf{1}_{[0,T_n]} \in \mathcal{S}$ with $\|H^n \mathbf{1}_{[0,T_n]}\| \to 0$. By right-continuity of $(J_X(H^n))_t$:

$$\begin{aligned}\mathbb{P}\Big(\sup_{0\le s\le t} |(J_X(H^n))_s| > \delta\Big) &\le \mathbb{P}\Big(|(J_X(H^n))_{T_n\wedge t}| > \delta\Big)\\ &= \mathbb{P}\Big(|(J_X(H^n\mathbf{1}_{[0,T_n]}))_t| > \delta\Big)\\ &= \mathbb{P}\Big(|I_X(H^n\mathbf{1}_{[0,T_n\wedge t]})| > \delta\Big) \to 0\,,\end{aligned}$$

such that $J_X(H^n) \to 0$ follows.
Step 2 For $\|H^n\| \xrightarrow{ucp} 0$, $J_X(H^n) \xrightarrow{ucp} 0$ holds.
Let $\|H^n\|_{ucp} \to 0$ and $\delta > 0, \eta > 0, t > 0$, $R_n = \inf\{s \mid |H^n(s)| > \eta\}$, $\tilde{H}^n = H^n \mathbf{1}_{[0,R_n]}$. Hence, $\tilde{H}^n \in \mathcal{S}$ and $\|\tilde{H}^n\| \le \eta$ by left-continuity. If $R_n \ge t$ we have that $\sup_{0\le s\le t} |(J_X(H^n))_s| = \sup_{0\le s\le t} |(J_X(\tilde{H}^n))_s|$. We conclude that

$$\begin{aligned}\mathbb{P}\Big(\sup_{0\le s\le t} |(J_X(H^n))_s| > \delta\Big) &\le \mathbb{P}\Big(\sup_{0\le s\le t} |(J_X(\tilde{H}^n))_s| > \delta\Big) + \mathbb{P}(R_n < t)\\ &\mathbb{P}\Big(\sup_{0\le s\le t} |(J_X(\tilde{H}^n))_s| > \delta\Big) + \mathbb{P}\Big(\sup_{0\le s\le t} |H^n(s)| > \eta\Big)\,.\end{aligned}$$

By step 1, for any $\varepsilon > 0$, we can choose η small enough to bound the probability in the first summand from above by $\varepsilon/2$. For large enough n the second probability can be bounded from above by $\varepsilon/2$ what yields the result. ■

Since $\mathcal{S}_{ucp}$ is dense in $\mathbb{L}_{ucp}$ and $\mathbb{D}_{ucp}$ is a Banach space the result above that $J_X : \mathcal{S}_{ucp} \to \mathbb{D}_{ucp}$ is continuous extends to $J_X : \mathbb{L}_{ucp} \to \mathbb{D}_{ucp}$.

Definition 2.5.12 Let X be a semimartingale. The continuous linear functional $J_X : \mathbb{L}_{ucp} \to \mathbb{D}_{ucp}$ is called the *stochastic integral*.

R If one restricts to integrators which are continuous L^2-martingales (like Brownian motion), the class of integrands can be taken even larger. In this situation all elements of V, see Definition 2.1.2, are possible integrands.

2.6 Properties of stochastic integrals

Proposition 2.6.1 Let X_1, X_2 be semimartingales and $Y_1, Y_2 \in \mathbb{L}$. It holds true that:

(i) $$\int (Y_1 + Y_2)\, d(X_1 + X_2) = \int Y_1\, d(X_1 + X_2) + \int Y_2\, d(X_1 + X_2) = \int Y_1\, dX_1 + \int Y_1\, dX_2 + \int Y_2\, dX_1 + \int Y_2\, dX_2\,.$$

(ii) $\Delta\big(\int Y_1\, dX_1\big) = Y_1\, \Delta X_1\,.$

(iii) $\int Y_1\, dX_1$ is a semimartingale and

$$\int Y_2\, d\Big(\int Y_1\, dX_1\Big) = \int Y_1 Y_2\, dX_1\,.$$

(iv) If X has paths of finite variation, $\int Y_1\, dX_1$ is indistinguishable from the Lebesgue-Stieltjes integral.

(v) $\mathbb{P} \ll \mathbb{Q}$. Then, $Y_1 \bullet_{\mathbb{P}} X_1 = Y_1 \bullet_{\mathbb{Q}} X_1$.

(vi) If X_1 is an L_2-martingale and $Y_1 \in \mathbf{b}\mathbb{L}$, then $\int Y_1\, dX_1$ is a martingale.

Proof. For (i)–(v) see Protter [1992], Chapter II.5. We shall prove (vi).
Let $|Y_1(t)| \le L$, $H^n \subseteq \mathcal{S}$ with $H^n \xrightarrow{ucp} Y_1$. We can assume that $|H^n(t)| \le L$. Then

$$\begin{aligned}
\mathbb{E}\Big[\Big(\Big(\int H^n\, dX_1\Big)(t)\Big)^2\Big] &= \mathbb{E}\Big[\sum_{i=1}^{n}\Big(H^n_{i-1}(X_1(T_i) - X_1(T_{i-1}))\Big)^2\Big] \\
&\le L^2\, \mathbb{E}\Big[\sum_{i=1}^{n}\big(X_1(T_i) - X_1(T_{i-1})\big)^2\Big] \\
&= L^2\, \mathbb{E}\big[X_\infty^2\big] < \infty\,.
\end{aligned}$$

Since $\|H^n \bullet X_1 - Y_1 \bullet X_1\|_{ucp} \to 0$, there exists for fix t a subsequence (n_k) with

$$\sup_{0\le s\le t}\Big(\big(H^{n_k} - Y_1\big) \bullet X_1\Big)(s) \to 0 \ \text{ a.s.}$$

Therefore, by Fatou's lemma

$$\begin{aligned}
\mathbb{E}\big[\big((Y_1 \bullet X_1)(t)\big)^2\big] &= \mathbb{E}\big[\lim_{n_k}\big((H^{n_k} \bullet X_1)(t)\big)^2\big] \\
&\le \liminf_{n_k}\mathbb{E}\big[\big((H^{n_k} \bullet X_1)(t)\big)^2\big] \\
&\le L^2\, \mathbb{E}\big[X_\infty^2\big]\,.
\end{aligned}$$

Thus, $Y_1 \bullet X_1$ is in L_2. It easily follows that $(H^n \bullet X_1)_{n\geq 1}$ are martingales.

$$\mathbb{P}\left(\left|\mathbb{E}\left[(Y_1 \bullet X_1)(t)|\mathcal{F}_s\right] - (Y_1 \bullet X_1)(s)\right| \geq \varepsilon\right) \leq \mathbb{P}\left(\mathbb{E}\left[\left((Y_1 - H^{n_k})\bullet X_1\right)(t)|\mathcal{F}_s\right] \geq \varepsilon/2\right) + \mathbb{P}\left(\left((Y_1 - H^{n_k})\bullet X_1\right)(s) \geq \varepsilon/2\right).$$

The second term tends to zero and for the first term we deduce that

$$\mathbb{E}\left[\mathbb{E}\left[\left|(Y_1 - H^{n_k})\bullet X_1\right|(t)|\mathcal{F}_s\right]\right] \leq \left(\mathbb{E}\left[\left(\left|(Y_1 - H^{n_k})\bullet X_1\right|(t)\right)^2\right]\mathbb{P}\left(\left|(Y_1 - H^{n_k})\bullet X_1\right|(t) > 1\right)\right)^{1/2} + \mathbb{E}\left[\left|(Y_1 - H^{n_k})\bullet X_1\right|(t)\mathbf{1}\{\left|(Y_1 - H^{n_k})\bullet X_1\right|(t) \leq 1\}\right]$$

and thus tends to zero by dominated convergence. We derive that

$$\mathbb{E}\left[\left|(Y_1 - H^{n_k})\bullet X_1\right|(t)|\mathcal{F}_s\right] \xrightarrow{p} 0.$$

what completes the proof. ■

We write that a sequence of random partitions $\sigma_n : T_1^n \leq \ldots \leq T_{k_n}^n$ *tends to identity* if

- $\lim_{n\to\infty} T_{k_n}^n = \infty$ a.s.
- The mesh $\sup_k |T_k^n - T_{k-1}^n| \to 0$ a.s.

For some given process Y and a random partition σ, we define the simple predictable process

$$Y^\sigma = \sum_{i=1}^{k} Y(\omega, T_{i-1})\mathbf{1}_{(T_{i-1},T_i]}(\omega,t). \tag{2.6.1}$$

Theorem 2.6.2 Let X be a semimartingale and $Y \in \mathbb{L}$ or $Y \in \mathbb{D}$. Let σ_n be a sequence of random partitions which tends to identity. Then, the elemenatry integrals

$$\int Y^\sigma(s)\,dX(s) = \sum_{i=1}^{k} Y(T_{i-1})\big(X(T_i \wedge \bullet) - X(T_{i-1}\wedge \bullet)\big)$$

converge ucp to the stochastic integral $\int Y_-\,dX$. ■

Proof. Without loss of generality, let $X(0) = 0$ and $Y_- \in \mathbb{L}$. Suppose $(Y^k) \subseteq \mathcal{S}$, such that $Y^k \xrightarrow{ucp} Y_-$. We have that

$$\int (Y_- - Y^{\sigma_n})(s)\,dX(s) = \int (Y_- - Y^k)(s)\,dX(s) + \int \left(Y^k - (Y^k)^{\sigma_n}\right)(s)\,dX(s) + \int \left((Y^k)^{\sigma_n} - Y^{\sigma_n}\right)(s)\,dX(s)$$

with X a càdlàg version. The first and the third addends tend to zero and it remains to show that

$$\int \left(Y^k - (Y^k)^{\sigma_n}\right)(s)\,dX(s) \xrightarrow{ucp} 0.$$

For fix k and as $n \to \infty$, this is ensured by right-continuity of X and since $\sup|T_i^n - T_{i-1}^n| \to 0$. ■

2.7 Quadratic variation

Definition 2.7.1 Let X and Y be semimartingales. The *quadratic variation process* $[X,X](t)$ of X is defined by

$$[X,X] := X^2 - 2\int X_- \, dX - (X(0))^2 . \tag{2.7.1a}$$

The *quadratic covariation process* $[X,Y](t)$ of X and Y is defined by

$$[X,Y] := XY - \int X_- \, dY - \int Y_- \, dX - X(0)Y(0) . \tag{2.7.1b}$$

R

- Bilinearity and symmetry gives the *polarization identity*

$$[X,Y] = \frac{1}{2}\Big([X+Y,X+Y] - [X,X] - [Y,Y]\Big) . \tag{2.7.2}$$

- Definition 2.7.1 is equivalent to the definition of quadratic variation used in Proposition 2.3.3, see below.

■ **Example 2.3** For a standard Brownian motion B, we obtain

$$[B,B](t) = (B(t))^2 - 2\int B(t)\, dB(t) = t \quad \forall\, t \in \mathbb{R}_+ .$$

■

Proposition 2.7.2 For a semimartingale X, the quadratic variation process $[X,X]$ is an adapted, monotone increasing, càdlàg process that satisfies

(i) $[X,X](t) - [X,X]_-(t) = \big(\Delta X(t)\big)^2$.

(ii) For any sequence of random partitions $0 = T_0^n \le T_1^n \le \ldots \le T_{k_n}^n$ which tends to identity, it holds true that

$$\sum_{i=1}^{k_n} \big(X(T_i^n \wedge \bullet) - X(T_{i-1}^n \wedge \bullet)\big)^2 \xrightarrow{ucp} [X,X] .$$

(iii) For a stopping time τ, we have

$$[X(\tau \wedge \bullet), X] = [X, X(\tau \wedge \bullet)] = [X,X](\tau \wedge \bullet) .$$

(iv) $[X,X]$ is of bounded variation.

Proof. (i)
$$\begin{aligned} [X,X](t) - [X,X]_-(t) &= \Delta(X^2(t)) - 2X(t-)\Delta X(t) \\ &= (X(t))^2 - (X(t-))^2 - 2X(t-)\Delta X(t) \\ &= \big(X(t) - X(t-)\big)^2 = \big(\Delta X(t)\big)^2 . \end{aligned}$$

(ii)
$$\begin{aligned}\sum_{i=1}^{k_n}\big(X(T_i^n\wedge t)-X(T_{i-1}^n\wedge t)\big)^2 &= \sum_{i=1}^{k_n}\big((X(T_i^n\wedge t))^2-(X(T_{i-1}^n\wedge t))^2\big)\\ &\quad -2\sum_{i=1}^{k_n}X(T_{i-1}^n\wedge t)\big(X(T_i^n\wedge t)-X(T_{i-1}^n\wedge t)\big)\\ &= (X(t))^2-(X(0))^2-2\sum_{i=1}^{k_n}X(T_{i-1}^n\wedge t)\big(X(T_i^n\wedge t)-X(T_{i-1}^n\wedge t)\big)\\ &\xrightarrow{ucp}(X(t))^2-(X(0))^2-2\int_0^t X_-\,dX\,.\end{aligned}$$

(ii) implies (iii). (iv) follows by monotonicity, for the covariation with polarization. ■

For a semimartingale X_t, we decompose the quadratic variation

$$[X,X](t)=[X,X]^c(t)+\sum_{0\le s\le t}(\Delta X_s)^2\,, \tag{2.7.3}$$

in the quadratic variation of the continuous part and the sum of squared jumps.
We conclude that a continuous semimartingale of bounded variation is almost surely constant.

Corollary 2.7.3 — Partial integration. For semimartingales X and Y, XY is a semimartingale and

$$XY-X_0Y_0=\int X_-\,dY+\int Y_-\,dX+[X,Y]\,.$$

Proposition 2.7.4 Let X,Y be semimartingales and $H,K\in\mathbb{L}$. It holds true that

$$[H\bullet X,K\bullet Y](t)=\int_0^t H(s)K(s)\,d[X,Y](s)\,. \tag{2.7.4}$$

Proof. It suffices to show that

$$[H\bullet X,Z]=H\bullet[X,Z]$$

for a semimartingale Z. Without loss of generality, $X_0=Y_0=0$. For $H(s)=\mathbf{1}_{[0,T]}(s)$ with a stopping time T, we have by Proposition 2.7.2 (iii)

$$[X(T\wedge\bullet),Z]=[X,Z](T\wedge\bullet)\,.$$

For $H(s)=U\,\mathbf{1}_{(S,T]}(s)$ with stopping times S,T and $U\in\mathfrak{F}_S$, it holds that

$$\int H(s)\,dX(s)=U\big(X(T\wedge\bullet)-X(S\wedge\bullet)\big)$$

and

$$\begin{aligned}[H\bullet X,Z]&=[U\big(X(T\wedge\bullet)-X(S\wedge\bullet)\big),Z]\\ &=U\Big([X,Z](T\wedge\bullet)-[X,Z](S\wedge\bullet)\Big)\\ &=\int H(s)\,d[X,Z](s)=H\bullet[X,Z]\,.\end{aligned}$$

For $H \in \mathcal{S}$, the claim follows by linearity. For $H \in \mathbb{L}$ there exists a sequence $(H^n), H^n \in \mathcal{S}$ with $H^n \xrightarrow{ucp} H$. Thereby

$$\begin{aligned}[H^n \bullet X, Z] &= (H^n \bullet X)Z - \int Z_- \, d(H^n \bullet X) - \int (H^n \bullet X)_- \, dZ \\ &= (H^n \bullet X)Z - \int Z_- H^n - dX - \int (H^n X)_- \, dZ ,\end{aligned}$$

using partial integration. Now, $(H^n X)_- \to (HX)_-$ in $\mathbb{L}_{ucp}$, $H^n \to H$ in $\mathbb{L}_{ucp}$, such that

$$\lim_{n\to\infty} [H^n \bullet X, Z] = [H \bullet X, Z] = \lim_{n\to\infty} H^n \bullet [X, Z] ,$$

and the proposition follows. ■

Corollary 2.7.5 — Itô isometry. Let W be a standard Brownian motion and $H \in \mathbb{L}$. It holds that

$$[H \bullet W, H \bullet W](t) = \int_0^t H(s)^2 \, ds$$

as well as

$$\mathbb{E}\left[\left(\int_0^t H(s)\, dW(s)\right)^2\right] = \int_0^t \mathbb{E}\left[H^2(s)\right] ds$$

with Fubini's theorem.

2.8 Multidimensional Case

Definition 2.8.1

1. An m-dimensional process is a semimartingale if each component is a one-dimensional semimartingale.
2. An $\mathbb{R}^m$-valued $(\mathcal{F}_t)$-adapted stochastic process $W(t) = (W_1(t), \dots, W_m(t))^T$ is an *m-dimensional Brownian motion* if each component W_i, $i = 1, \dots, m$, is a one-dimensional $(\mathcal{F}_t)$-Brownian motion and all components are independent.
3. If Y is an $\mathbb{R}^{d\times m}$-valued stochastic process such that each component Y_{ij}, $1 \le i \le d$, $1 \le j \le m$, is an element of V^*, see Definition 2.1.2, then the multidimensional Itô integral $\int Y \, dW$ for m-dimensional Brownian motion W is an $\mathbb{R}^d$-valued random variable with components

$$\left(\int_0^\infty Y(t)\, dW(t)\right)_i := \sum_{j=1}^m \int_0^\infty Y_{ij}(t)\, dW_j(t), \quad 1 \le i \le d ,$$

and analogously for m-dimensional semimartingales.

Proposition 2.8.2 The Itô isometry extends to the multidimensional case such that for $\mathbb{R}^{d\times m}$-valued processes X, Y with components in V and m-dimensional Brownian motion W

$$\mathbb{E}\left[\left\langle \int_0^\infty X(t)\, dW(t), \int_0^\infty Y(t)\, dW(t) \right\rangle\right] = \int_0^\infty \sum_{i=1}^d \sum_{j=1}^m \mathbb{E}[X_{ij}(t) Y_{ij}(t)]\, dt .$$

Proof. The term in the brackets on the left hand side is equal to

$$\sum_{i=1}^{d}\sum_{j=1}^{m}\sum_{k=1}^{m}\int_0^\infty X_{ij}(t)\,dW_j(t)\int_0^\infty Y_{ik}(t)\,dW_k(t)$$

and the result follows from the one-dimensional Itô isometry once the following claim has been proved: stochastic integrals with respect to independent Brownian motions are uncorrelated (attention: they may well be dependent).

For this, let us consider two independent Brownian motions W_1 and W_2 and two simple processes Y_1, Y_2 in V on the same filtered probability space with

$$Y_k(t) = \sum_{i=0}^{\infty} \eta_{ik}(\omega)\mathbf{1}_{[t_i,t_{i+1})}(t), \quad k \in \{1,2\}.$$

The common partition of the time axis can always be achieved by taking a common refinement of the two partitions. Then by the $\mathcal{F}_{t_i}$-measurability of η_{ik} we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^\infty Y_1(t)\,dW_1(t)\int_0^\infty Y_2(t)\,dW_2(t)\Big]\\
&= \sum_{0\le i\le j<\infty} \mathbb{E}\big[\eta_{i1}\eta_{j2}(W_1(t_{i+1})-W_1(t_i))(W_2(t_{j+1})-W_2(t_j))\big]\\
&\quad + \sum_{0\le j<i<\infty} \mathbb{E}\big[\eta_{i1}\eta_{j2}(W_1(t_{i+1})-W_1(t_i))(W_2(t_{j+1})-W_2(t_j))\big]\\
&= \sum_{0\le i\le j<\infty} \mathbb{E}\big[\eta_{i1}\eta_{j2}(W_1(t_{i+1})-W_1(t_i))\big]\,\mathbb{E}\big[(W_2(t_{j+1})-W_2(t_j))\big]\\
&\quad + \sum_{0\le j<i<\infty} \mathbb{E}\big[\eta_{i1}\eta_{j2}(W_1(t_{j+1})-W_1(t_j))\big]\,\mathbb{E}\big[(W_2(t_{i+1})-W_2(t_i))\big]\\
&= 0
\end{aligned}$$

For each process in V there exists a sequence of simple processes such that the corresponding stochastic integrals converge in $L^2(\mathbb{P})$, which implies that the respective covariances converge, too. This density argument proves the general case. ■

2.9 The Itô formula

Theorem 2.9.1 — Itô's lemma. Let X be a one-dimensional continuous semimartingale and $F \in \mathcal{C}^{2,1}(\mathbb{R}\times\mathbb{R}_+,\mathbb{R})$. Then, $F(X(t),t)$ is a continuous semimartingale with

$$\begin{aligned}
F(X(t),t) = F(X(0),0) &+ \int_0^t \frac{\partial F}{\partial t}(X(s),s)\,ds + \int_0^t \frac{\partial F}{\partial x}(X(s),s)\,dX(s) \qquad (2.9.1)\\
&+ \frac{1}{2}\int_0^t \frac{\partial^2 F}{\partial x^2}(X(s),s)\,d[X,X](s)\,.
\end{aligned}$$

For a d-dimensional continuous semimartingale and $F \in \mathcal{C}^{2,1}(\mathbb{R}^d\times\mathbb{R}_+,\mathbb{R})$, $F(X(t),t)$

is a continuous semimartingale with

$$F(X(t),t) = F(X(0),0) + \int_0^t \frac{\partial F}{\partial t}(X(s),s)\,ds + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X(s),s)\,dX_i(s)$$
$$+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X(s),s)\,d[X_i,X_j](s)\,. \tag{2.9.2}$$

Proof. The equations hold for constant F. We first restrict to $F(X(t))$ with $F \in \mathcal{C}^2(\mathbb{R}^d,\mathbb{R})$. Let F be a polynomial of degree n and $G(X(t)) = X_l(t)\,F(X(t))$, $1 \le l \le d$. By virtue of additivity it suffices to consider monomials. Partial integration yields

$$X_l(t)\,F(X(t)) = X_l(0)\,F(X(0)) + \int_0^t X_l(s)\,dF(X(s)) + [F(X),X_l](t) + \int_0^t F(X(s))\,dX_l(s).$$

We pursue an induction. Using associativity and by the induction hypothesis, we have that

$$\int_0^t X_l(s)\,dF(X(s)) = \sum_{i=1}^d \int_0^t X_l(s)\frac{\partial F}{\partial x_i}(X(s))\,dX_i(s) + \frac{1}{2}\sum_{i,j=1}^d \int_0^t X_l(s)\frac{\partial^2 F}{\partial x_i \partial x_j}(X(s))\,d[X_i,X_j](s).$$

By Corollary 2.7.3:

$$\begin{aligned}
[F(X),X_l](t) &= \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X(s))\,d[X_i,X_l](s)\\
&\quad + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X(s))\,d[[X_i,X_j],X_l](s)\\
&= \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X(s))\,d[X_i,X_l](s)\,.
\end{aligned}$$

We obtain that

$$\begin{aligned}
G(X(t)) = X_l(0)\,F(X(0)) &+ \sum_{i=1}^d \int_0^t X_l(s)\frac{\partial F}{\partial x_i}(X(s))\,dX_i(s)\\
&+ \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X(s))\,d[X_i,X_l](s)\\
&+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t X_l(s)\frac{\partial^2 F}{\partial x_i \partial x_j}(X(s))\,d[X_i,X_j](s)\\
&+ \int_0^t F(X(s))\,dX_l(s)\,,
\end{aligned}$$

such that

$$\frac{\partial G}{\partial x_i}(X(s)) = X_l(s)\frac{\partial F}{\partial x_i}(X(s)) + \delta_{il}F(X(s))\,,$$

with δ_{il} being Kronecker's delta. Furthermore, we derive that

$$\frac{\partial^2 G}{\partial x_i \partial x_j}(X(s)) = X_l(s)\frac{\partial^2 F}{\partial x_i \partial x_j}(X(s)) + (\delta_{il} + \delta_{jl})\frac{\partial F}{\partial x_l}(X(s)).$$

The above identities imply the formula (2.9.2).
By the Stone-Weierstrass-Theorem (cf. Klenke [2008], page 302) for any continuous function $F \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ there exists a sequence of polynomials which uniformly converges to F. Continuity of the stochastic integral thus ensures ucp-convergence. The general case for $F \in \mathcal{C}^{2,1}(\mathbb{R} \times \mathbb{R}_+, \mathbb{R})$ follows with continuity of F and the usual convergence to the Riemann integral. ■

We refer to Theorem II.32 on pages 78-79 of Protter [1992] for a more general version of Itô's lemma (resp. the Itô formula) including jumps.

II Stochastic differential equations

3. Strong solutions of SDEs

3.1 The strong solution concept

The first definition of a solution of a stochastic differential equation reflects the interpretation that the solution process X at time t is determined by the equation and the exogenous input of the initial condition and the path of the Brownian motion up to time t. Mathematically, this is translated into a measurability condition on X_t or equivalently into the smallest reasonable choice of the filtration to which X should be adapted, see condition (a) below.

Definition 3.1.1 A *strong solution* X of the stochastic differential equation

$$dX(t) = b(X(t),t)\,dt + \sigma(X(t),t)\,dW(t), \quad t \geq 0, \tag{3.1.1}$$

with $b : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}^{d\times m}$ measurable, on the given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the fixed m-dimensional Brownian motion W and the independent initial condition X_0 over this probability space is a stochastic process $(X(t), t \geq 0)$ satisfying:

(a) X is adapted to the filtration $(\mathcal{G}_t)$, where $\mathcal{G}_t^0 := \sigma(W(s), 0 \leq s \leq t) \vee \sigma(X_0)$ and $\mathcal{G}_t$ is the completion of $\bigcap_{s>t} \mathcal{G}_s^0$ with $\mathbb{P}$-null sets;
(b) X is a continuous process;
(c) $\mathbb{P}(X(0) = X_0) = 1$;
(d) $\mathbb{P}\left(\int_0^t (\|b(X(s),s)\| + \|\sigma(X(s),s)\|^2)\,ds < \infty\right) = 1$ holds for all $t > 0$;
(e) With probability one we have

$$X(t) = X(0) + \int_0^t b(X(s),s)\,ds + \int_0^t \sigma(X(s),s)\,dW(s), \quad \forall t \geq 0.$$

R It can be shown [Karatzas and Shreve, 1991, Section 2.7] that the completion of the filtration of Brownian motion (or more generally of any strong Markov process) is right-continuous. This means that $\mathcal{G}_t$ equals already the completion of $\mathcal{G}_t^0$.

With this definition at hand the notion of the existence of a strong solution is clear. We will say that strong uniqueness of a solution holds, only if the construction of a strong solution is unique on any probability space carrying the random elements W and X_0, where X_0 is an arbitrary initial condition.

Definition 3.1.2 Suppose that, whenever $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space equipped with a Brownian motion W and an independent random variable X_0, any two strong solutions X and X' of (3.1.1) with initial condition X_0 satisfy $\mathbb{P}(\forall t \geq 0 : X_t = X'_t) = 1$. Then we say that *strong uniqueness* holds for equation (3.1.1) or more precisely for the pair (b, σ).

R Since solution processes are by definition continuous and $\mathbb{R}_+$ is separable, it suffices to have the weaker condition $\mathbb{P}(X_t = X'_t) = 1$ for all $t \geq 0$ in the above definition.

3.2 Uniqueness

Example 3.1 Consider the one-dimensional equation

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

with a bounded, Borel-measurable function $b : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ that is non-increasing in the first variable. Then strong uniqueness holds for this equation, that is for the pair $(b, 1)$. To prove this define for two strong solutions X and X' on the same filtered probability space the process $D(t) := X(t) - X'(t)$. This process is (weakly) differentiable with

$$\frac{d}{dt} D^2(t) = 2D(t)\dot{D}(t) = 2(X(t) - X'(t))(b(X(t), t) - b(X'(t), t)) \leq 0, \text{ a.e.}$$

From $X(0) = X'(0)$ we infer $D^2(t) = 0$ for all $t \geq 0$. ■

Already for deterministic differential equations examples of non-uniqueness are well known. For instance, the differential equation $\dot{x}(t) = |x(t)|^\alpha$ with $0 < \alpha < 1$ and $x(0) = 0$ has the family of solutions $x_\tau(t) = ((t - \tau)/\beta)^\beta$ for $t \geq \tau$, $x_\tau(t) = 0$ for $t \leq \tau$ with $\beta = 1/(1 - \alpha)$ and $\tau \geq 0$. The usual sufficient condition for uniqueness in the deterministic theory is Lipschitz continuity of the analogue of the drift function in the space variable. Also for SDEs Lipschitz continuity, even in its local form, suffices. First, we recall the classical Gronwall Lemma.

Lemma 3.2.1 Let $T > 0$, $c \geq 0$ and $u, v : [0, T] \to \mathbb{R}_+$ be measurable functions. If u is bounded and v is integrable, then

$$u(t) \leq c + \int_0^t u(s)v(s)\,ds \quad \forall t \in [0, T]$$

implies

$$u(t) \leq c \exp\Big(\int_0^t v(s)\,ds\Big), \quad t \in [0, T].$$

Proof. Suppose $c > 0$ and set

$$z(t) := c + \int_0^t u(s)v(s)\,ds, \quad t \in [0, T].$$

Then $u(t) \le z(t)$, $z(t)$ is weakly differentiable and for almost all t

$$\frac{\dot{z}(t)}{z(t)} = \frac{u(t)v(t)}{z(t)} \le v(t)$$

holds so that $\log(z(t)) \le \log(z(0)) + \int_0^t v(s)\,ds$ follows. This shows that

$$u(t) \le z(t) \le c \exp\Big(\int_0^t v(s)\,ds\Big), \quad t \in [0,T].$$

For $c = 0$ apply the inequality for $c_n > 0$ with $\lim_n c_n = 0$ and take the limit. ∎

Theorem 3.2.2 Suppose that b and σ are locally Lipschitz continuous in the space variable, that is, for all $n \in \mathbb{N}$ there is a $K_n > 0$ such that for all $t \ge 0$ and all $x, y \in \mathbb{R}^d$ with $\|x\|, \|y\| \le n$

$$\|b(x,t) - b(y,t)\| + \|\sigma(x,t) - \sigma(y,t)\| \le K_n \|x - y\|$$

holds. Then strong uniqueness holds for equation (3.1.1).

Proof. Let two solutions X and X' of (3.1.1) with the same initial condition X_0 be given on some common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We define the stopping times $\tau_n := \inf\{t > 0 \,|\, \|X(t)\| \ge n\}$ and τ_n' in the same manner for X', $n \in \mathbb{N}$. Then $\tau_n^* := \tau_n \wedge \tau_n'$ converges $\mathbb{P}$-almost surely to infinity. The difference $X(t \wedge \tau_n^*) - X'(t \wedge \tau_n^*)$ equals $\mathbb{P}$-almost surely

$$\int_0^{t \wedge \tau_n^*} (b(X(s),s) - b(X'(s),s))\,ds + \int_0^{t \wedge \tau_n^*} (\sigma(X(s),s) - \sigma(X'(s),s))\,dW(s).$$

We conclude by Itô isometry and Cauchy-Schwarz inequality:

$$\begin{aligned}
&\mathbb{E}[\|X(t \wedge \tau_n^*) - X'(t \wedge \tau_n^*)\|^2] \\
&\le 2\mathbb{E}\Big[\Big(\int_0^{t \wedge \tau_n^*} \|b(X(s),s) - b(X'(s),s)\|\,ds\Big)^2\Big] + 2\mathbb{E}\Big[\int_0^{t \wedge \tau_n^*} \|\sigma(X(s),s) - \sigma(X'(s),s)\|^2\,ds\Big] \\
&\le 2TK_n^2 \int_0^t \mathbb{E}[\|X(s \wedge \tau_n^*) - X'(s \wedge \tau_n^*)\|^2]\,ds + 2K_n^2 \int_0^t \mathbb{E}[\|X(s \wedge \tau_n^*) - X'(s \wedge \tau_n^*)\|^2]\,ds.
\end{aligned}$$

By Gronwall's inequality we conclude $\mathbb{P}(X(t \wedge \tau_n^*) = X'(t \wedge \tau_n^*)) = 1$ for all $n \in \mathbb{N}$ and $t \in [0,T]$. Letting $n, T \to \infty$, we see that $X(t) = X'(t)$ holds $\mathbb{P}$-almost surely for all $t \ge 0$ and with the remark from page 30 strong uniqueness follows. ∎

R In the one-dimensional case strong uniqueness already holds for Hölder-continuous diffusion coefficient σ of order $1/2$, see [Karatzas and Shreve, 1991, Proposition 5.2.13] for more details and refinements.

3.3 Existence

In the deterministic theory differential equations are usually solved locally around the initial condition. In the stochastic framework one is rather interested in global solutions and then uses appropriate stopping in order to solve an equation up to some random explosion time. To exclude explosions in finite time, the linear growth of the coefficients suffices. The standard example for explosion is the ODE

$$\dot{x}(t) = x(t)^2, \quad t \geq 0, \qquad x(0) \neq 0.$$

Its solution is given by $x(t) = 1/(x_0^{-1} - t)$ which explodes for $x_0 > 0$ and $t \uparrow x_0^{-1}$. Note already here that with the opposite sign $\dot{x}(t) = -x(t)^2$ the solution $x(t) = x(0)/(1+t)$ exists globally. Intuitively, the different behaviour is clear because in the first case x grows the faster the further away from zero it is ("positive feedback"), while in the second case x monotonically converges to zero ("negative feedback").

We shall first establish an existence theorem under rather strong growth and Lipschitz conditions and then later improve on that.

Theorem 3.3.1 Suppose that the coefficients satisfy the global Lipschitz and linear growth conditions

$$\|b(x,t) - b(y,t)\| + \|\sigma(x,t) - \sigma(y,t)\| \leq K\|x-y\| \quad \forall x,y \in \mathbb{R}^d, t \geq 0 \tag{3.3.1}$$

$$\|b(x,t)\| + \|\sigma(x,t)\| \leq K(1+\|x\|) \quad \forall x \in \mathbb{R}^d, t \geq 0 \tag{3.3.2}$$

with some constant $K > 0$. Moreover, suppose that on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ there exists an m-dimensional Brownian motion W and an initial condition X_0 with $\mathbb{E}[\|X_0\|^2] < \infty$. Then there exists a strong solution of the SDE (3.1.1) with initial condition X_0 on this probability space, which in addition satisfies with some constant $C > 0$ the moment bound

$$\mathbb{E}[\|X(t)\|^2] \leq C(1+\mathbb{E}[\|X_0\|^2])e^{Ct}, \quad t \geq 0.$$

■

Proof. As in the deterministic case we perform successive approximations and apply a Banach fixed point argument ("Picard-Lindelöf iteration"). Define recursively

$$X^0(t) := X_0, \quad t \geq 0 \tag{3.3.3}$$

$$X^{n+1}(t) := X_0 + \int_0^t b(X^n(s),s)\,ds + \int_0^t \sigma(X^n(s),s)\,dW(s), \quad t \geq 0. \tag{3.3.4}$$

Obviously, the processes X^n are continuous and adapted to the filtration generated by X_0 and W. Let us fix some $T > 0$. We are going to show that for arbitrary $t \in [0,T]$

$$\mathbb{E}\Big[\sup_{0\leq s\leq t} \|X^{n+1}(s) - X^n(s)\|^2\Big] \leq C_1 \frac{(C_2 t)^n}{n!} \tag{3.3.5}$$

holds with suitable constants $C_1, C_2 > 0$ independent of t and n and $C_2 = O(T)$. Let us see how we can derive the theorem from this result. From Chebyshev's inequality we obtain

$$\mathbb{P}\Big(\sup_{0\leq s\leq T} \|X^{n+1}(s) - X^n(s)\| > 2^{-n-1}\Big) \leq 4C_1 \frac{(4C_2T)^n}{n!}.$$

The term on the right hand side is summable over n, whence by the Borel-Cantelli Lemma we conclude

$$\mathbb{P}\Big(\text{for infinitely many } n\text{: } \sup_{0\le s\le T}\|X^{n+1}(s)-X^n(s)\| > 2^{-n-1}\Big)=0.$$

Therefore, by summation $\sup_{m\ge 1}\sup_{0\le s\le T}\|X^{n+m}(s)-X^n(s)\|\le 2^{-n}$ holds for all $n\ge N(\omega)$ with some $\mathbb{P}$-almost surely finite random index $N(\omega)$. In particular, the random variables $X^n(s)$ form a Cauchy sequence $\mathbb{P}$-almost surely and converge to some limit $X(s)$, $s\in[0,T]$. Obviously, this limiting process X does not depend on T and is thus defined on $\mathbb{R}_+$. Since the convergence is uniform over $s\in[0,T]$, the limiting process X is continuous. Of course, it is also adapted by the adaptedness of X^n. Taking the limit $n\to\infty$ in equation (3.3.4), we see that X solves the SDE (3.1.1) up to time T because of

$$\sup_{0\le s\le T}\|b(X^n(s),s)-b(X_T(s),s)\|\le K\sup_{0\le s\le T}\|X^n(s)-X_T(s)\|\to 0 \text{ (in } L^2(\mathbb{P}))$$
$$\mathbb{E}\Big[\|\sigma(X^n(\bullet),\bullet)-\sigma(X(\bullet),\bullet)\|^2_{V([0,T])}\Big]\le K^2T\sup_{0\le s\le T}\mathbb{E}[\|X_n(s)-X(s)\|^2]\to 0.$$

Since $T>0$ was arbitrary, the equation (3.1.1) holds for all $t\ge 0$. From estimate (3.3.5) and the asymptotic bound $C_2=O(T)$ we finally obtain by summation over n and putting $T=t$ the asserted estimate on $\mathbb{E}[\|X(t)\|^2]$.

It thus remains to establish the claimed estimate (3.3.5), which follows essentially from Doob's martingale inequality and the type of estimates used for proving Theorem 3.2.2. Proceeding inductively, we infer from the linear growth condition that (3.3.5) is true for $n=0$ with some $C_1>0$. Assuming it to hold for $n-1$, we obtain with a constant $D>0$ from Doob's inequality (see Appendix 8.1):

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{0\le s\le t}\|X^{n+1}(s)-X^n(s)\|^2\Big]\\
&\le 2\mathbb{E}\Big[\sup_{0\le s\le t}\Big\|\int_0^s b(X^n(u),u)-b(X^{n-1}(u),u)\,du\Big\|^2\Big]\\
&\quad+2\mathbb{E}\Big[\sup_{0\le s\le t}\Big\|\int_0^s \sigma(X^n(u),u)-\sigma(X^{n-1}(u),u)\,dW(u)\Big\|^2\Big]\\
&\le 2K^2t\int_0^t\mathbb{E}[\|X^n(u)-X^{n-1}(u)\|^2]\,du+2DK^2\int_0^t\mathbb{E}[\|X^n(u)-X^{n-1}(u)\|^2]\,du\\
&\le (2K^2TC_1+2DC_1K^2)C_2^{n-1}\frac{t^n}{n!}.
\end{aligned}$$

The choice $C_2=2K^2(T+D)=O(T)$ thus gives the result. ■

The last theorem is the key existence theorem that allows generalizations into many directions. The most powerful one is essentially based on conditions such that a solution X exists locally and $\|X(t)\|^2$ remains bounded for all $t\ge 0$ ($L(x)=x^2$ is a Lyapunov function). Our presentation follows Durrett [1996].

Lemma 3.3.2 Suppose X_1 and X_2 are adapted continuous processes with $X_1(0)=X_2(0)$ and $\mathbb{E}[\|X_1(0)\|^2]<\infty$. Let $\tau_R:=\inf\{t\ge 0\,|\,\|X_1(t)\|\ge R \text{ or } \|X_2(t)\|\ge R\}$. If both X_1 and X_2 satisfy the stochastic differential equation (3.1.1) on the random time interval $[0,\tau_R]$ with Lipschitz conditions on the coefficients b and σ, then $X_1(t\wedge\tau_R)=X_2(t\wedge\tau_R)$ holds $\mathbb{P}$-almost surely for all $t\ge 0$.

Proof. We proceed as in the proof of inequality (3.3.5) and obtain for $0 \le t \le T$:

$$\mathbb{E}\Big[\sup_{0\le s\le t\wedge\tau_R} \|X_1(s)-X_2(s)\|^2\Big] \le 2K^2(t+D)\int_0^t \mathbb{E}[\|X_1(u\wedge\tau_R)-X_2(u\wedge\tau_R)\|^2]\,du$$
$$\le 2K^2(T+D)\int_0^t \mathbb{E}\Big[\sup_{0\le s\le u\wedge\tau_R}\|X_1(s)-X_2(s)\|^2\Big]\,du.$$

Hence, Gronwall's Lemma implies that the expectation is zero and the result follows. ■

Theorem 3.3.3 Suppose the drift and diffusion coefficients b and σ are locally Lipschitz continuous in the space variable and satisfy for some $B \ge 0$

$$2\langle x, b(x,t)\rangle + \mathrm{trace}(\sigma(x,t)\sigma(x,t)^T) \le B(1+\|x\|^2), \quad \forall x\in\mathbb{R}^d, t\ge 0,$$

then the stochastic differential equation (3.1.1) has a strong solution for any initial condition X_0 satisfying $\mathbb{E}[\|X_0\|^2] < \infty$. ■

Proof. We extend the previous theorem by a suitable cut-off scheme. For any $R > 0$ define coefficient functions b_R, σ_R such that

$$b_R(x) = \begin{cases} b(x), & \|x\|\le R,\\ 0, & \|x\|\ge 2R,\end{cases} \text{ and } \sigma_R(x) = \begin{cases} \sigma(x), & \|x\|\le R,\\ 0, & \|x\|\ge 2R,\end{cases}$$

and b_R and σ_R are interpolated for $\|x\| \in (R, 2R)$ in such a way that they are Lipschitz continuous in the state variable. Then let X_R be the by Theorem 3.3.1 unique strong solution to the stochastic differential equation with coefficients b_R and σ_R. Introduce the stopping time $\tau_R := \inf\{t\ge 0 \,|\, \|X_R(t)\| \ge R\}$. Then by Lemma 3.3.2 $X_R(t)$ and $X_S(t)$ coincide for $t \le \min(\tau_R, \tau_S)$ and we can define

$$X_\infty(t) := X_R(t) \text{ for } t \le \tau_R.$$

The process X_∞ will be a strong solution of the stochastic differential equation (3.1.1) if we can show $\lim_{R\to\infty}\tau_R = \infty$ $\mathbb{P}$-almost surely.

Put $\varphi(x) = 1 + \|x\|^2$. Then Itô's formula yields for any $t, R > 0$

$$\begin{aligned}
&e^{-Bt}\varphi(X_R(t)) - \varphi(X_R(0))\\
&= -B\int_0^t e^{-Bs}\varphi(X_R(s))\,ds + \sum_{i=1}^d \int_0^t e^{-Bs}2X_{R,i}(s)\,dX_{R,i}(s)\\
&\quad + \frac{1}{2}\sum_{i=1}^d\int_0^t e^{-Bs}2\sum_{j=1}^d \sigma_{ij}(X_R(s),s)^2\,ds\\
&= \text{local martingale}\\
&\quad + \int_0^t e^{-Bs}\Big(-B\varphi(X_R(s)) + 2\langle x, b_R(X_R(s),s)\rangle + \mathrm{trace}(\sigma_R(X_R(s),s)\sigma_R^T(X_R(s),s))\Big)\,ds.
\end{aligned}$$

Our assumption implies that $(e^{-B(t\wedge\tau_R)}\varphi(X_R(t\wedge\tau_R)))_{t\geq 0}$ is a supermartingale by the optional stopping theorem. We conclude

$$\begin{aligned}\mathbb{E}[\varphi(X_0)] &\geq \mathbb{E}\big[e^{-B(t\wedge\tau_R)}\varphi(X_R(t\wedge\tau_R))\big]\\ &= \mathbb{E}\big[e^{-B(t\wedge\tau_R)}\varphi(X_\infty(t\wedge\tau_R))\big]\\ &\geq e^{-Bt}\,\mathbb{P}(\tau_R\leq t)\min_{\|x\|=R}\varphi(x).\end{aligned}$$

Because of $\lim_{\|x\|\to\infty}\varphi(x)=\infty$ we have $\lim_{R\to\infty}\mathbb{P}(\tau_R\leq t)=0$. Since the events $(\{\tau_R\leq t\})_{R>0}$ decrease, there exists for all $t>0$ and $\mathbb{P}$-almost all ω an index R_0 such that $\tau_R(\omega)\geq t$ for all $R\geq R_0$, which is equivalent to $\tau_R\to\infty$ $\mathbb{P}$-almost surely. ■

3.4 Explicit solutions

3.4.1 Linear Equations

In this paragraph we want to study the linear or affine equations

$$dX(t)=\big(A(t)X(t)+a(t)\big)\,dt+\sigma(t)\,dW(t),\quad t\geq 0. \tag{3.4.1}$$

Here, A is a $d\times d$-matrix, a is a d-dimensional vector and σ is a $d\times m$-dimensional matrix, where all objects are deterministic as well as measurable and locally bounded in the time variable. As usual, W is an m-dimensional Brownian motion and X a d-dimensional process.

The corresponding deterministic linear equation

$$\dot{x}(t)=A(t)x(t)+a(t),\quad t\geq 0, \tag{3.4.2}$$

has for every initial condition x_0 an absolutely continuous solution x, which is given by

$$x(t)=\Phi(t)\big(x_0+\int_0^t\Phi^{-1}(s)a(s)\,ds\big),\quad t\geq 0,$$

where Φ is the so-called *fundamental solution*. This means that Φ solves the matrix equation

$$\dot{\Phi}(t)=A(t)\Phi(t),\quad t\geq 0,\qquad \text{with } \Phi(0)=\mathrm{Id}.$$

In the case of a matrix A that is constant in time, the fundamental solution is given by

$$\Phi(t)=e^{At}:=\sum_{k=0}^{\infty}\frac{(tA)^k}{k!}.$$

For a thorough analysis of the deterministic case we refer to page 38 of Kong [2014].

Proposition 3.4.1 The strong solution X of equation (3.4.1) with initial condition X_0 is given by

$$X(t)=\Phi(t)\Big(X_0+\int_0^t\Phi^{-1}(s)a(s)\,ds+\int_0^t\Phi^{-1}(s)\sigma(s)\,dW(s)\Big),\quad t\geq 0.$$

Proof. Apply Itô's formula. ■

R

1. The function $\mu(t) := \mathbb{E}[X(t)]$ satisfies under the hypothesis $\mathbb{E}[|X(0)|] < \infty$ the deterministic linear differential equation (3.4.2).
2. The variance $v(t) = \mathbb{E}\left[(X(t)-\mu(t))(X(t)-\mu(t))^\top\right]$ obeys the linear equation

$$\dot{v}(t) = A(t)\,v(t) + v(t)A^\top(t) + \sigma(t)\sigma^\top(t)\,.$$

3.4.2 Transformation methods

We follow the presentation by Kloeden and Platen [1992], page 120, and consider scalar equations that can be solved explicitly by suitable transformations.

Consider the scalar stochastic differential equation

$$dX(t) = \tfrac{1}{2}b(X(t))b'(X(t))\,dt + b(X(t))\,dW(t), \tag{3.4.3}$$

where $b:\mathbb{R}\to\mathbb{R}$ is continuously differentiable with derivative b' and does not vanish and W is a one-dimensional Brownian motion. This equation is equivalent to the Fisk-Stratonovich equation

$$dX(t) = b(X(t))\circ dW(t).$$

Define

$$h(x) := \int_c^x \frac{1}{b(y)}\,dy \text{ for some } c\in\mathbb{R}\,.$$

Then $X(t) := h^{-1}(W(t)+h(X_0))$, where h^{-1} denotes the inverse of h which exists by monotonicity, solves the equation (3.4.3). This follows easily from $(h^{-1})'(W(t)+h(X_0)) = b(X(t))$ and $(h^{-1})''(W(t)+h(X_0)) = b'(X(t))b(X(t))$.

■ **Example 3.2**

1. (geometric Brownian motion) $dX(t) = \frac{\alpha^2}{2}X(t)\,dt + \alpha X(t)\,dW(t)$ has the solution $X(t) = X_0\exp(\alpha W(t))$.
2. The choice $b(x) = \beta|x|^\alpha$ for $\alpha,\beta\in\mathbb{R}$ corresponds formally to the equation

$$dX(t) = \tfrac{1}{2}\alpha\beta^2|X(t)|^{2\alpha-1}\operatorname{sgn}(X(t))\,dt + \beta|X(t)|^\alpha\,dW(t).$$

For $\alpha<1$ we obtain formally the solution

$$X(t) = \Big|\beta(1-\alpha)W(t) + |X_0|^{1-\alpha}\operatorname{sgn}(X_0)\Big|^{\frac{1}{1-\alpha}}\operatorname{sgn}(\beta(1-\alpha)W(t) + |X_0|^{1-\alpha}\operatorname{sgn}(X_0)).$$

This is well defined and indeed a strong solution if $\frac{1}{1-\alpha}$ is nonnegative. The specific choice $\alpha = \frac{n-1}{n}$ with $n\in\mathbb{N}$ odd gives

$$X(t) = (\beta n^{-1}W(t) + \sqrt[n]{X_0})^n.$$

For even n and $X_0\geq 0$ this formula defines a solution of

$$dX(t) = \frac{(n-1)\beta^2}{n}X(t)^{(n-2)/n}\,dt + \beta X(t)^{(n-1)/n}\,dW(t),$$

and X remains nonnegative for all times $t\geq 0$. Observe that a solution exists, although the coefficients are not locally Lipschitz. One can show that for $n=2$ strong uniqueness holds, whereas for $n>2$ also the trivial process $X(t)=0$ is a solution.

3. The equation

$$dX(t) = -a^2 \sin(X(t))\cos^3(X(t))\,dt + a\cos^2(X(t))\,dW(t)$$

has for $X_0 \in (-\frac{\pi}{2}, \frac{\pi}{2})$ the solution $X(t) = \arctan(aW(t) + \tan(X_0))$, which remains contained in the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$. This can be explained by the fact that for $x = \pm\frac{\pi}{2}$ the coefficients vanish and for values x close to this boundary the drift pushes the process towards zero more strongly than the diffusion part can possibly disturb.
4. The equation

$$dX(t) = a^2 X(t)(1 + X(t)^2)\,dt + a(1 + X(t)^2)\,dW(t)$$

is solved by $X(t) = \tan(aW(t) + \arctan X_0)$ and thus explodes $\mathbb{P}$-almost surely in finite time.

■

The transformation idea allows certain generalizations. With the same assumptions on b and the same definition of h we can solve the equation

$$dX(t) = \Big(\alpha b(X(t)) + \frac{1}{2} b(X(t)) b'(X(t))\Big)\,dt + b(X(t))\,dW(t)$$

by $X(t) = h^{-1}(\alpha t + W(t) + h(X_0))$. Equations of the type

$$dX(t) = \Big(\alpha h(X(t)) b(X(t)) + \frac{1}{2} b(X(t)) b'(X(t))\Big)\,dt + b(X(t))\,dW(t)$$

are solved by $X(t) = h^{-1}(e^{\alpha t} h(X_0) + e^{\alpha t} \int_0^t e^{-\alpha s}\,dW(s))$.

Finally, we consider for $n \in \mathbb{N}$, $n \geq 2$, the equation

$$dX(t) = (aX(t)^n + bX(t))\,dt + cX(t)\,dW(t).$$

Writing $Y(t) = X(t)^{1-n}$ we obtain

$$\begin{aligned} dY(t) &= (1-n)X(t)^{-n}\,dX(t) + \tfrac{1}{2}(1-n)(-n)X(t)^{-n-1}c^2X^2(t)\,dt \\ &= (1-n)(a + (b - \tfrac{c^2}{2}n)Y(t))dt + (1-n)cY(t)\,dW(t). \end{aligned}$$

Hence, Y is a geometric Brownian motion and we obtain after transformation for all $X_0 \neq 0$

$$X(t) = e^{(b-\frac{c^2}{2})t + cW(t)} \Big(X_0^{1-n} + a(1-n) \int_0^t e^{(n-1)(b-\frac{c^2}{2})s + c(n-1)W(s)}\,ds \Big)^{1/(1-n)}.$$

In addition to the trivial solution $X(t) = 0$ we therefore always have a nonnegative global solution in the case $X_0 \geq 0$ and $a \leq 0$. For odd integers n and $a \leq 0$ a global solution exists for any initial condition, cf. Theorem 3.3.3. In the other cases it is easily seen that the solution explodes in finite time.

4. Weak solutions of SDEs

4.1 The weak solution concept

We start with the famous example of H. Tanaka. Consider the scalar SDE

$$dX(t) = \operatorname{sgn}(X(t))\,dW(t), \quad t \geq 0, \qquad X(0) = 0, \tag{4.1.1}$$

where $\operatorname{sgn}(x) = \mathbf{1}_{(0,\infty)}(x) - \mathbf{1}_{(-\infty,0]}(x)$. Any adapted process X satisfying (4.1.1) is a continuous martingale with quadratic variation $[X,X](t) = t$. Lévy's characterization implies that X has the law of Brownian motion. If X satisfies this equation, then so does $-X$, since the Lebesgue measure of $\{t \in [0,T] \,|\, X(t) = 0\}$ vanishes almost surely for any Brownian motion. Hence strong uniqueness cannot hold.

We now invert the roles of X and W, for equation (4.1.1) obviously implies $dW(t) = \operatorname{sgn}(X(t))\,dX(t)$. Hence, we take a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a Brownian motion X and consider the filtration $(\mathcal{F}_t^X)_{t\geq 0}$ generated by X and completed under $\mathbb{P}$. Then we define the process

$$W(t) := \int_0^t \operatorname{sgn}(X(s))\,dX(s), \quad t \geq 0.$$

W is a continuous $(\mathcal{F}_t^X)$-adapted martingale with quadratic variation $[W,W](t) = t$, hence also an $(\mathcal{F}_t^X)$-Brownian motion. The couple (X,W) then solves the Tanaka equation. However, X is not a strong solution because the filtration $(\mathcal{F}_t^W)_{t\geq 0}$ generated by W and completed under $\mathbb{P}$ satisfies $\mathcal{F}_t^W \subsetneqq \mathcal{F}_t^X$ as we shall see.

For the proof let us take a sequence (f_n) of continuously differentiable functions on the real line that satisfy $f_n(x) = \operatorname{sgn}(x)$ for $|x| \geq \frac{1}{n}$ and $|f_n(x)| \leq 1$, $f_n(-x) = -f_n(x)$ for all $x \in \mathbb{R}$. If we set $F_n(x) = \int_0^x f_n(y)\,dy$, then $F_n \in C^2(\mathbb{R})$ and $\lim_{n\to\infty} F_n(x) = |x|$ holds uniformly on compact intervals. By Itô's formula for any solution X of (4.1.1)

$$F_n(X(t)) - \int_0^t f_n(X(s))dX(s) = \frac{1}{2}\int_0^t f_n'(X(s))ds, \quad t \geq 0,$$

follows and by Lebesgue's Theorem the left hand side converges in probability for $n \to \infty$ to $|X(t)| - \int_0^t \operatorname{sgn}(X(s))dX(s) = |X(t)| - W(t)$. By symmetry, $f_n'(x) = f_n'(|x|)$ and we have for $t \geq 0$ $\mathbb{P}$-almost surely

$$W(t) = |X(t)| - \lim_{n\to\infty} \frac{1}{2}\int_0^t f_n'(|X(s)|)ds.$$

Hence, $\mathcal{F}_t^W \subseteqq \mathcal{F}_t^{|X|}$ holds with obvious notation. The event $\{X(t) > 0\}$ has probability $\frac{1}{2} > 0$ and is not $\mathcal{F}_t^{|X|}$-measurable. Therefore $\mathcal{F}_t^X \setminus \mathcal{F}_t^{|X|}$ is non-void and $\mathcal{F}_t^W \subsetneqq \mathcal{F}_t^X$ holds for any solution X, which is thus not a strong solution in our definition. Note that the above derivation would be clearer with the aid of Tanaka's formula and the concept of local time.

Definition 4.1.1 A *weak solution* of the stochastic differential equation (3.1.1) is a triple (X,W), $(\Omega,\mathcal{F},\mathbb{P})$, $(\mathcal{F}_t)_{t\geq 0}$ where

(a) $(\Omega,\mathcal{F},\mathbb{P})$ is a probability space equipped with the filtration $(\mathcal{F}_t)_{t\geq 0}$ that satisfies the usual conditions;

(b) X is a continuous, $(\mathcal{F}_t)$-adapted $\mathbb{R}^d$-valued process and W is an m-dimensional $(\mathcal{F}_t)$-Brownian motion on the probability space;

(c) conditions (d) and (e) of Definition 3.1.1 are fulfilled.

The distribution $\mathbb{P}^{X(0)}$ of $X(0)$ is called *initial distribution* of the solution X.

R Any strong solution is also a weak solution with the additional filtration property $\mathcal{F}_t^X \subseteqq \mathcal{F}_t^W \vee \sigma(X(0))$. The Tanaka equation provides a typical example of a weakly solvable SDE that has no strong solution.

Definition 4.1.2 We say that *pathwise uniqueness* for equation (3.1.1) holds whenever two weak solutions (X,W), $(\Omega,\mathcal{F},\mathbb{P})$, $(\mathcal{F}_t)_{t\geq 0}$ and (X',W), $(\Omega,\mathcal{F},\mathbb{P})$, $(\mathcal{F}_t')_{t\geq 0}$ on a common probability space with a common Brownian motion with respect to both filtrations $(\mathcal{F}_t)$ and $(\mathcal{F}_t')$, and with $\mathbb{P}(X(0) = X'(0)) = 1$ satisfy $\mathbb{P}(\forall t \geq 0 : X(t) = X'(t)) = 1$.

Definition 4.1.3 We say that *uniqueness in law* (or weak uniqueness) holds for equation (3.1.1) whenever two weak solutions (X,W), $(\Omega,\mathcal{F},\mathbb{P})$, $(\mathcal{F}_t)_{t\geq 0}$ and (X',W'), $(\Omega',\mathcal{F}',\mathbb{P}')$, $(\mathcal{F}_t')_{t\geq 0}$ with the same initial distribution have the same law, that is $\mathbb{P}(X(t_1) \in B_1,\dots,X(t_n) \in B_n) = \mathbb{P}'(X'(t_1) \in B_1,\dots,X'(t_n) \in B_n)$ holds for all $n \in \mathbb{N}$, $t_1,\dots,t_n > 0$ and Borel sets $B_1,\dots,B_n$.

■ **Example 4.1** For the Tanaka equation pathwise uniqueness fails because X and $-X$ are at the same time solutions. We have, however, seen that X must have the law of a Brownian motion and thus uniqueness in law holds. ■

4.2 The two concepts of uniqueness

Let us discuss the notion of pathwise uniqueness and of uniqueness in law in some detail. When we consider weak solutions we are mostly interested in the law of the solution process so that uniqueness in law is usually all we require. However, as we shall see, the concept of pathwise uniqueness is stronger than that of uniqueness in law and if we reconsider the proof of Theorem 3.2.2 we immediately see that we have not used the special filtration properties of strong uniqueness and we obtain:

Corollary 4.2.1 Suppose that b and σ are locally Lipschitz continuous in the space variable, that is, for all $n \in \mathbb{N}$ there is a $K_n > 0$ such that for all $t \geq 0$ and all $x, y \in \mathbb{R}^d$ with $\|x\|, \|y\| \leq n$

$$\|b(x,t) - b(y,t)\| + \|\sigma(x,t) - \sigma(y,t)\| \leq K_n \|x - y\|$$

holds. Then pathwise uniqueness holds for equation (3.1.1).

The same remark applies to Example 3.1. As Tanaka's example has shown, pathwise uniqueness can fail when uniqueness in law holds. It is not clear, though, that the converse implication is true.

Theorem 4.2.2 Pathwise uniqueness implies uniqueness in law. ▪

Proof. We have to show that two weak solutions (X_i, W_i), $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$, $(\mathcal{F}_t^i)$, $i = 1, 2$ on possibly different filtered probability spaces agree in distribution. The main idea is to define two weak solutions with the same law on a common space with the same Brownian motion and to apply the pathwise uniqueness assumption. To this end we set

$$S := \mathbb{R}^d \times C(\mathbb{R}_+, \mathbb{R}^m) \times C(\mathbb{R}_+, \mathbb{R}^d), \qquad \mathcal{S} = \text{Borel } \sigma\text{-field of } S$$

and consider the image measures

$$\mathbb{Q}_i(A) := \mathbb{P}_i((X_i(0), W_i, X_i) \in A), \quad A \in \mathcal{S}, \quad i = 1, 2.$$

Since $X_i(t)$ is by definition $\mathcal{F}_t^i$-measurable, $X_i(0)$ is independent of W_i under $\mathbb{P}_i$. If we call μ the law of $X_i(0)$ under $\mathbb{P}_i$ (which by assumption does not depend on i), we thus have that the product measure $\mu \otimes \mathbb{W}$ is the law of the first two coordinates $(X_i(0), W_i)$ under $\mathbb{P}_i$, where $\mathbb{W}$ denotes the Wiener measure. Since $C(\mathbb{R}_+, \mathbb{R}^k)$ is a Polish space, a regular conditional distribution (Markov kernel) K_i of X_i under $\mathbb{P}_i$ given $(X_i(0), W_i)$ exists [Karatzas and Shreve, 1991, Section 5.3D] and we may write for Borel sets $F \subset \mathbb{R}^d \times C(\mathbb{R}_+, \mathbb{R}^m)$, $G \subset C(\mathbb{R}_+, \mathbb{R}^d)$

$$\mathbb{Q}_i(F \times G) = \int_F K_i(x_0, w; G)\, \mu(dx_0)\, \mathbb{W}(dw).$$

Let us now define

$$T = S \times C(\mathbb{R}_+, \mathbb{R}^d), \qquad \mathcal{T} = \text{Borel } \sigma\text{-field of } T$$

and equip this space with the probability measure

$$\mathbb{Q}(d(x_0, w, y_1, y_2)) = K_1(x_0, w; dy_1) K_2(x_0, w; dy_2) \mu(dx_0)\, \mathbb{W}(dw).$$

Finally, denote by $\mathcal{T}^*$ the completion of $\mathcal{T}$ under $\mathbb{Q}$ and consider the filtrations

$$\mathcal{T}_t = \sigma((x_0, w(s), y_1(s), y_2(s)), s \leq t)$$

and its $\mathbb{Q}$-completion $\mathcal{T}_t^*$ and its right-continuous version $\mathcal{T}^{**} = \bigcap_{s>t} \mathcal{T}_s^*$. Then the projection on the first coordinate has under $\mathbb{Q}$ the law of the initial distribution of X_i and the projection on the second coordinate is under $\mathbb{Q}$ an $\mathcal{T}_t^{**}$-Brownian motion (recall the remark from page 29). Moreover, the distribution of the projection (w, y_i) under $\mathbb{Q}$ is the same as that of (W_i, X_i) under $\mathbb{P}_i$ such that we have constructed two weak solutions on the same probability space with the same initial condition and the same Brownian motion.

Pathwise uniqueness now implies $\mathbb{Q}(\{(x_0, w, y_1, y_2) \in T \mid y_1 = y_2\}) = 1$. This entails

$$\mathbb{P}_1((W_1, X_1) \in A) = \mathbb{Q}((w, y_1) \in A) = \mathbb{Q}((w, y_2) \in A) = \mathbb{P}_2((W_2, X_2) \in A).$$

■

The same methodology allows to prove the following, at a first glance rather striking result.

Theorem 4.2.3 The existence of a weak solution and pathwise uniqueness imply the existence of a strong solution on any sufficiently rich probability space. ■

Proof. See [Karatzas and Shreve, 1991, Cor. 5.3.23]. ■

4.3 Existence via Girsanov's theorem

The Girsanov theorem is one of the main tools of stochastic analysis. In the theory of stochastic differential equations it often allows to extend results for a particular equation to those with more general drift coefficients. Abstractly seen, a Radon-Nikodym density for a new measure is obtained, under which the original process behaves differently. We only work in dimension one and start with a lemma on conditional Radon-Nikodym densities.

Lemma 4.3.1 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{H} \subset \mathcal{F}$ be a sub-σ-algebra and $f \in L^1(\mathbb{P})$ be a density, that is nonnegative and integrating to one. Then a new probability measure $\mathbb{Q}$ on $\mathcal{F}$ is defined by $\mathbb{Q}(d\omega) = f(\omega)\mathbb{P}(d\omega)$ and for any $\mathcal{F}$-measurable random variable X with $\mathbb{E}_{\mathbb{Q}}[|X|] < \infty$ we obtain

$$\mathbb{E}_{\mathbb{Q}}[X \mid \mathcal{H}]\,\mathbb{E}_{\mathbb{P}}[f \mid \mathcal{H}] = \mathbb{E}_{\mathbb{P}}[Xf \mid \mathcal{H}] \quad \mathbb{P}\text{-a.s.}$$

R In the unconditional case we obviously have

$$\mathbb{E}_{\mathbb{Q}}[X] = \int X \, d\mathbb{Q} = \int X f \, d\mathbb{P} = \mathbb{E}_{\mathbb{P}}[Xf].$$

Proof. We show that the left-hand side is a version of the conditional expectation on the right. Since it is obviously $\mathcal{H}$-measurable, it suffices to verify

$$\int_H \mathbb{E}_{\mathbb{Q}}[X \mid \mathcal{H}]\,\mathbb{E}_{\mathbb{P}}[f \mid \mathcal{H}]\, d\mathbb{P} = \int_H X f \, d\mathbb{P} = \int_H X \, d\mathbb{Q} \quad \forall H \in \mathcal{H}.$$

By the projection property of conditional expectations we obtain

$$\mathbb{E}_{\mathbb{P}}[\mathbf{1}_H \mathbb{E}_{\mathbb{Q}}[X \mid \mathcal{H}]\,\mathbb{E}_{\mathbb{P}}[f \mid \mathcal{H}]] = \mathbb{E}_{\mathbb{P}}[\mathbf{1}_H \mathbb{E}_{\mathbb{Q}}[X \mid \mathcal{H}] f] = \mathbb{E}_{\mathbb{Q}}[\mathbf{1}_H \mathbb{E}_{\mathbb{Q}}[X \mid \mathcal{H}]] = \mathbb{E}_{\mathbb{Q}}[\mathbf{1}_H X],$$

which is the above identity. ■

Lemma 4.3.2 Let $(\beta(t), 0 \le t \le T)$ be an $(\mathcal{F}_t)$-adapted process with $\beta \mathbf{1}_{t \le T} \in V^*$. Then

$$M(t) := \exp\Big(-\int_0^t \beta(s)\,dW(s) - \frac{1}{2}\int_0^t \beta^2(s)\,ds\Big), \quad 0 \le t \le T,$$

is an $(\mathcal{F}_t)$-supermartingale. It is a martingale if and only if $\mathbb{E}[M(T)] = 1$ holds.

Proof. If we apply Itô's formula to M, we obtain

$$dM(t) = -\beta(t)M(t)\,dW(t), \quad 0 \le t \le T.$$

Hence, M is always a nonnegative local $\mathbb{P}$-martingale. By Fatou's lemma for conditional expectations we infer that M is a supermartingale and a proper martingale if and only if $\mathbb{E}_{\mathbb{P}}[M(T)] = \mathbb{E}_{\mathbb{P}}[M(0)] = 1$. ∎

Lemma 4.3.3 M is a martingale if β satisfies one of the following conditions (whereas the implications 1. ⇒ 2. ⇒ 3. hold):

1. β is uniformly bounded;
2. Novikov's condition:
$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^T \beta^2(t)\,dt\Big)\Big] < \infty;$$
3. Kazamaki's condition:
$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^T \beta(t)\,dW(t)\Big)\Big] < \infty.$$

Proof. By the previous proof we know that M solves the linear SDE

$$dM(t) = -\beta(t)M(t)\,dW(t) \text{ with } M(0) = 1\,.$$

Since $\beta(t)$ is uniformly bounded, the diffusion coefficient satisfies the linear growth and Lipschitz conditions and we could modify Theorem 3.3.1 to cover also stochastic coefficients and obtain equally that $\sup_{0 \le t \le T} \mathbb{E}[M(t)^2]$ is finite. This implies $\beta M \mathbf{1}_{[0,T]} \in V$ and M is a martingale.

Alternatively, we prove $\beta M \mathbf{1}_{[0,T]} \in V$ by hand: If β is uniformly bounded by some $K > 0$, then we have for any $p > 0$ and any partition $0 = t_0 \le t_1 \le \cdots \le t_n = t$

$$\begin{aligned}
&\mathbb{E}\Big[\exp\Big(p\sum_{i=1}^n \beta(t_{i-1})(W(t_i) - W(t_{i-1}))\Big)\Big] \\
&= \mathbb{E}\Big[\exp\Big(p\sum_{i=1}^{n-1} \beta(t_{i-1})(W(t_i) - W(t_{i-1}))\Big)\,\mathbb{E}[\exp(p\beta(t_{n-1})(W(t_n) - W(t_{n-1}))\,|\,\mathcal{F}_{t_{n-1}}]\Big] \\
&= \mathbb{E}\Big[\exp\Big(p\sum_{i=1}^{n-1} \beta(t_{i-1})(W(t_i) - W(t_{i-1}))\Big)\exp(p^2\beta(t_{n-1})^2(t_n - t_{n-1})\Big] \\
&\le \mathbb{E}\Big[\exp\Big(p\sum_{i=1}^{n-1} \beta(t_{i-1})(W(t_i) - W(t_{i-1}))\Big)\exp(p^2K^2(t_n - t_{n-1})\Big] \\
&\le \exp\Big(\sum_{i=1}^n p^2K^2(t_i - t_{i-1})\Big) \\
&= \exp(p^2K^2t).
\end{aligned}$$

This shows that the random variables $\exp\Big(\sum_{i=1}^n \beta(t_{i-1})(W(t_i)-W(t_{i-1}))\Big)$ are uniformly bounded in any $L^p(\mathbb{P})$-space and thus uniformly integrable. Since by taking finer partitions these random variables converge to $\exp(\int_0^t \beta(s)\,dW(s))$ in $\mathbb{P}$-probability, we infer that $M(t)$ has finite expectation and even moments of all orders. Consequently, $\int_0^T \mathbb{E}[(\beta(t)M(t))^2]\,dt$ is finite and M is a martingale.

For the sufficiency of Novikov's and Kazamaki's condition we refer to Liptser and Shiryaev [2001] and the references there. ■

Theorem 4.3.4 Let $(X(t), 0 \le t \le T)$ be a stochastic (Itô) process on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying

$$X(t) = \int_0^t \beta(s)\,ds + W(t), \quad 0 \le t \le T,$$

with a Brownian motion W and a process $\beta \mathbf{1}_{t\le T} \in V^*$. If β is such that M is a martingale, then $(X(t), 0 \le t \le T)$ is a Brownian motion under the measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t))$ defined by $\mathbb{Q}(d\omega) = M(T,\omega)\,\mathbb{P}(d\omega)$. ■

Proof. We use Lévy's characterization of Brownian motion. Since M is a martingale, $M(T)$ is a density and $\mathbb{Q}$ is well-defined.

We put $Z(t) = M(t)X(t)$ and obtain by Itô's formula (or partial integration)

$$\begin{aligned} dZ(t) &= M(t)\,dX(t) + X(t)\,dM(t) + d\langle M, X\rangle_t \\ &= M(t)\Big(\beta(t)\,dt + dW(t) - X(t)\beta(t)\,dW(t) - \beta(t)dt\Big) \\ &= M(t)(1 - X(t)\beta(t))\,dW(t). \end{aligned}$$

This shows that Z is a local martingale. If Z is a martingale, then we accomplish the proof using the preceding lemma:

$$\mathbb{E}_{\mathbb{Q}}[X(t)\,|\,\mathcal{F}_s] = \frac{\mathbb{E}_{\mathbb{P}}[M(t)X(t)\,|\,\mathcal{F}_s]}{\mathbb{E}_{\mathbb{P}}[M(t)\,|\,\mathcal{F}_s]} = \frac{Z(s)}{M(s)} = X(s), \quad s \le t,$$

implies that X is a $\mathbb{Q}$-martingale which by its very definition has quadratic variation t. Hence, X is a Brownian motion under $\mathbb{Q}$.

If Z is only a local martingale with associated stopping times (τ_n), then the above relation holds for the stopped processes $X^{\tau_n}(t) = X(t \wedge \tau_n)$, which shows that X is a local $\mathbb{Q}$-martingale and Lévy's theorem applies. ■

Proposition 4.3.5 Suppose X is a stochastic process on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying for some $T > 0$ and measurable functions b and σ

$$dX(t) = b(X(t),t)\,dt + \sigma(X(t),t)\,dW(t), \quad 0 \le t \le T, \qquad X(0) = X_0.$$

Assume further that $u(x,t) := -c(x,t)/\sigma(x,t)$, c measurable, is such that

$$M(t) = \exp\Big(-\int_0^t u(X(s),s)\,dW(s) - \frac{1}{2}\int_0^t u^2(X(s),s)\,ds\Big), \quad 0 \le t \le T,$$

is an $(\mathcal{F}_t)$-martingale.

Then the stochastic differential equation

$$dY(t) = (b(Y(t),t) + c(Y(t),t))\,dt + \sigma(Y(t),t)\,dW(t), \quad 0 \le t \le T, \qquad Y(0) = X_0, \tag{4.3.1}$$

has a weak solution given by $((X,\widehat{W}),(\Omega,\mathcal{F},\mathbb{Q}),(\mathcal{F}_t))$ for the $\mathbb{Q}$-Brownian motion

$$\widehat{W}(t) := W(t) + \int_0^t u(X(s),s)\,ds, \quad t \ge 0,$$

and the probability $\mathbb{Q}$ given by $\mathbb{Q}(d\omega) := M(T,\omega)\,\mathbb{P}(d\omega)$.

R The martingale condition is for instance satisfied if σ is bounded away from zero and c is uniformly bounded. Putting $\sigma(x,t) = 1$ and $b(x,t) = 0$ we have weak existence for the equation $dX(t) = c(X(t),t)\,dt + dW(t)$ if c is Borel-measurable and satisfies a linear growth condition in the space variable, but without continuity assumption [Karatzas and Shreve, 1991, Prop. 5.36].

Proof. From Theorem 4.3.4 we infer that $\widehat{W}$ is a $\mathbb{Q}$-Brownian motion. Hence, we can write

$$dX(t) = \Big(b(X(t),t) - \sigma(X(t),t)u(X(t),t)\Big)\,dt + \sigma(X(t),t)\,d\widehat{W}(t),$$

which by definition of u shows that $(X,\widehat{W})$ solves under $\mathbb{Q}$ equation (4.3.1). ■

The Girsanov Theorem also allows statements concerning uniqueness in law. The following is a typical version, which is proved in [Karatzas and Shreve, 1991, Prop. 5.3.10, Cor 5.3.11].

Proposition 4.3.6 Let two weak solutions $((X_i,W_i),(\Omega_i,\mathcal{F}_i,\mathbb{P}_i),(\mathcal{F}^i_t))$, $i = 1,2$, of

$$dX(t) = b(X(t),t)\,dt + dW(t), \quad 0 \le t \le T,$$

with $b : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ measurable be given with the same initial distribution. If

$$\mathbb{P}_i\Big(\int_0^T |b(X_i(t),t)|^2\,dt < \infty\Big) = 1$$

holds for $i = 1,2$, then (X_1,W_1) and (X_2,W_2) have the same law under the respective probability measures. In particular, if b is uniformly bounded, then uniqueness in distribution holds.

5. The Markov properties

5.1 General facts about Markov processes

Let us fix the measurable space (state space) $(S,\mathcal{S})$ and the filtered probability space $(\Omega,\mathcal{F},\mathbb{P};(\mathcal{F}_t)_{t\geq 0})$ until further notice. We present certain notions and results concerning Markov processes without proof and refer e.g. to Kallenberg [2002] for further information. We specialise immediately to processes in continuous time and later on also to processes with continuous paths.

Definition 5.1.1 An S-valued stochastic process $(X(t), t\geq 0)$ is called *Markov process* if X is $(\mathcal{F}_t)$-adapted and satisfies

$$\forall 0\leq s\leq t, B\in\mathcal{S}: \ \mathbb{P}(X(t)\in B\,|\,\mathcal{F}_s)=\mathbb{P}(X(t)\in B\,|\,X(s)) \quad \mathbb{P}\text{-a.s.}$$

In the sequel we shall always suppose that regular conditional transition probabilities (Markov kernels) $\mu_{s,t}$ exist, that is for all $s\leq t$ the functions $\mu_{s,t}: S\times\mathcal{S}\to\mathbb{R}$ are measurable in the first component and probability measures in the second component and satisfy

$$\mu_{s,t}(X(s),B)=\mathbb{P}(X(t)\in B\,|\,X(s))=\mathbb{P}(X(t)\in B\,|\,\mathcal{F}_s) \quad \mathbb{P}\text{-a.s.} \tag{5.1.1}$$

Lemma 5.1.2 The Markov kernels $(\mu_{s,t})$ satisfy the Chapman-Kolmogorov equation

$$\mu_{s,u}(x,B)=\int_S \mu_{t,u}(y,B)\,\mu_{s,t}(x,dy) \quad \forall 0\leq s\leq t\leq u, x\in S, B\in\mathcal{S}.$$

Definition 5.1.3 Any family of regular conditional probabilities $(\mu_{s,t})_{s\leq t}$ satisfying the Chapman-Kolmogorov equation is called a semigroup of *Markov kernels*. The kernels (or the associated process) are called *time-homogeneous* if $\mu_{s,t}=\mu_{0,t-s}$ holds. In this case we just write μ_{t-s}.

Theorem 5.1.4 For any initial distribution ν on $(S,\mathcal{S})$ and any semigroup of Markov kernels $(\mu_{s,t})$ there exists a Markov process X such that $X(0)$ is ν-distributed and equation (5.1.1) is satisfied.
If S is a metric space with Borel σ-algebra $\mathcal{S}$ and if the process has a continuous version,

then the process can be constructed on the path space $\Omega = C(\mathbb{R}_+, S)$ with its Borel σ-algebra $\mathfrak{B}$ and canonical right-continuous filtration $\mathcal{F}_t = \bigcap_{s>t} \sigma(X(u), u \leq s)$, where $X(u,\omega) := \omega(u)$ are the coordinate projections. The probability measure obtained is called $\mathbb{P}_\nu$ and it holds

$$\mathbb{P}_\nu(A) = \int_S \mathbb{P}_x(A)\,\nu(dx), \quad A \in \mathfrak{B},$$

with $\mathbb{P}_x := \mathbb{P}_{\delta_x}$. ■

For the formal statement of the strong Markov property we introduce the shift operator ϑ_t that induces a left-shift on the function space Ω.

Definition 5.1.5 The *shift operator* ϑ_t on the canonical space Ω is given by $\vartheta_t : \Omega \to \Omega$, $\vartheta_t(\omega) = \omega(t + \bullet)$ for all $t \geq 0$.

Lemma 5.1.6

1. ϑ_t is measurable for all $t \geq 0$.
2. For $(\mathcal{F}_t)$-stopping times σ and τ the random time $\gamma := \sigma + \tau \circ \vartheta_\sigma$ is again an $(\mathcal{F}_t)$-stopping time.

Theorem 5.1.7 Let X be a time-homogeneous Markov process and let τ be an $(\mathcal{F}_t)$-stopping time with at most countably many values. Then we have for all $x \in S$

$$\mathbb{P}_x(X \circ \vartheta_\tau \in A \,|\, \mathcal{F}_\tau) = \mathbb{P}_{X(\tau)}(A) \quad \mathbb{P}_x\text{-a.s.} \quad \forall A \in \mathfrak{B}. \tag{5.1.2}$$

If X is the canonical process on the path space, then this is just an identity concerning the image measure under $\omega \mapsto \vartheta_{\tau(\omega)}(\omega)$: $\mathbb{P}_x(\bullet \,|\, \mathcal{F}_\tau) \circ (\vartheta_\tau)^{-1} = \mathbb{P}_{X(\tau)}$. ■

Definition 5.1.8 A process X satisfying (5.1.2) for any finite (or equivalently bounded) stopping time τ is called *strong Markov*.

R The strong Markov property entails the Markov property by setting $\tau = t$ and $A = \{X(s) \in B\}$ for some $B \in \mathcal{S}$ in (5.1.2).

5.2 The martingale problem

We specify now to the state space $S = \mathbb{R}^d$. As before we work on the path space $\Omega = C(\mathbb{R}_+, \mathbb{R}^d)$ with its Borel σ-algebra $\mathfrak{B}$.

Definition 5.2.1 A probability measure $\mathbb{P}$ on the path space $(\Omega, \mathfrak{B})$ is a solution of the *local martingale problem* for (b, σ) if

$$M^f(t) := f(X(t)) - f(X(0)) - \int_0^t A_s f(X(s))\,ds, \quad t \geq 0,$$

where

$$A_s f(x) := \frac{1}{2}\sum_{i,j=1}^{d} (\sigma\sigma^T(x,s))_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \langle b(x,s), \mathrm{grad}(f)(x)\rangle,$$

$b : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}^{d\times m}$ measurable, is a local martingale under $\mathbb{P}$ for all functions $f \in C_K^\infty(\mathbb{R}^d, \mathbb{R})$, that is, f has derivatives of all orders and compact support.

R If b and σ are bounded, then $\mathbb{P}$ even solves the martingale problem, for which M^f is required to be a proper martingale.

Theorem 5.2.2 The stochastic differential equation

$$dX(t) = b(X(t),t)\,dt + \sigma(X(t),t)\,dW(t), \quad t \geq 0$$

has a weak solution $((X,W),(\Omega,\mathfrak{A},\mathbb{P}),(\mathcal{F}_t))$ if and only if a solution to the local martingale problem (b,σ) exists. In this case the law $\mathbb{P}^X$ of X on the path space equals the solution of the local martingale problem. ■

Proof. For simplicity we only give the proof for the one-dimensional case, the multi-dimensional method of proof follows the same ideas.

1. Given a weak solution, Itô's rule yields for any $f \in C_K^\infty(\mathbb{R})$

$$\begin{aligned} df(X(t)) &= f'(X(t))\,dX(t) + \tfrac{1}{2}f''(X(t))\,d[X,X](t) \\ &= f'(X(t))\sigma(X(t),t)\,dW(t) + A_t f(X(t))\,dt. \end{aligned}$$

   Hence, M^f is a local martingale; just note that $\sigma(X(\bullet)) \in V^*$ is required for the weak solution and f' is bounded such that the stochastic integral is indeed well defined and a local martingale under $\mathbb{P}$. Of course, this remains true, when considered on the path space under the image measure $\mathbb{P}^X$.
2. Conversely, let $\mathbb{P}$ be a solution of the local martingale problem and consider functions $f_n \in C_K^\infty(\mathbb{R})$ with $f_n(x) = x$ for $|x| \leq n$. Then the standard stopping argument applied to M^{f_n} for $n \to \infty$ shows that

$$M(t) := X(t) - X(0) - \int_0^t b(X(s),s)\,ds, \quad t \geq 0,$$

   is a local martingale. Similarly approximating $g(x) = x^2$, we obtain that

$$N(t) := X^2(t) - X^2(0) - \int_0^t \big(\sigma^2(X(s),s) + b(X(s),s)2X(s)\big)\,ds, \quad t \geq 0,$$

   is a local martingale. By Itô's formula, $dX^2(t) = 2X(t)dX(t) + d[X,X](t)$ holds and shows

$$N(t) = \int_0^t 2X(s)\,dM(s) + [M,M](t) - \int_0^t \sigma^2(X(s),s)\,ds, \quad t \geq 0.$$

Therefore $[M,M](t) - \int_0^t \sigma^2(X(s),s)\,ds$ is a continuous local martingale of bounded variation. By [Revuz and Yor, 1999, Prop. IV.1.2] it must therefore vanish identically and $d[M,M](t) = \sigma^2(X(t),t)\,dt$ follows. By the representation theorem for continuous local martingales [Kallenberg, 2002, Thm. 18.12] there exists a Brownian motion W such that $M(t) = \int_0^t \sigma(X(s),s)\,dW(s)$ holds for all $t \geq 0$. Consequently (X,W) solves the stochastic differential equation.

■

R A stochastic differential equation has (in distribution) a unique weak solution if and only if the corresponding local martingale problem is uniquely solvable, given some initial distribution.

5.3 The strong Markov property

We immediately start with the main result that solutions of stochastic differential equations are under mild conditions strong Markov processes. This entails that the solutions are diffusion processes in the sense of Feller [1971].

Theorem 5.3.1 Let $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be time-homogeneous measurable coefficients such that the local martingale problem for (b,σ) has a unique solution $\mathbb{P}_x$ for all initial distributions δ_x, $x \in \mathbb{R}^d$. Then the family $(\mathbb{P}_x)$ satisfies the strong Markov property.

Proof. In order to state the strong Markov property we need that $(\mathbb{P}_x)_{x\in\mathbb{R}^d}$ are Markov kernels. Theorem 21.10 of Kallenberg [2002] shows by abstract arguments that $x \mapsto \mathbb{P}_x(B)$ is measurable for all $B \in \mathfrak{B}$.

We thus have to show

$$\mathbb{P}_x(X \circ \vartheta_\tau \in B \,|\, \mathcal{F}_\tau) = \mathbb{P}_{X(\tau)}(B) \quad \mathbb{P}_x\text{-a.s. } \forall\, B \in \mathfrak{B},\ \text{bounded stopping time } \tau.$$

By the unique solvability of the martingale problem it suffices to show that the *random* probability measure $\mathbb{Q}_\tau := \mathbb{P}_x((\vartheta_\tau)^{-1}\bullet \,|\, \mathcal{F}_\tau)$ solves $\mathbb{P}_x$-almost surely the martingale problem for (b,σ) with initial distribution $\delta_{X(\tau)}$. Concerning the initial distribution we find for any Borel set $A \subseteq \mathbb{R}^d$ by the stopping time property of τ

$$\begin{aligned}
\mathbb{P}_x((\vartheta_\tau)^{-1}\{\omega' \,|\, \omega'(0) \in A\} \,|\, \mathcal{F}_\tau)(\omega) &= \mathbb{P}_x(\{\omega' \,|\, \omega'(\tau(\omega')) \in A\} \,|\, \mathcal{F}_\tau)(\omega)\\
&= \mathbf{1}_A(\omega(\tau(\omega)))\\
&= \mathbf{1}_A(X(\tau(\omega),\omega))\\
&= \mathbb{P}_{X(\tau(\omega),\omega)}(\{\omega' \,|\, \omega'(0) \in A\}).
\end{aligned}$$

It remains to prove the local martingale property of M^f under $\mathbb{Q}_\tau$, that is the martingale property of $M^{f,n}(t) := M^f(t \wedge \tau_n)$ with $\tau_n := \inf\{t \geq 0 \,|\, \|M^f(t)\| \geq n\}$. By its very definition $M^f(t)$ is always $\mathcal{F}_t$-measurable, so we prove that $\mathbb{P}_x$-almost surely

$$\int_F M^{f,n}(t,\omega')\,\mathbb{Q}_\tau(d\omega') = \int_F M^{f,n}(s,\omega')\,\mathbb{Q}_\tau(d\omega') \quad \forall F \in \mathcal{F}_s,\ s \leq t.$$

By the separability of Ω and the continuity of $M^{f,n}$ it suffices to prove this identity for countably many F, s and t [Kallenberg, 2002, Thm. 21.11]. Consequently, we need not worry about $\mathbb{P}_x$-null sets. We obtain

$$\begin{aligned}\int_F M^{f,n}(t,\omega')\,\mathbb{Q}_\tau(d\omega') &= \int \mathbf{1}_F(\vartheta_\tau(\omega''))M^{f,n}(t,\vartheta_\tau(\omega''))\,\mathbb{P}_x(d\omega''\,|\,\mathcal{F}_\tau)\\ &= \mathbb{E}_x[\mathbf{1}_{(\vartheta_\tau)^{-1}F}M^{f,n}(t,\vartheta_\tau)\,|\,\mathcal{F}_\tau].\end{aligned}$$

Because of $M^{f,n}(t,\vartheta_\tau) = M^f((t+\tau)\wedge\sigma_n)$ with $\sigma_n := \tau_n\circ\vartheta_\tau+\tau$, which is by Lemma 5.1.6 a stopping time, the process $M^{f,n}(t,\vartheta_\tau)$ is a martingale under $\mathbb{P}_x$ adapted to $(\mathcal{F}_{t+\tau})_{t\ge 0}$. Since $(\vartheta_\tau)^{-1}F$ is an element of $\mathcal{F}_{s+\tau}$, we conclude by optional stopping that $\mathbb{P}_x$-almost surely

$$\begin{aligned}\int_F M^{f,n}(t,\omega')\,\mathbb{Q}_\tau(d\omega') &= \mathbb{E}_x[\mathbf{1}_{(\vartheta_\tau)^{-1}F}\,\mathbb{E}_x[M^{f,n}(t,\vartheta_\tau)\,|\,\mathcal{F}_{s+\tau}]\,|\,\mathcal{F}_\tau]\\ &= \mathbb{E}_x[\mathbf{1}_{(\vartheta_\tau)^{-1}F}M^{f,n}(s+\tau)\,|\,\mathcal{F}_\tau]\\ &= \int_F M^{f,n}(s,\omega')\,\mathbb{Q}_\tau(d\omega').\end{aligned}$$

Consequently, we have shown that with $\mathbb{P}_x$-probability one $\mathbb{Q}_\tau$ solves the martingale problem with initial distribution $X(\tau)$ and therefore equals $\mathbb{P}_{X(\tau)}$. ∎

▪ **Example 5.1** A famous application is the reflection principle for Brownian motion W. By the strong Markov property, for any finite stopping time τ the process $(W(t+\tau)-W(\tau), t\ge 0)$ is again a Brownian motion independent of $\mathcal{F}_\tau$, cf. Exercise 6. ▪

5.4 The infinitesimal generator

We first gather some facts concerning Markov transition operators and their semigroup property, see Kallenberg [2002] or Revuz and Yor [1999].

Lemma 5.4.1 Given a family $(\mu_t)_{t\ge 0}$ of time-homogeneous Markov kernels, the operators

$$T_t f(x) := \int f(y)\mu_t(x,dy),\; f: S\to\mathbb{R} \text{ bounded, measurable,}$$

form a semigroup, that is $T_t\circ T_s = T_{t+s}$ holds for all $t,s\ge 0$.

Proof. Use the Chapman-Kolmogorov equation. ∎

We now specialise to the state space $S=\mathbb{R}^d$ with its Borel σ-algebra.

Definition 5.4.2 If the operators $(T_t)_{t\ge 0}$ satisfy (a) $T_t f\in C_0(\mathbb{R}^d)$ for all $f\in C_0(\mathbb{R}^d)$ and (b) $\lim_{h\to 0} T_h f(x) = f(x)$ for all $f\in C_0(\mathbb{R}^d)$, $x\in\mathbb{R}^d$, then (T_t) is called a *Feller semigroup*.

For rigorous proofs of the following relations, we refer to Thm. 17.6 and Lemma 17.5 of Kallenberg [2002]. A Feller semigroup $(T_t)_{t\ge 0}$ is a strongly continuous operator semigroup on $C_0(\mathbb{R}^d)$, that is $\lim_{h\to 0} T_h f = f$ holds in supremum norm. It is uniquely determined by its (infinitesimal) generator $A:\mathcal{D}(A)\subset C_0(\mathbb{R}^d)\to\mathbb{R}^d$ with

$$Af := \lim_{h\to 0}\frac{T_h f - f}{h},\quad \mathcal{D}(A) := \{f\in C_0(\mathbb{R}^d)\,|\,\lim_{h\to 0}\tfrac{T_h f - f}{h}\text{ exists}\}.$$

Moreover, the semigroup uniquely defines the Markov kernels and thus the distribution of the associated Markov process (which is called *Feller process*). We have for all $f \in \mathcal{D}(A)$

$$\tfrac{d}{dt}T_t f = AT_t f = T_t Af.$$

If b and σ are bounded and satisfy the conditions of Theorem 5.3.1, then the Markov kernels $(\mathbb{P}_x)_{x\in\mathbb{R}^d}$ solving the martingale problem for (b,σ) give rise to a Feller semigroup (T_t). Any function $f \in C_0^2(\mathbb{R}^d)$ lies in $\mathcal{D}(A)$ and fulfills

$$Af(x) = \frac{1}{2}\sum_{i,j=1}^{d}(\sigma\sigma^T(x))_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \langle b(x), \operatorname{grad}(f)(x)\rangle.$$

We shall even prove a stronger result under less restrictive conditions, which turns out to be a very powerful tool in calculating certain distributions for the solution processes.

Theorem 5.4.3 — Dynkin's formula. Assume that b and σ are measurable, locally bounded and such that the SDE (3.1.1) with time-homogeneous coefficients has a (in distribution) unique weak solution. Then for all $x \in \mathbb{R}^d$, $f \in C_K^2(\mathbb{R}^d)$ and all bounded stopping times τ, we have

$$\mathbb{E}_x[f(X(\tau))] = f(x) + \mathbb{E}_x\Big[\int_0^\tau Af(X(s))\,ds\Big].$$

Proof. By Theorem 5.2.2 the process M^f is a local martingale under $\mathbb{P}_x$. By the compact support of f and the local boundedness of b and σ we infer that $M^f(t)$ is uniformly bounded and therefore M^f is a martingale. Then the optional stopping result $\mathbb{E}[M^f(\tau)] = \mathbb{E}[M^f(0)] = 0$ yields Dynkin's formula. ■

Example 5.2

1. Let W be an m-dimensional Brownian motion starting in some point a and $\tau_R := \inf\{t \geq 0 \,|\, \|W(t)\| \geq R\}$. Then $\mathbb{E}_a[\tau_R] = (R^2 - \|a\|^2)/m$ holds for $\|a\| < R$. To infer this from Dynkin's formula put $f(x) = \|x\|^2$ for $\|x\| \leq R$ and extend f outside of the ball such that $f \in C^2(\mathbb{R})$ with compact support. Then $Af(x) = m$ for $\|x\| \leq R$ and therefore Dynkin's formula yields $\mathbb{E}_a[f(W(\tau_R \wedge n))] = f(a) + m\mathbb{E}_a[\tau_R \wedge n]$. By montone convergence,

$$\mathbb{E}_a[\tau_R] = \lim_{n\to\infty}\mathbb{E}_a[\tau_R \wedge n] = \lim_{n\to\infty}(\mathbb{E}_a[\|W(\tau_R \wedge n)\|^2] - \|a\|^2)/m$$

holds and we can conclude by dominated convergence ($\|W(\tau_R \wedge n)\| \leq R$).

2. Consider the one-dimensional stochastic differential equation

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

Suppose a weak solution exists for some initial value $X(0)$ with $\mathbb{E}[X(0)^2] < \infty$ and that $\sigma^2(x) + 2xb(x) \leq C(1+x^2)$ holds. Then $\mathbb{E}[X(t)^2] \leq (\mathbb{E}[X(0)^2]+1)e^{Ct} - 1$

follows. To prove this, use the same f and put $\kappa_t := \tau_R \wedge t$ with τ_R from above for all $t \geq 0$ such that by Dynkin's formula

$$\begin{aligned}\mathbb{E}_x[X(\kappa_t)^2] &= x^2 + \mathbb{E}_x\Big[\int_0^{\kappa_t}\big(\sigma^2(X(s)) + 2b(X(s))X(s)\big)\,ds\Big]\\ &\leq x^2 + \int_0^t C(1+X(s\wedge\tau_R)^2)\,ds\,.\end{aligned}$$

By Gronwall' s lemma, we obtain $\mathbb{E}_x[1+X(\kappa_t)^2] \leq (x^2+1)e^{Ct}$. Since this is valid for any $R>0$ we get $\mathbb{E}_x[X(t)^2] \leq (x^2+1)e^{Ct}-1$ and averaging over the initial condition yields $\mathbb{E}[X(t)^2] \leq (\mathbb{E}[X(0)^2]+1)e^{Ct}-1$. Note that this kind of approach was already used in Theorem 3.3.3 and improves significantly on the moment estimate of Theorem 3.3.1.

3. For the solution process X of a one-dimensional SDE as before we consider the stopping time $\tau := \inf\{t \geq 0 \,|\, X(t) = 0\}$. We want to decide whether $\mathbb{E}_a[\tau]$ is finite or infinite for $a>0$. For this set $\tau_R := \tau \wedge \inf\{t \geq 0 \,|\, X(t) \geq R\}$, $R>a$, and consider a function $f \in C^2(\mathbb{R})$ with compact support, $f(0)=0$ and solving $Af(x)=1$ for $x \in [0,R]$. Then Dynkin's formula yields

$$\mathbb{E}_a[f(X(\tau_R\wedge n))] = f(a) + \mathbb{E}_a[\tau_R \wedge n].$$

For a similar function g with $Ag=0$ and $g(0)=0$ we obtain $\mathbb{E}_a[g(X(\tau_R\wedge n))] = g(a)$. Hence,

$$\begin{aligned}\mathbb{E}_a[\tau_R\wedge n] &= \mathbb{E}_a[f(X(\tau_R\wedge n))] - f(a)\\ &= \mathbb{P}_a(X(\tau_R\wedge n)=R)f(R) + \mathbb{E}_a[f(X(n))\mathbf{1}_{\{\tau_R>n\}}] - f(a)\\ &= \Big(g(a) - \mathbb{E}_a[g(X(n))\mathbf{1}_{\{\tau_R>n\}}]\Big)\frac{f(R)}{g(R)} + \mathbb{E}_a[f(X(n))\mathbf{1}_{\{\tau_R>n\}}] - f(a)\end{aligned}$$

follows. Using the uniform boundedness of f and g we infer by montone and dominated convergence for $n\to\infty$

$$\mathbb{E}_a[\tau_R] = g(a)\frac{f(R)}{g(R)} - f(a).$$

Monotone convergence for $R\to\infty$ thus gives $\mathbb{E}_a[\tau]<\infty$ if and only if $\lim_{R\to\infty}\frac{f(R)}{g(R)}$ is finite. The functions f and g can be determined in full generality, but we restrict ourselves to the case of vanishing drift $b(x)=0$ and strictly positive diffusion coefficient $\inf_{0\leq y\leq x}\sigma(y)>0$ for all $x>0$. Then

$$f(x) = \int_0^x\int_0^y \frac{2}{\sigma^2(z)}\,dz\,dy \text{ and } g(x) = x$$

will do. Since $f(x)\to\infty$, $g(x)\to\infty$ hold for $x\to\infty$, L'Hopital's rule gives

$$\lim_{R\to\infty}\frac{f(R)}{g(R)} = \lim_{R\to\infty}\frac{f'(R)}{g'(R)} = \int_0^\infty \frac{2}{\sigma^2(z)}\,dz.$$

We conclude that the solution of $dX(t) = \sigma(X(t))\,dW(t)$ with $X(0)=a$ satisfies $\mathbb{E}_a[\tau]<\infty$ if and only if σ^{-2} is integrable. For constant σ we obtain a multiple of

Brownian motion which satisfies $\mathbb{E}_a[\tau] = \infty$. For $\sigma(x) = x + \varepsilon$, $\varepsilon > 0$, $\mathbb{E}_a[\tau] < \infty$ holds, but in the limit $\varepsilon \to 0$ the expectation tends to infinity. This can be understood when observing that a solution of $dX(t) = (X(t) + \varepsilon)dW(t)$ is given by the translated geometric Brownian motion $X(t) = \exp(W(t) - \frac{t}{2}) - \varepsilon$, which tends to $-\varepsilon$ almost surely, but never reaches the value $-\varepsilon$. Concerning the behaviour of $\sigma(x)$ for $x \to \infty$ we note that $\mathbb{E}_a[\tau]$ is finite as soon as $\sigma(x)$ grows at least like x^α for some $\alpha > \frac{1}{2}$ such that the rapid fluctuations of X for large x make excursions towards zero more likely.

■

5.5 The Kolmogorov equations

The main object one is usually interested in to calculate for the solution process X of an SDE is the transition probability $\mathbb{P}(X(t) \in B \,|\, X(s) = x)$ for $t \geq s \geq 0$ and any Borel set B. A concise description is possible, if a transition density $p(x,y;t)$ exists satisfying

$$\mathbb{P}(X(t) \in B \,|\, X(s) = x) = \int_B p(x,y;t-s)\,dy.$$

Here we shall present analytical tools to determine this transition density if it exists. The proof of its existence usually either relies completely on analytical results or on Malliavin calculus, both being beyond our scope.

Lemma 5.5.1 Assume that b and σ are continuous and such that the SDE (3.1.1) has a (in distribution) unique weak solution for any deterministic initial value. For any $f \in C_K^2(\mathbb{R}^d)$ set $u(x,t) := \mathbb{E}_x[f(X(t))]$. Then u is a solution of the parabolic partial differential equation

$$\frac{\partial u}{\partial t}(x,t) = (Au(\bullet,t))(x), \quad \forall x \in \mathbb{R}^d, t \geq 0, \text{ with } u(x,0) = f(x) \quad \forall x \in \mathbb{R}^d.$$

Proof. Dynkin's formula for $\tau = t$ yields by the Fubini-Tonelli theorem

$$u(x,t) = f(x) + \int_0^t \mathbb{E}_x[Af(X(s))]\,ds \quad \forall x \in \mathbb{R}^d, t \geq 0.$$

Since the coefficients b and σ are continuous, the integrand is continuous and u is continuously differentiable with respect to t satisfying $\frac{\partial u}{\partial t}(x,t) = \mathbb{E}_x[Af(X(t))]$. On the other hand we obtain by the Markov property for $t,h > 0$

$$\mathbb{E}_x[u(X(h),t)] = \mathbb{E}_x[\mathbb{E}_{X(h)}[f(X(t))]] = \mathbb{E}_x[f(X(t+h))] = u(x,t+h).$$

For fixed $t > 0$ we infer that the left hand side of

$$\frac{u(x,t+h) - u(x,t)}{h} = \frac{\mathbb{E}_x[u(X(h),t)] - u(x,t)}{h}$$

converges for $h \to 0$ to $\frac{\partial u}{\partial t}$ and therefore also the right-hand side. Therefore u lies in the domain $\mathcal{D}(A)$ and the assertion follows. ■

Corollary 5.5.2 If the transition density $p(x,y;t)$ exists, is twice continuously differentiable with respect to x and continuously differentiable with respect to t, then $p(x,y;t)$ solves for all $y \in \mathbb{R}^d$ the backward Kolmogorov equation

$$\frac{\partial u}{\partial t}(x,t) = (Au(\bullet,t))(x), \quad \forall x \in \mathbb{R}^d, t \geq 0, \text{ with } u(x,0) = \delta_y(x).$$

In other words, for fixed y the transition density is the fundamental solution of this parabolic partial differential equation.

Proof. Writing the identity in the preceding lemma in terms of p, we obtain for any $f \in C_K^2(\mathbb{R}^d)$

$$\frac{\partial}{\partial t} \int_{\mathbb{R}^d} f(y) p(x,y;t)\, dy = A\Big(\int_{\mathbb{R}^d} f(y) p(x,y;t)\, dy\Big).$$

By the compact support of f and the smoothness properties of p, we may interchange integration and differentiation on both sides. From $\int (\frac{\partial}{\partial t} - A) p(x,y;t) f(y) dy = 0$ for any test function f we then conclude by a continuity argument. ■

Corollary 5.5.3 If the transition density $p(x,y;t)$ exists, is twice continuously differentiable with respect to y and continuously differentiable with respect to t, then $p(x,y;t)$ solves for all $x \in \mathbb{R}^d$ the forward Kolmogorov equation

$$\frac{\partial u}{\partial t}(y,t) = (A^* u(\bullet,t))(y), \quad \forall y \in \mathbb{R}^d, t \geq 0, \text{ with } u(y,0) = \delta_x(y),$$

where

$$A^* f(y) = \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2}{\partial y_i y_j} \Big((\sigma \sigma^T(y))_{ij} f(y) \Big) - \sum_{i=1}^d \frac{\partial}{\partial y_i} \Big(b_i(y) f(y) \Big)$$

is the formal adjoint of A. Hence, for fixed x the transition density is the fundamental solution of the parabolic partial differential equation with the adjoint operator.

Proof. Let us evaluate $\mathbb{E}_x[Af(X(t))]$ for any $f \in C_K^2(\mathbb{R}^d)$ in two different ways. First, we obtain by definition

$$\mathbb{E}_x[Af(X(t))] = \int Af(y) p(x,y;t)\, dy = \int f(y) (A^* p(x,\bullet;t))(y)\, dy.$$

On the other hand, by dominated convergence and by Dynkin's formula we find

$$\int f(y) \frac{\partial}{\partial t} p(x,y;t)\, dy = \frac{\partial}{\partial t} \mathbb{E}_x[f(X(t))] = \mathbb{E}_x[Af(X(t))].$$

We conclude again by testing this identity with all $f \in C_K^2(\mathbb{R}^d)$. ■

R The preceding results are in a sense not very satisfactory because we had to postulate properties of the unknown transition density in order to derive a determining equation. Karatzas and Shreve [1991] state on page 368 sufficient conditions on the coefficients b and σ, obtained from the analysis of the partial differential equations, under which the transition density is the unique classical solution of the forward and backward Kolmogorov equation, respectively. Main hypotheses are ellipticity of the diffusion coefficient and boundedness of both coefficients together with certain Hölder-continuity requirements. In the case of the forward equation in addition the first two derivatives of σ and the first derivative of b have to have these properties, which is intuitively explained by the form of the adjoint A^*.

Example 5.3 We have seen that a solution of the scalar Ornstein-Uhlenbeck process

$$dX(t) = \alpha X(t)\,dt + \sigma\,dW(t), \quad t \geq 0,$$

is given by $X(t) = X(0)e^{\alpha t} + \int_0^t e^{\alpha(t-s)}\sigma\,dW(s)$. Hence, the transition density is given by the normal density

$$p(x,y;t) = \frac{1}{\sqrt{2\pi\sigma^2(2\alpha)^{-1}(e^{2\alpha t}-1)}} \exp\Big(-(y-xe^{\alpha t})^2/(\sigma^2\alpha^{-1}(e^{2\alpha t}-1))\Big).$$

It can be checked that p solves the Kolmogorov equations

$$\frac{\partial u}{\partial t}(x,t) = \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2}(x,t) + \alpha\frac{\partial u}{\partial x}(x,t) \qquad \text{and} \qquad \frac{\partial u}{\partial t}(y,t) = \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial y^2}(y,t) - \alpha\frac{\partial u}{\partial y}(y,t).$$

For $\alpha = 0$ and $\sigma = 1$ we obtain the Brownian motion transition density $p(x,y;t) = (2\pi t)^{-1/2}\exp(-(y-x)^2/(2t))$ which is the fundamental solution of the classical heat equation $\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}$ in both variables x and y. ■

5.6 The Dirichlet problem and Brownian motion

Let U be an open and bounded subset of $\mathbb{R}^d$ with ∂U its boundary. Denote by $\Phi : \partial U \to \mathbb{R}$ a continuous function on the boundary. A function $u : \overline{U} \to \mathbb{R}$, which is twice continuously differentiable on U and continuous on the closure $\overline{U}$ is a solution to the Dirichlet problem with boundary condition Φ, if

$$\Delta u(x) = 0 \;\; \forall x \in U \;,\; u(x) = \Phi(x) \;\; \forall x \in \partial U\,. \tag{5.6.1}$$

$\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$ is the Laplacian. A function with $\Delta u(x) = 0$ for all $x \in U$ is called harmonic on U. The problem was formulated by Gauss who thought that there is always a solution, which is not true. However, under suitable conditions on U there is a solution and in this case it can be represented and simulated exploiting a Brownian motion. The deep connection between Brownian motion and harmonic functions is understood applying Itô's formula. In the sequel, denote by $B^x(t) = x + B(t)$ with a standard Brownian motion B, such that B^x is a Brownian motion started in $x \in \mathbb{R}^d$.

Lemma 5.6.1 Let $u : \mathbb{R}^d \to \mathbb{R}$ be harmonic on U and

$$T = \inf\{t \in \mathbb{R}_+ \,|\, B^x(t) \notin U\}\,.$$

$u(B^x(t))$ is a local martingale on $[0,T)$ and

$$u(B^x(t)) = u(x) + \int_0^t \sum_{i=1}^d \frac{\partial}{\partial x_i} u(B(s))\, dB_s^i\,.$$

Proof. Applying Itô's formula yields

$$u(B^x(t)) - u(x) = \int_0^t \sum_{i=1}^d \frac{\partial}{\partial x_i} u(B(s))\, dB_s^i + \frac{1}{2}\int_0^t \Delta u(B(s))\, ds\,.$$

When u is harmonic, the last addend vanishes. ∎

Theorem 5.6.2 Let u be a solution of the Dirichlet problem on U with boundary condition Φ. Let $T = \inf\{t \in \mathbb{R}_+ \mid B^x(t) \notin U\}$. Then, for all $x \in U$

$$u(x) = \mathbb{E}_x\left[\Phi(B^x(T))\right],$$

where $\mathbb{E}_x$ refers to the expectation with respect to the starting value x. In particular, the boundary uniquely determines the solution. ∎

Proof. First, we show that T is almost surely finite and has even moments of all orders, $\sup_{x\in U} \mathbb{E}_x[T^p] < \infty$ for all $p < \infty$. We can write

$$\mathbb{E}_x[T^p] = \int_0^\infty \mathbb{P}_x(T^p \geq t)\, dt = \int_0^\infty p s^{p-1}\, \mathbb{P}_x(T \geq s)\, ds\,.$$

If $\mathbb{P}_x(T \geq n) \leq p^n$ with some $p < 1$ and for all $n \in \mathbb{N}$, the above expectation exists. This can be proved by induction. First

$$\begin{aligned}\mathbb{P}_x(T < 1) &\geq \mathbb{P}_x\left(|B^x(1) - x| > \sup_{x,y\in U} |x-y|\right)\\ &= \mathbb{P}_0\left(|B(1)| > \sup_{x,y\in U} |x-y|\right) = 1 - q \in (0,1)\,.\end{aligned}$$

With the induction step

$$\begin{aligned}\mathbb{P}_x(T \geq n) &\geq (2\pi)^{-d/2} \int_U e^{-|x-y|^2/2}\, \mathbb{P}_y(T > n-1)\, dy\\ &\leq p^{n-1}\, \mathbb{P}_x(B_1 \in U) \leq p^n\,,\end{aligned}$$

the claim follows.
Now, apply Itô's formula to see that

$$u(B^x(t \wedge T)) = u(B^x(0)) + \int_0^{t\wedge T} \nabla u(B(s))\, dB_s\,.$$

The integrand $\nabla u(B(s))$ is bounded and we conclude that

$$\mathbb{E}_x\left[u(B(t \wedge T))\right] = u(x)\,.$$

Since T is almost surely finite, with dominated convergence it follows that

$$\mathbb{E}_x\left[u(B(T))\right] = u(x)\,.$$

Since $B(T) \in \partial U$, it holds that $u(B(T)) = \Phi(B(T))$. ∎

R This approach is not very useful to show existence of solutions of the Dirichlet problem, because if defined as above, it will be difficult to show differentiability of u. However, if differentiability is assumed, then u solves the Dirichlet problem.
The solution can be simulated using this formula, by running independent Brownian motions starting in x until they hit the boundary of U. Then, one sets $u(x)$ to the average of the values of Φ on the hitting points.

We apply Theorem 5.6.2 to prove the next useful Lemma about hitting times of Brownian motion.

Lemma 5.6.3 Let $a < x < b$ and $T = \inf\{t \in \mathbb{R}_+ \mid B^x(t) \notin (a,b)\}$ with a one-dimensional Brownian motion B^x started in $x \in \mathbb{R}$. We have that

$$\mathbb{P}_x\left(B^x(T) = a\right) = \frac{b-x}{b-a}, \mathbb{P}_x\left(B^x(T) = b\right) = \frac{x-a}{b-a}. \tag{5.6.2}$$

Proof. With $u(x) = (b-a)^{-1}(b-x)$, $u''(x) = 0 \in U = (a,b)$ and u continuous, the solution of the Dirichlet problem with boundary condition $\Phi(a) = 1$ and $\Phi(b) = 0$ satisfies

$$u(x) = \mathbb{E}_x\left[\Phi(B^x(T))\right] = \mathbb{P}_x(B^x(T) = a).$$

Analogous reasoning for $B^x(T) = b$ proves the Lemma. ■

Corollary 5.6.4 Let $T_y = \inf\{t \in \mathbb{R}_+ \mid B^x(t) = y\}$. It holds that $\mathbb{P}_y(T_x < \infty) = 1$ for all $x, y \in \mathbb{R}$.

Proof. Without loss of generality we can assume $x = 0$ and $y > 0$. With the preceding lemma, we obtain

$$\mathbb{P}_y\left(T_0 < \infty\right) \geq \lim_{N\to\infty} \mathbb{P}_y\left(T_0 < T_{Ny}\right) = \lim_{N\to\infty} \frac{N-1}{N} = 1.$$

■

Based upon the preceding results, we can prove the following important property of the one-dimensional Brownian motion.

Theorem 5.6.5 The one-dimensional standard Brownian motion is recurrent, that is, for all $x \in \mathbb{R}$ there exists a sequence of random times $\tau_n \to \infty$ with $B(\tau_n) = x$. That means each point x is visited infinitely often. ■

Proof. The preceding Corollary suffices to establish that B visits almost surely each point. We use the strong Markov property. Thereby, $\left(B(T_x+t)\right)_{t\geq 0}$ for $T_x = \inf\{t \in \mathbb{R}_+ \mid B(t) = x\}$ is again a Brownian motion. We have that $\mathbb{P}(T_x < 1) > 0$ and can use Borel-Cantelli to conclude the result, since we have after each visit an independent event for the Brownian motion $\left(B(T_x+t)\right)_{t\geq 0}$. ■

One can prove that for $d = 2$ each open set around any point is visited infinitely often. For dimension $d \geq 3$ instead, Brownian motion is transient. This can be proved based on the same ingredients, but also in various ways. We refer to Example 7.4.20 of Øksendal [1998] for a proof using Dynkin's formula.

5.7 The Feynman-Kac formula

We continue to study partial differential equations exploiting probabilistic structures. We consider the heat equation with a dissipation rate $c : (0,\infty) \times \mathbb{R}^d \to \mathbb{R}$. The equation describes the temperature in space and time of a system under cooling or heating. Thereto, consider the partial differential equation

$$\frac{\partial}{\partial t} u(t,x) = \frac{1}{2}\Delta u(t,x) - c(t,x)u(t,x) \ , \ u(0,x) = f(x) \, . \tag{5.7.1}$$

Lemma 5.7.1 If u is a solution of (5.7.1),

$$M_s = u(t-s, B(s)) \exp\Big(- \int_0^s c(t-r, B(r))\, dr \Big)$$

with a Brownian motion B is a local martingale on $[0,t)$.

Proof. We apply Itô's formula.

$$\begin{aligned}
&u(t-s, B(s)) \exp\Big(- \int_0^s c(t-r, B(r))\, dr \Big) - u(t, B(0)) \\
&\quad = \int_0^s \exp\Big(- \int_0^r c(t-u, B(u))\, du \Big) \Big(- \frac{\partial}{\partial r} u(t-r, B(r)) - u(t-r, B(r)) c(t-r, B(r)) \Big)\, dr \\
&\qquad + \sum_{i=1}^d \int_0^s \exp\Big(- \int_0^r c(t-u, B(u))\, du \Big) \frac{\partial}{\partial x_i} u(t-r, B(r))\, dB_r^i \\
&\qquad + \frac{1}{2} \int_0^s \exp\Big(- \int_0^r c(t-u, B(u))\, du \Big) \Delta u(t-r, B(r))\, dr \, ,
\end{aligned}$$

where we use that $[B^i, B^i](r) = r$ and $[B^i, B^j](r) = 0$ for $i \neq j$. Since $\Delta u/2 - cu - \frac{\partial}{\partial t} u = 0$, the claim follows. ■

> **Theorem 5.7.2 — Feynman-Kac.** Let c be a bounded dissipation rate and u a solution of (5.7.1) which is bounded on all sets $[0,t) \times \mathbb{R}^d$. Then, u satisfies
>
> $$u(t,x) = \mathbb{E}_x \Big[f(B(t)) \exp\Big(- \int_0^t c(t-u, B(u))\, du \Big) \Big] \tag{5.7.2}$$
>
> and is uniquely determined by (5.7.2).

Proof. Under the assumptions, M is a bounded martingale on $[0,t)$. By

$$M_t = f(B(t)) \exp\Big(- \int_0^t c(t-u, B(u))\, du \Big) = \lim_{s \to t, s < t} M_s$$

we derive that

$$\mathbb{E}_x \Big[f(B(t)) \exp\Big(- \int_0^t c(t-u, B(u))\, du \Big) \Big] = \mathbb{E}_x[M_t] = \mathbb{E}_x[M_0] = u(t,x) \, .$$

■

For a Feller process $(X(t))_{t \in \mathbb{R}_+}$ with generator A, the formula generalizes in the same way substituting $X(t)$ for $B(t)$ for a solution u of

$$\frac{\partial u}{\partial t} = Au - cu \ , \ u(0,x) = f(x) \, .$$

This is proved as Theorem 8.2.1 in Øksendal [1998].

III

Applications

6. Applications in Mathematical Finance

6.1 Stochastic market model

We denote by $S_t^j, j = 1,\dots,d, t \in [0,T]$, d stock price processes. One (usually risk-less) financial asset S_t^0 is used as a benchmark such that the other prices are expressed relative to this so-called numéraire. The discounted price processes are $\tilde{S}_t^j = S_t^j / S_t^0$, where we assume $S_t^0 > 0$ for all $t \geq 0$.

Definition 6.1.1 A (time-continuous) *market* (model) is a $(d+1)$-dimensional stochastic process $(S_t)_{t\in[0,T]}$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, where $(\mathcal{F}_t)$ satisfies the usual conditions. It holds that $\inf_{0\le t\le T} S_t^0 > 0$ $\mathbb{P}$-almost surely. The market is normalized if $S_t^0 \equiv 1$. A *trading strategy* or portfolio is a predictable $(d+1)$-dimensional process $(\vartheta_t), \vartheta^j \in \mathbb{L}, 0 \le j \le d$. The value (process) of a portfolio (ϑ_t) is given by

$$V_t(\vartheta) = \langle \vartheta_t, S_t \rangle = \sum_{j=0}^{d} \vartheta_t^j S_t^j .$$

The portfolio $\vartheta = (\vartheta(t))_{t\in[0,T]}$ is called *self-financing* if for all $t \in [0,T]$:

$$V_t(\vartheta) = V_0(\vartheta) + \int_0^t \sum_{j=0}^{d} \vartheta_s^j \, dS_s^j .$$

Restricting to continuous processes for price models, we only require that (ϑ_t) is adapted.

Lemma 6.1.2 Given that ϑ is self-financing, we have that

$$d\tilde{V}(\vartheta) = \sum_{j=1}^{d} \vartheta_-^j \, d\tilde{S}^j = \sum_{j=1}^{d} \vartheta^j \, d\tilde{S}^j .$$

Hence ϑ is also self-financing for the normalized market.

Proof. We apply the partial integration formula from Corollary 2.7.3 :

$$\begin{aligned} d\tilde{V}(\vartheta) &= d\big(S_0^{-1}V(\vartheta)\big) \\ &= S_{0-}^{-1}\,dV(\vartheta) + V(\vartheta)_-\,dS_0^{-1} + d[V(\vartheta),S_0^{-1}] \\ &= S_{0-}^{-1}\,\langle\vartheta,dS\rangle + \langle\vartheta,S_-\rangle\,dS_0^{-1} + \langle\vartheta,d[S,S_0^{-1}]\rangle \\ &= \langle\vartheta,S_{0-}^{-1}\,dS + S_-\,dS_0^{-1} + d[S,S_0^{-1}]\rangle \\ &= \langle\vartheta,d\big(S_0^{-1}S\big)\rangle = \langle\vartheta,d\tilde{S}\rangle\,. \end{aligned}$$

■

Definition 6.1.3 The trading strategy ϑ is called *admissible*, when $\tilde{V}_t(\vartheta)$ is bounded from below, $\tilde{V}_t(\vartheta) \geq -K, 0 \leq t \leq T$, $\mathbb{P}$-almost surely. A self-financing trading strategy ϑ is called an *arbitrage*, if it is admissible and

$$V_0(\vartheta) \leq 0\,, V_T(\vartheta) \geq 0 \;\mathbb{P}-a.s.\,, \mathbb{P}\big(V_T(\vartheta) > 0\big) > 0\,.$$

Theorem 6.1.4 — Fundamental theorem of asset pricing. If there exists an equivalent measure $\mathbb{Q} \sim \mathbb{P}$, such that $\tilde{S}$ is a square-integrable local martingale with respect to $\mathbb{Q}$, then the market has no arbitrage. ■

Proof. Apparently, $\tilde{V}_t(\vartheta) = \tilde{V}_0(\vartheta) + \int_0^t \sum_{j=1}^d \vartheta_s^j\, d\tilde{S}_s^j$ is a local martingale. Since $\tilde{V}_t(\vartheta) \geq -K$, it is a supermartingale. Hence,

$$\mathbb{E}_{\mathbb{Q}}\left[\tilde{V}_T(\vartheta)\right] \leq \mathbb{E}_{\mathbb{Q}}\left[\tilde{V}_0(\vartheta)\right] \leq 0\,.$$

Thus, if $\tilde{V}_T(\vartheta) \geq 0$ $\mathbb{P}$-a.s., we conclude that $\tilde{V}_T(\vartheta) \geq 0$ $\mathbb{Q}$-a.s. The above requirements imply that $\mathbb{Q}\big(\tilde{V}_T(\vartheta) > 0\big) = 0$, which implies $\mathbb{P}\big(\tilde{V}_T(\vartheta) > 0\big) = 0$. We obtain the result that the market has no arbitrage. ■

The measure $\mathbb{Q}$ is called an equivalent (local) martingale measure. Theorem 6.1.4 states the relevant direction that if an equivalent local martingale measure exists, then the market has no arbitrage. It can be shown that the market satisfies an even stronger condition of 'no free lunch with vanishing risk (NFLVR)'. Moreover, an equivalence holds true. If the market satisfies the NFLVR condition, then there exists an equivalent martingale measure, see Delbaen and Schachermayer [1994]. This result asks for a much deeper proof and is nowadays referred to as the *fundamental theorem of asset pricing*.

Consider in the sequel the market model

$$\begin{aligned} dS_t^j &= b_t^j\,dt + \sum_{i=1}^d \sigma_t^{ji}\,dB_t^i \\ &= b_t^j\,dt + \langle\sigma_t^j\,dB_t\rangle \end{aligned}$$

with $dS_t^0 = \rho_t S_t^0\,dt$, $S_0^0 = 1$, and with a $(d \times d)$ adapted matrix process σ and a d-dimensional standard Brownian motion B. Based on Theorem 4.3.4 and the results from Section 4.3, we obtain that if a d-dimensional process u_t exists such that

$$\sigma_t u_t = b_t - \rho_t S_t \text{ and } \mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \langle u_t,u_t\rangle\,dt\right)\right] < \infty\,,$$

the market has no arbitrage.

Proposition 6.1.5 Let $\vartheta^1, \dots, \vartheta^j$ be chosen. Choosing a specific ϑ^0, one can always make the portfolio self-financing. In particular, this is possible for any $V_0(\vartheta)$.

Proof. By virtue of the the self-financing property, it holds that

$$\begin{aligned} V_t(\vartheta) = \langle \vartheta_t, S_t \rangle &= \vartheta_t^0 S_t^0 + \sum_{i=1}^d \vartheta_t^i S_t^i \\ &= V_0(\vartheta) + \int_0^t \langle \vartheta, dS \rangle \\ &= V_0(\vartheta) + \int_0^t \vartheta_s^0 \, dS_s^0 + \sum_{i=1}^d \int_0^t \vartheta_s^i \, dS_s^i \, . \end{aligned}$$

Thereby, we derive that

$$d(\vartheta_t^0 S_t^0) = \vartheta_t^0 \, dS_t^0 + dA_t \text{ with } A_t = \sum_{i=1}^d \left(\int_0^t \vartheta_s^i \, dS_s^i - \vartheta_t^i S_t^i \right) .$$

Since $S_t^0 = \exp\left(\int_0^t \rho_s \, ds \right)$,

$$d(\vartheta_t^0 S_t^0) = \rho_t \left(\vartheta_t^0 S_t^0 \right) dt + dA_t \, .$$

This affine SDE admits the solution (cf. Proposition 3.4.1)

$$\vartheta_t^0 S_t^0 = S_t^0 \vartheta_0^0 + S_t^0 \int_0^t \left(S_s^0 \right)^{-1} dA_s \, .$$

Dividing by S_t^0 and applying partial integration yields

$$\begin{aligned} \vartheta_t^0 &= \vartheta_0^0 + \int_0^t \left(S_s^0 \right)^{-1} dA_s \\ &= \vartheta_0^0 + \left(S_t^0 \right)^{-1} A_t - A_0 - \int_0^t A_s \, d\left(S_s^0 \right)^{-1} \\ &= V_0(\vartheta) + \left(S_t^0 \right)^{-1} A_t + \int_0^t \rho_s \left(S_s^0 \right)^{-1} A_s \, ds \, . \end{aligned} \tag{6.1.1}$$

Choosing ϑ_0 according to (6.1.1), the portfolio is self-financing. ■

6.2 Applications for the Black-Scholes model

Definition 6.2.1 A (contingent) T-claim is a lower bounded $\mathcal{F}_T$-measurable random variable H_T. We say that the claim H_T is attainable if there exists an admissible self-financing strategy ϑ such that $V_T(\vartheta) = H_T$.

Important examples of contingent claims are options and futures. The *European call option* gives the owner the right (but not the obligation) to buy one stock at the specified price K (strike price) at maturity time T. If the price of this stock satisfies $S_T^j(\omega) > K$, then the owner of the option obtains the payoff $S_T^j(\omega) - K$ at maturity time T, while $S_T^j(\omega) \leq K$

results in the payoff 0 and the owner will not exercise his option in this case. Hence, the payoff is $C_T = (S^j_T(\omega) - K)^+$. The European put option gives the owner the right (but not the obligation) to sell one stock at strike price K at time T. This option has the payoff $P_T = (K - S^j_T(\omega))^+$.
More complicated financial instruments include American options, where the option can be executed up to time T at a fix strike price. Exotic options as barrier or Asian options depend on the path $(S^j_t)_{t\in[0,T]}$.

Definition 6.2.2 For an attainable T-claim H_T, the smallest value of admissible self-financing strategies ϑ with $V_T(\vartheta) = H_T$,

$$\pi_t(H_T) = \operatorname{essinf}_{\vartheta} V_t(\vartheta)$$

is called hedge price (also fair price) of H_T at time t.

Proposition 6.2.3 Let H_T be a T-claim and $\mathbb{Q}$ an equivalent martingale measure. If there exists an attainable strategy ϑ for H_T, such that $\tilde{V}_t(\vartheta)$ is a $\mathbb{Q}$-martingale, then

$$\pi_t(\tilde{H}_T) = \mathbb{E}_{\mathbb{Q}}\left[\tilde{H}_T|\mathcal{F}_t\right].$$

Such a portfolio/strategy is called a *replicating or hedging* portfolio/strategy for H_T.

Proof. $\tilde{S}$ is a $\mathbb{Q}$-local martingale. We have that

$$\mathbb{E}_{\mathbb{Q}}\left[\tilde{H}_T|\mathcal{F}_t\right] = \mathbb{E}_{\mathbb{Q}}\left[\tilde{V}_T(\vartheta)|\mathcal{F}_t\right] \le \tilde{V}_t(\vartheta).$$

Thus, $\mathbb{E}_{\mathbb{Q}}\left[\tilde{H}_T|\mathcal{F}_t\right] \le \pi_t(\tilde{H}_T)$. If $\tilde{V}(\vartheta)$ is a $\mathbb{Q}$-martingale, then

$$\tilde{V}_t(\vartheta) = \mathbb{E}_{\mathbb{Q}}\left[\tilde{H}_T|\mathcal{F}_t\right].$$

■

If $\mathbb{Q}$ is not unique $\pi(\tilde{H}_T)$ is independent of $\mathbb{Q}$. If any T-claim $\tilde{H}_T \le 1$ has a replicating strategy, it follows that $\mathbb{Q}$ is unique on $\mathcal{F}_T$ (set $\tilde{H}_T = \mathbf{1}_A$ for $A \in \mathcal{F}_T$). The market is called *complete* if for every bounded T-claim there exists a hedging strategy. If $\mathbb{Q}$ is unique the market is complete. The last assertion, however, is not easy to establish in general, see, for instance, Chapter 6 of Duffie [2010]. Instead, we establish at the end of this section completeness of the Black-Scholes model directly.

In the sequel, we focus on the *Black-Scholes model* with two assets S^0 and S^1, where

$$dS^1_t = S^1_t \mu \, dt + S^1_t \sigma \, dB_t \tag{6.2.1a}$$

with a one-dimensional standard Brownian motion B. S^0_t is a risk-free asset with interest rate r such that

$$S^0_t = \exp(rt) \,, \; S^1_t = S^1_0 \exp\left(\sigma B_t + \left(\mu t - \frac{\sigma^2 t}{2}\right)\right), 0 \le t \le T\,. \tag{6.2.1b}$$

Proposition 6.2.4 In the Black-Scholes model, (6.2.1b), $\mathbb{Q}$ is unique and

$$d\,\mathbb{Q} = \exp\left(\frac{r-\mu}{\sigma} B_T - \frac{1}{2}\left(\frac{r-\mu}{\sigma}\right)^2 T\right) d\,\mathbb{P}\,. \tag{6.2.2}$$

Proof. We illustrate $\tilde{S}_t^1$ as $\mathbb{Q}$-local martingale:

$$\tilde{S}_t^1 = S_0^1 + \int_0^t h_s \, dW_s$$

with $\mathbb{Q}$-Brownian motion W. If such an illustration exists, $\tilde{S}_t^1$ is a $\mathbb{Q}$-local martingale. We deduce with partial integration that

$$\begin{aligned} d\tilde{S}_t^1 &= d\big(e^{-rt} S_t^1\big) = S_t^1 \, d\big(e^{-rt}\big) + e^{-rt} \, dS_t^1 + d[e^{-rt}, S_t^1] \\ &= -rS_t^1 e^{-rt} \, dt + e^{-rt} S_t^1 \mu \, dt + e^{-rt} S_t^1 \sigma \, dB_t \\ &= \tilde{S}_t^1 \big((\mu - r) \, dt + \sigma \, dB_t\big) \\ &= \sigma \tilde{S}_t^1 \, dW_t \end{aligned} \tag{6.2.3}$$

with $W_t = \sigma^{-1}(\mu - r)t + B_t$. Theorem 4.3.4 and the results from Section 4.3 grant that W is a $\mathbb{Q}$-Brownian motion with

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = M(T) = \exp\Big(\frac{r-\mu}{\sigma} B_T - \frac{(r-\mu)^2}{2\sigma^2} T\Big)$$

if $M(T) = \mathbb{E}_{\mathbb{P}}\left[M(T) | \mathcal{F}_t\right]$ is a martingale. This is ensured by the Novikov criterion. The solution of (6.2.3) is

$$\tilde{S}_t^1 = S_0^1 \exp\big(\sigma W_t - \frac{1}{2}\sigma^2 t\big) .$$

$\tilde{S}_t^1$ is a martingale. ■

Corollary 6.2.5 The Black-Scholes model has no arbitrage.

Lemma 6.2.6 In the Black-Scholes model it holds that $\pi_t(H_T) = V_t(\vartheta)$ for a T-claim $H_T = f(S_T^1), f \geq 0, \mathbb{E}_{\mathbb{Q}}[H_T^2] < \infty$, with $V_t(\vartheta) = F(t, S_t^1)$ and

$$F(t,x) = e^{-r(T-t)} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f\big(x \exp\big(r - \frac{\sigma^2}{2}\big)(T-t) + \sigma u(T-t)\big) e^{-u^2/2} \, du \,, \tag{6.2.4}$$

where $0 \leq t \leq T$ and $x > 0$.

Proof. We have that

$$\begin{aligned} S_T^1 &= \exp\big(\sigma B_T + \big(\mu - \frac{\sigma^2}{2}\big) T\big) \\ &= \exp\big(\sigma W_T + \big(r - \frac{\sigma^2}{2}\big) T\big) \\ &= e^{\sigma W_t + \left(r - \frac{\sigma^2}{2}\right) t} e^{\sigma(W_T - W_t) + \left(r - \frac{\sigma^2}{2}\right)(T-t)} \\ &= S_t^1 e^{\sigma(W_T - W_t) + \left(r - \frac{\sigma^2}{2}\right)(T-t)} . \end{aligned}$$

$\tilde{V}(\vartheta)$ is a $\mathbb{Q}$-martingale for $H_T = f(S_T^1)$, when

$$\mathbb{E}_{\mathbb{Q}}\left[f(\tilde{S}_T^1)|\mathcal{F}_t\right] = \mathbb{E}_{\mathbb{Q}}\left[e^{-rT} f(S_T^1)|\mathcal{F}_t\right] = \tilde{V}_t(\vartheta) = e^{-rt} V_t(\vartheta)\,.$$

We derive that

$$V_t(\vartheta) = \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)} f(S_T^1)|\mathcal{F}_t\right]$$

and that $V_t(\vartheta) = F(t, S_t^1)$ with

$$F(t,x) = \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)} f\big(x\exp\big((r - \frac{\sigma^2}{2})(T-t) + \sigma(W_T - W_t)\big)\big)|\mathcal{F}_t\right],$$

since $W_T - W_t$ is independent from $\mathcal{F}_t$. Using $W_T - W_t \sim N(0, T-t)$ yields the result. ■

In the following the cdf of the standard normal distribution is denoted by Φ.

Theorem 6.2.7 — Black-Scholes formula. In the Black-Scholes model, the hedge price of a European call-option with strike price K is given by

$$C^{BS}(t, S_t^1) = S_t^1 \Phi(d_+) - K e^{-r(T-t)} \Phi(d_-)\,, \tag{6.2.5a}$$

$$\text{with} \;\; d_\pm = \frac{\log(S_t^1/K) + \left(r \pm \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\,. \tag{6.2.5b}$$

■

Proof. With $f(x) = (x-K)^+$ it holds that $H_T = f(S_T^1)$. Applying Lemma 6.2.6 yields

$$C^{BS}(t,x) = \mathbb{E}\left[\left(x\exp\left(-\frac{\sigma^2}{2}(T-t) + \sigma\sqrt{T-t}Z\right) - Ke^{-r(T-t)}\right)^+\right],$$

with $Z \sim N(0,1)$ a standard normal random variable. We rewrite this

$$C^{BS}(t,x) = \mathbb{E}\left[\left(x\exp\left(-\frac{\nu^2}{2} + \nu Z\right) - k\right)^+\right],$$

with $k = Ke^{-r(T-t)}$ and $\nu = \sigma\sqrt{T-t}$. The integrand above is positive if

$$Z \geq \nu^{-1}\big(-\log(x/k) + \nu^2/2\big) = -d_-\,,$$

such that

$$\begin{aligned} C^{BS}(t,x) &= \frac{x}{\sqrt{2\pi}} \int_{-d_-}^{\infty} \exp\left(-\frac{\nu^2}{2} + \nu u - u^2/2\right) du - \frac{k}{\sqrt{2\pi}} \int_{-d_-}^{\infty} e^{-u^2/2}\, du \\ &= x\big(1 - \Phi(-d_- - \nu)\big) - k\Phi(d_-) = x\Phi(d_+) - k\Phi(d_-)\,, \end{aligned}$$

with

$$d_- = \frac{\log(x/k) - \frac{\nu^2}{2}}{\nu}\,,\; d_+ = d_- + \nu\,.$$

■

In the Black-Scholes model, the hedge price of a European put-option with strike price K is determined by Theorem 6.2.7 and the *put-call parity*

$$C^{BS}(t,S_t^1) - P^{BS}(t,S_t^1) = S_t^1 - Ke^{-r(T-t)}\,.$$

Proposition 6.2.8 — Δ-hedge. Let $H_T = f(S_T^1)$ be a T-claim with $\mathbb{E}_{\mathbb{Q}}[H_T^2] < \infty$ and $V_t(\vartheta) = F(t,S_t^1)$. If $F(t,x)$ is twice continuously differentiable with respect to x, then

$$\vartheta_t^0 = e^{-rt}\big(F(t,S_t^1) - \vartheta_t^1 S_t^1\big)\,,$$
$$\vartheta_t^1 = \frac{\partial F}{\partial x}(t,S_t^1)$$

is a hedging strategy with continuous paths. $\Delta = \frac{\partial F}{\partial x} = \frac{\partial V}{\partial S^1}$ is called the option's delta and the above hedging strategy Δ-hedge.

Proof. It holds that

$$\tilde{V}_t(\vartheta) = \tilde{F}(t,\tilde{S}_t^1) = e^{-rt}F\big(t,\tilde{S}_t^1 e^{rt}\big) = e^{-rt}F(t,S_t^1)\,.$$

An application of Itô's formula gives

$$d\tilde{V}_t(\vartheta) = \Big(\frac{\partial \tilde{F}}{\partial t}(t,\tilde{S}_t^1) + \frac{1}{2}\frac{\partial^2 \tilde{F}}{\partial y^2}(t,\tilde{S}_t^1)\sigma^2(\tilde{S}_t^1)^2\Big)dt + \frac{\partial \tilde{F}}{\partial y}(t,\tilde{S}_t^1)\,d\tilde{S}_t^1\,.$$

Since $\tilde{V}(\vartheta)$ is a $\mathbb{Q}$-martingale and the second addend of the right-hand side is as well a local $\mathbb{Q}$-martingale, the integral over the first addend vanishes:

$$\int_0^t \Big(\frac{\partial \tilde{F}}{\partial s}(s,\tilde{S}_s^1) + \frac{1}{2}\frac{\partial^2 \tilde{F}}{\partial y^2}(s,\tilde{S}_s^1)\sigma^2(\tilde{S}_s^1)^2\Big)ds = 0\,.$$

Hence, we conclude that

$$d\tilde{V}_t(\vartheta) = \frac{\partial \tilde{F}}{\partial y}(t,\tilde{S}_t^1)\,d\tilde{S}_t^1\,.$$

Variables y and x are linked by $y = e^{-rt}x$. Putting

$$\vartheta_t^1 = \frac{\partial \tilde{F}}{\partial y}(s,\tilde{S}_t^1) = \frac{\partial F}{\partial x}(t,S_t^1)$$

and $\vartheta_t^0 = \tilde{F}(t,\tilde{S}_t^1) - \vartheta_t^1\tilde{S}_t^1$, ϑ is self-financing with $\tilde{V}_t(\vartheta) = \tilde{F}(t,\tilde{S}_t^1)$ and thus $V_T(\vartheta) = f(S_T^1)$. ■

Corollary 6.2.9 — Black-Scholes PDE. Let $H_T = f(S_T^1)$ and ϑ be the Δ-hedge. Then $V_t(\vartheta) = F(t, S_t^1)$ with F satisfying the partial differential equation

$$\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 F}{\partial x^2} + rx\frac{\partial F}{\partial x} = rF$$

and boundary condition $F(T,x) = f(x)$.

Proof. The preceding proposition implies that

$$\frac{\partial \tilde{F}}{\partial t} + \frac{1}{2}\sigma^2 y^2 \frac{\partial^2 \tilde{F}}{\partial y^2} = 0\,.$$

Inserting $\tilde{F} = e^{-rt}F(t, xe^{rt})$ and carefully considering the variable transform $y = e^{-rt}x$ yields the result. ■

The Black-Scholes partial differential equation can be obtained by the backward Kolmogorov equation from Corollary 5.5.2 as well. Consider the stochastic differential equation

$$dS_t^1 = rS_t^1\,dt + \sigma S_t^1\,dW_t$$

under $\mathbb{Q}$. The transition density p of the solution process satisfies according to Corollary 5.5.2 the partial differential equation

$$\frac{\partial p}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial x^2} + r\frac{\partial p}{\partial x} = 0\,.$$

Based On Lemma 5.5.1, we derive for

$$V_t(\vartheta) = F(t, S_t^1) = \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)} f(S_T^1)|\mathcal{F}_t\right]$$

the Black-Scholes partial differential equation

$$\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 F}{\partial x^2} + rx\frac{\partial F}{\partial x} - rF = 0\,.$$

Proposition 6.2.10 Let $a > 0$. In the Black-Scholes model there exists a self-financing strategy ϑ such that $V_0(\vartheta) = 0$ and $\tilde{V}_T(\vartheta) = a$. This ϑ is not admissible.

Proof. Set

$$Z_t^1 = \frac{1}{\sigma \tilde{S}_t^1 \sqrt{T-t}}\,,\ 0 \le t < T\,.$$

Choose Z_t^0 such that $Z = (Z_t^0, Z_t^1)$ is self-financing and $V_0(\vartheta) = 0$. $\tilde{S}_t^1$ is a $\mathbb{Q}$-local martingale and by (6.2.3):

$$d\tilde{S}_t^1 = \sigma \tilde{S}_t^1\,dW_t\,,$$

with $\mathbb{Q}$-Brownian motion W. Therefore

$$d\tilde{V}_t(Z) = Z_t^1\, d\tilde{S}_t^1 = \frac{1}{\sqrt{T-t}}\, dW_t\,.$$

Since $\tilde{V}_t(Z) = Z_t^0 + Z_t^1 \tilde{S}_t^1$, we obtain that

$$Z_t^0 = \int_0^t \frac{dW}{\sqrt{T-s}} - \frac{1}{\sigma\sqrt{T-t}}\;,\; 0 \le t < T\,.$$

Z is self-financing. $\tilde{V}_t(Z)$ is a Gaussian process with

$$\mathrm{Cov}_{\mathbb{Q}}\left(\tilde{V}_s(Z), \tilde{V}_t(Z)\right) = \int_0^{t\wedge s} \frac{1}{T-u}\, du = -\log\left(1 - \frac{s\wedge t}{T}\right).$$

The time-shifted process

$$\tilde{W}_t = \tilde{V}_{T(1-e^{-t})}(Z)$$

is a Brownian motion using Lévy's theorem since

$$\mathrm{Cov}_{\mathbb{Q}}\left(\tilde{W}_s, \tilde{W}_t\right) = t \wedge s\,.$$

Define the stopping times

$$\begin{aligned} T_1 &= \inf\left\{s \,|\, \tilde{V}_s(Z) = a\right\}, \\ T_2 &= \inf\left\{s \,|\, \tilde{W}_s = a\right\}. \end{aligned}$$

The are related by $T_1 = T(1-e^{-T_2})$. Since $\tilde{W}$ is a $\mathbb{Q}$-Brownian motion:

$$\mathbb{Q}(T_2 < \infty) = 1 \;\Rightarrow\; \mathbb{Q}(T_1 < T) = 1 \;\Rightarrow\; \mathbb{P}(T_1 < T) = 1\,.$$

The strategy that obeys $V_0(\vartheta) = 0$ and $\tilde{V}_T(\vartheta) = a$ is now defined as

$$\vartheta_t^0 = \begin{cases} Z_t^0 & , 0 \le t \le T_1 \\ a & , T_1 < t \le T \end{cases}\;,\; \vartheta_t^1 = \begin{cases} Z_t^1 & , 0 \le t \le T_1 \\ 0 & , T_1 < t \le T \end{cases}\,.$$

Note however that $\tilde{V}(\vartheta)$ is not bounded from below. ■

Proposition 6.2.11 The Black-Scholes market model is complete.

Proof.

$$M(t) = \mathbb{E}_{\mathbb{Q}}\left[\tilde{H}_T | \mathcal{F}_t\right] \ge 0$$

is a $\mathbb{Q}$-martingale. By the representation theorem for continuous local martingales [Kallenberg, 2002, Thm. 18.12] there exists a Brownian motion W such that

$$M(t) = M(0) + \int_0^t h(s)\, dW(s)$$

holds for all $t \geq 0$ $\mathbb{Q}$-almost surely. We derive that

$$dM(t) = h(t)\,dW(t) = h(t)\big(\sigma \tilde{S}_t^1\big)^{-1}\,d\tilde{S}_t^1$$

such that

$$\vartheta_t^1 = h(t)\big(\sigma \tilde{S}_t^1\big)^{-1}\,,\ \vartheta_t^0 = M(t) - \vartheta_t^1 \tilde{S}_t^1$$

satisfies $\tilde{V}_t(\vartheta) = M(t) \geq 0$ and

$$d\tilde{V}_t(\vartheta) = dM(t) = \vartheta_t^1\,d\tilde{S}_t^1\,.$$

such that ϑ is a hedge. ■

7. Stochastic control theory

In this chapter we briefly present the main approach for solving optimal control problems for dynamical systems described by stochastic differential equations: Bellman's principle of dynamic programming and the resulting Hamilton-Jacobi-Bellman equation.

7.1 Stochastic control problems and dynamic programming

We consider the controlled stochastic differential equation

$$dX^u(t) = b(t, X^u(t), u(t))\,dt + \sigma(t, X^u(t), u(t))\,dW(t), \quad t \in [0,T], \quad X^u(0) = x,$$

where X is d-dimensional, $x \in \mathbb{R}^d$, W is a m-dimensional Brownian motion and the coefficients $b : [0,T] \times \mathbb{R}^d \times U \to \mathbb{R}^d$, $\sigma : [0,T] \times \mathbb{R}^d \times U \to \mathbb{R}^{d \times m}$ are regular, say Lipschitz continuous, in x and depend on the control function $u : [0,T] \to U$, for some space U, usually a subset of $\mathbb{R}^k, k \in \mathbb{N}$.

Definition 7.1.1 A control function u is called admissible if it satisfies

1. $u(t)$ is $\mathcal{F}_t$-adapted.
2. $u(t,\omega) \in U$ for all t and ω.
3. the stochastic differential equation has a (strong) unique solution.

One usually considers a fixed Wiener process W and strong solutions for admissible u. However, when strong solutions may not exist such a definition may be too restrictive and it is not uncommon to generalize the above definition for admissible u only demanding for a weak solution.

Definition 7.1.2 An admissible control u is called a Markov control if it is of the form $u(t,\omega) = \alpha(t, X^u(t,\omega))$ with some $\alpha : [0,T] \times \mathbb{R}^d \to U$.

For a Markov strategy, $X^u(t)$ is a Markov process. This is not necessarily true in general, where the control can depend on the entire past. We focus on Markov controls directly, but in fact the methods obtained in this chapter automatically give rise to Markov controls.

The goal is to choose the control u in such a way that a given cost functional

$$J(u) := \mathbb{E}\Big[\int_0^T f(t, X^u(t), u(t))\,dt + h(X^u(T))\Big], \tag{7.1.1}$$

where $f: [0,T] \times \mathbb{R}^d \times U \to \mathbb{R}$ and $h: \mathbb{R}^d \to \mathbb{R}$ are certain continuous (or at first measurable) functions, is minimized (or equivalently an output functional maximized). The function f determines the running costs while h is tantamount to the terminal cost.

In the sequel we suppose $T \in \mathbb{R}_+$ fixed (finite time). Below, we also consider T a certain type of stopping time (indefinite time). Other control problems are stated for infinite time with cost functionals of the form

$$J_\lambda(u) = \mathbb{E}\left[\int_0^\infty e^{-\lambda s} f\big(X^u(s), u(s)\big)\,ds\right]$$

with discounted cost criterion or

$$J(u) = \limsup_{T\to\infty} T^{-1}\,\mathbb{E}\left[\int_0^\infty f\big(X^u(s), u(s)\big)\,ds\right]$$

with a time-average cost criterion.

■ **Example 7.1** A standard example of stochastic control is to select a self-financing portfolio of assets, which is in some sense optimal. Suppose a riskless asset S_0 like a bond grows by a constant rate $r > 0$ over time

$$dS_0(t) = rS_0(t)\,dt,$$

while a risky asset S_1 like a stock follows the scalar diffusion equation of a geometric Brownian motion (the Black-Scholes model)

$$dS_1(t) = S_1(t)\Big(b\,dt + \sigma\,dW(t)\Big).$$

Since this second asset is risky, it is natural to suppose $b > r$. The agent has at each time t the possibility to trade, that is to decide the fraction $u(t)$ of his wealth $X(t)$ which is invested in the risky asset S_1. Under this model we can derive the stochastic differential equation governing the dynamics of the agent's wealth:

$$\begin{aligned} dX(t) &= u(t)X(t)\Big(b\,dt + \sigma\,dW(t)\Big) + (1-u(t))X(t)r\,dt \\ &= (r + (b-r)u(t))X(t)\,dt + \sigma u(t)X(t)\,dW(t). \end{aligned}$$

Without short selling or borrowing, necessarily $u(t) \in [0,1]$ has to hold for all t and we choose $U = [0,1]$. More generally, allowing for short selling or borrowing, we take $U = \mathbb{R}$. Suppose the investor wants to maximize his average utility at time $T > 0$, where the utility is usually assumed to be a concave function $\tilde{U}$ of the wealth. Then the cost functional is

$$J(u) = -\,\mathbb{E}[\tilde{U}(X(T))]\,.$$

Note that the expectation depends of course on the initial wealth endowment $X(0)$ and the chosen investment strategy u. ■

The problem of optimal stochastic control can now be stated as follows: Find for given (s,x) a control process u^* such that

$$J(u^*) = \inf_u J(u) \, .$$

The existence of an optimal control process is not always ensured, but in many cases follows from the setup of the problem or by compactness arguments.

For a Markov control and finite time, $T < \infty$, consider the following cost-to-go functionals:

$$\begin{aligned} J_t^u(X^u(t)) &= \mathbb{E}\Big[\int_t^T f(s,X^u(s),u(s))\,ds + h(X^u(T))\big|\mathcal{F}_t\Big] \\ &= \mathbb{E}\Big[\int_t^T f(s,X^u(s),u(s))\,ds + h(X^u(T))\big|X^u(t)\Big] \, , \end{aligned} \tag{7.1.2}$$

where we use the Markov property. We can now state the main tool we want to use for solving this optimization problem. Conceptually, the idea is to study how the optimal cost changes over time and state. This means that we shall consider the so-called *value function*

$$V(s,y) := \inf_u J_s^u(y), \quad (s,y) \in [0,T) \times \mathbb{R}^d,$$

with its natural extension $V(T,y) = h(y)$. Assume u^* exists with $J_t^{u^*}(x) \le J_t^u(x)$ for all admissible u for all x and for all $t \in [0,T)$.

Intuitively, Bellman's dynamic programming principle asserts that a globally optimal control u over the period $[0,T]$ is also locally optimal for shorter periods. In other words, we cannot improve upon a globally optimal control by optimising separately on smaller subintervals. If this were the case, we could simply patch together these controls to obtain a globally better control.
The value function satisfies the recursion

$$V(r,X^u(r)) = \inf_{u'} \mathbb{E}\Big[\int_r^t f(s,X^{u'}(s),u'(s))\,ds + V(t,X^{u'}(t))\big|X^{u'}(r)\Big] \, ,$$

where the infimum is taken over all admissible Markov controls u' that coincide with u on the interval $[0,r)$. The minimum (existence assumed) is attained by the control which coincides with the optimal control u^* on the interval $[r,t]$. Then, we can partition the interval $[0,T]$ up into bins $[0,t_1]$, $[t_1,t_2], \ldots, [t_n,T]$. Setting $V(T,x) = h(x)$, we can obtain $V(t_n,x)$ by computing the minimum above with $r = t_n$, $t = T$, and this immediately gives us the optimal control on the interval $[t_n,T]$. Next, we compute the optimal control on the previous interval $[t_{n-1},t_n]$ by minimizing the above expression with $r = t_{n-1}$, $t = t_n$ (we know $V(t_n,x)$ from the previous minimization), and iterating this procedure gives the optimal control u^* on the entire interval $[0,T]$. This idea becomes particularly powerful if we let the partition size go to zero.
We derive the following recursion for the cost-to-go functionals:

$$\begin{aligned} J_r^u(X^u(r)) &= \mathbb{E}\Big[\int_r^T f(s,X^u(s),u(s))\,ds + h(X^u(T))\big|X^u(r)\Big] \\ &= \mathbb{E}\Big[\int_r^t f(s,X^u(s),u(s))\,ds + J_t^u(X^u(t))\big|X^u(r)\Big] \, , \end{aligned}$$

with iterated conditional expectations for all $0 \le r \le t \le T$. For some control u' with u' coinciding with u on $[0,t)$ and u' coinciding with u^* on $[t,T]$, we obtain that

$$V(r,X^u(r)) \le J_r^{u'}(X^{u'}(r)) = \mathbb{E}\Big[\int_r^t f(s,X^u(s),u(s))\,ds + V(t,X^{u^*}(t))\big|X^{u'}(r)\Big]. \quad (7.1.3)$$

If $V(t,x)$ is continuously differentiable with respect to t and twice continuously differentiable with respect to x, Itô's formula yields:

$$V(t,X^u(t)) = V(r,X^u(r)) + \int_r^t \Big(\frac{\partial V(s,X^u(s))}{\partial s} + A_s^u V(s,X^u(s))\Big)\,ds + \text{local martingale}$$

where we have denoted by A_s^u the infinitesimal generator associated to $X^u(s)$:

$$A_s^u f(y) = \frac{1}{2}\sum_{i,j=1}^d (\sigma\sigma^T(s,y,u))_{ij}\frac{\partial^2 f}{\partial y_i \partial y_j}(y) + \langle b(s,y,u), \mathrm{grad}(f)(y)\rangle\,.$$

If the local martingale is a martingale, we obtain that

$$V(r,X^u(r)) = \mathbb{E}\left[\int_r^t \Big(-\frac{\partial V(s,X^u(s))}{\partial s} - A_s^u V(s,X^u(s))\Big)\,ds + V(t,X^u(t))\big|X^u(r)\right]$$

and with (7.1.3) the inequality

$$\mathbb{E}\left[\int_r^t \Big(\frac{\partial V(s,X^u(s))}{\partial s} + A_s^u V(t,X^u(t)) + f(s,X^u(s),u(s))\Big)\,ds\big|X^u(r)\right] \ge 0\,,$$

for all $0 \le r \le t \le T$, which becomes an equality for $u = u^*$ on $[r,T]$. This gives formally for $t \downarrow r$

$$\inf_u\Big(\frac{\partial V(s,X^u(s))}{\partial s} + A_s^u V(s,X^u(s)) + f(s,X^u(s),u(s))\Big) = 0\,,$$

the *Bellman equation*. In terms of the so-called Hamiltonian

$$H(t,x,\partial_x V(s,x),\partial^2_{xx}V(s,x)) = A_s^\alpha V(s,x) + f(s,x,\alpha)$$

we arrive at the *Hamilton-Jacobi-Bellman (HJB) equation*

$$\frac{\partial V(s,x)}{\partial s} + \inf_\alpha H\big(s,x,\partial_x V(s,x),\partial^2_{xx}V(s,x)\big) = 0\,, \quad (s,x) \in [0,T) \times \mathbb{R}^d\,.$$

Though we restricted to establish the Bellman equation formally under some strong assumptions, the Bellman equation will still hold under surprisingly general conditions (provided that we introduce an appropriate theory of weak solutions). The theory of viscosity solutions is designed especially for this purpose. This direction is highly technical. We take the perpendicular direction by turning the story upside down. Rather than starting with the optimal control problem, and showing that the Bellman equation follows, we will start with the Bellman equation (regarded simply as a nonlinear PDE) and suppose that we have found a solution. We will then show that this solution does indeed coincide with the value function of an optimal control problem, and that the control function obtained from the minimum in the Bellman equation is optimal. This procedure is called verification. This allows to solve a variety of control problems, while avoiding many technicalities. It should be noted that stochastic optimal control problems are much better behaved, in general, than their deterministic counterparts. In particular, hardly any deterministic optimal control problem admits a nice solution to the Bellman equation, so that the approach of this chapter would be very restrictive in the deterministic case; however, the noise in our equations actually regularizes the Bellman equation, so that sufficiently smooth solutions are not uncommon.

7.2 Verification of the finite time case

Proposition 7.2.1 Let $V(t,x)$ be continuously differentiable with respect to t and twice continuously differentiable with respect to x, such that

$$\frac{\partial V(t,x)}{\partial t} + \min_{\alpha}\Big(A_t^{\alpha} V(t,x) + f(s,x,\alpha)\Big) = 0\,,$$

with $V(x,T) = h(x)$ and $|\mathbb{E}[V(0,X(0))]| < \infty$. Let $\alpha^* \in \operatorname{argmin}_{\alpha}\big(A_T^{\alpha} V(t,x) + f(t,x\alpha)\big)$ and $\mathcal{U}$ be the class of admissible controls for which

$$\sum_{i=1}^{d}\sum_{k=1}^{m}\int_0^t \frac{\partial V(t,X^u(s))}{\partial x_i}\sigma_{ik}(s,X^u(s),u(s))\,dW_k(s)$$

is a martingale and $u^*(t) = \alpha^*(t,X^{u^*}(t)) \in \mathcal{U}$. Then, $V(t,x) = J_t^{u^*}(x)$ solves the optimal control problem.

Proof. For any $u \in \mathcal{U}$, by Itô's formula, we obtain that

$$\mathbb{E}[V(0,X(0))] = \mathbb{E}\left[\int_0^T \Big(-\frac{\partial V(s,X^u(s))}{\partial s} - A_s^u V(s,X^u(s))\Big)\,ds + V(T,X^u(T))\right].$$

With $V(T,x) = h(x)$ and the Bellman equation we derive that

$$\mathbb{E}[V(0,X(0))] \leq \mathbb{E}\left[\int_0^T f(s,X^u(s),u(s))\,ds + h(X^u(T))\right] = J(u)\,.$$

In particular, for $u = u^*$, equality holds. ∎

We also state the dynamic programming principle with a reference for a more detailed proof in a general setting.

Theorem 7.2.2 (Bellman's dynamic programming principle) Under certain regularity conditions for any $(s,y) \in [0,T) \times \mathbb{R}^d$ and $z \in [0,T]$:

$$V(s,y) = \inf_u \mathbb{E}\left[\int_s^z f(t,X^u(t),u(t))\,dt + V(z,X^u(z))\right].$$

Proof. See Theorem 4.3.3 in Yong and Zhou [1999]. ∎

Let us also cite the standard classical verification theorem from [Yong and Zhou, 1999, Thm. 5.5.1].

Theorem 7.2.3 Suppose $W \in C^{1,2}([0,T],\mathbb{R}^d)$ solves the HJB equation together with its final value. Then

$$W(s,y) \leq J_s^u(x)$$

holds for all controls u and all (s,y), that is W is a lower bound for the value function. Furthermore, an admissible control u^* is optimal if and only if

$$\frac{\partial V(s,x)}{\partial s} + \inf_{\alpha} H\big(s,x,\partial_x V(s,x),\partial_{xx}^2 V(s,x)\big) = 0, \quad (s,x) \in [0,T) \times \mathbb{R}^d .$$

holds for $t \in [0,T]$ almost surely.

Example 7.1 Let us close this section by reconsidering the optimal investment example. The Hamiltonian in this case is given by

$$H(t,x,\alpha,p,P) = \frac{1}{2}\sigma^2\alpha^2x^2P + (r + (b-r)\alpha)xp$$

such that the HJB equation reads

$$0 = \partial_t V(t,x) + \min_{\alpha\in\mathbb{R}}\Big(\frac{1}{2}\sigma^2\alpha^2x^2\partial_{xx}V(t,x) + (r+(b-r)\alpha)x\partial_x V(t,x)\Big).$$

We find the optimizing value α^* in this equation by the first order condition

$$\sigma^2\alpha^*x^2\partial_{xx}V(t,x) + (b-r)x\partial_x V(t,x) = 0$$

leading to the more explicit HJB equation

$$\begin{aligned}\partial_t V(t,x) &= -\frac{1}{2}\frac{(r-b)^2x^2(\partial_x V)^2}{\sigma^2x^2\partial_{xx}} - rx\partial_x V + \frac{(r-b)^2x^2(\partial_x V)^2}{\sigma^2x^2\partial_{xx}V} \\ &= -rx\partial_x V + \frac{1}{2}\frac{(r-b)^2(\partial_x V)^2}{\sigma^2\partial_{xx}V}.\end{aligned}$$

For the choice $\tilde{U}(x) = x^\beta, \beta \in (0,1)$, of the cost functional we find for $\beta \in (0,1)$ a solution satisfying the HJB equation and having the correct final value to be

$$V(t,x) = e^{\lambda(T-t)}x^\beta \qquad \text{with} \qquad \lambda = r\beta + \frac{(b-r)^2\beta}{2\sigma^2(1-\beta)}.$$

This yields the optimal feedback function

$$u^*(x,t) = \frac{b-r}{\sigma^2(1-\beta)}.$$

Hence, if $u^* \in [0,1]$ is valid, we have found the optimal strategy just to have a constant fraction of the wealth invested in both assets. Some special choices of the parameters make the optimal choice clearer: for $b \downarrow r$ we will not invest in the risky asset because it does not offer a higher average yield, for $\sigma \to \infty$ the same phenomenon occurs due to the concavity of the utility function penalizing relative losses higher than gains, for $\sigma \to 0$ or $\alpha \to 1$ we do not run into high risk when investing in the stock and thus will do so (even with borrowing for $u^* > 1$). ■

7.3 Verification of the indefinite time case

We now consider the controlled stochastic differential equation

$$dX^u(t) = b(X^u(t),u(t))\,dt + \sigma(X^u(t),u(t))\,dB(t), \quad t \in [0,T_S], \quad X^u(0) = x,$$

with the first exit time

$$T_S = \inf\{t \in \mathbb{R}_+ \,|\, X^u(t) \notin S\}$$

for $S \subseteq U$ and $X^u(0) \in S$. Consider for $X^u(0) = x$ the cost functional

$$J_u(x) = \mathbb{E}_x\Big[\int_0^{T_S} f(X^u(t),u(t))\,dt + h(X^u(T_S))\mathbf{1}_{\{T_S<\infty\}}\Big]. \tag{7.3.1}$$

We focus on Markov controls $u(t,\omega) = u(t, X^u(t,\omega))$ and maximization of the functional

$$V(x) = J_{u^*}(x) = \sup_u J_u(x)\ .$$

Theorem 7.3.1 On the assumptions

(i) Let $V(x)$ be twice continuously differentiable with respect to x in $int(S)$ and continuous on ∂S.

(ii) For all bounded stopping times $\tau \le T_S$, for all $x \in S$ and all u:

$$\mathbb{E}_x\Big[V(X^u(\tau)) + \int_0^{\tau} \big|A^u V(X^u(s))\big|\,ds\Big] < \infty\,.$$

(iii) An optimal control u^* exists.

(iv) If $X^u(0) \in \partial S$, then $T_S = 0$ almost surely.

it holds true that:

1. $V(x) = h(x)$ for all $x \in \partial S$.
2. For all $x \in S$: $\sup_u (A^u V(x) + f(x,u)) = 0$.
3. The optimal control u^* is determined by

$$(A^{u^*} V(x) + f(x,u^*) = 0\,.$$

■

Proof. 1. readily follows from (iv). Define for some T

$$\tilde{u}_t(x) = \begin{cases} u(t)\ , & t \le T_s \wedge T\,, \\ u^*(t)\ , & t > T_s \wedge T\,. \end{cases}$$

Using the strong Markov property, we obtain that

$$\begin{aligned} J_u(x) &= \mathbb{E}_x\Big[\int_0^{T_S\wedge T} f(X^u(t),u(t))\,dt\Big] + \mathbb{E}_x\big[J_{u^*}(X^u(T_S \wedge T))\big] \\ &= \mathbb{E}_x\Big[\int_0^{T_S\wedge T} f(X^u(t),u(t))\,dt\Big] + \mathbb{E}_x\big[V(X^u(T_S \wedge T))\big]\,. \end{aligned}$$

An application of Dynkin's formula from Theorem 5.4.3 yields

$$\mathbb{E}_x\big[V(X^u(T_S \wedge T))\big] = V(x) + \mathbb{E}_x\Big[\int_0^{T_S \wedge T} A^u V(X^u(t))\,dt\Big].$$

Since $V(x) \geq J_u(x)$, we have that

$$\mathbb{E}_x\Big[\int_0^{T_S \wedge T} f(X^u(t), u(t))\,dt\Big] + \mathbb{E}_x\Big[\int_0^{T_S \wedge T} A^u V(X^u(t))\,dt\Big] \leq 0.$$

This implies that

$$\frac{\mathbb{E}_x\Big[\int_0^{T_S \wedge T} \Big(f(X^u(t), u(t)) + A^u V(X^u(t))\Big)\,dt\Big]}{\mathbb{E}_x[T_S \wedge T]} \leq 0.$$

If we let $T \to 0$, the above assumptions ensure that this implies

$$f(x,u) + A^u V(x) \leq 0.$$

In particular, equality holds for $u = u^*$ what finishes the proof. ■

IV

8. Appendix

8.1 Doob's Martingale Inequality

Theorem 8.1.1 Suppose $(X_n, \mathcal{F}_n)_{0 \le n \le N}$ is a discrete martingale. Then for every $p \ge 1$ and $\lambda > 0$

$$\lambda^p \, \mathbb{P}\Big(\sup_{0 \le n \le N} |X_n| \ge \lambda \Big) \le \mathbb{E}[|X_N|^p],$$

and for every $p > 1$

$$\mathbb{E}\Big[\sup_{0 \le n \le N} |X_n|^p \Big] \le \Big(\frac{p}{p-1} \Big)^p \mathbb{E}[|X_N|^p].$$

■

Proof. Introduce the stopping time $\tau := \inf\{n \,|\, |X_n| \ge \lambda\} \wedge N$. Since $(|X_n|^p)$ is a submartingale, the optional stopping theorem gives

$$\mathbb{E}[|X_N|^p] \ge \mathbb{E}[|X_\tau|^p] \ge \lambda^p \, \mathbb{P}\Big(\sup_n |X_n| \ge \lambda \Big) + \mathbb{E}[|X_N|^p \mathbf{1}_{\{\sup_n |X_n| < \lambda\}}],$$

which proves the first part. Moreover, we deduce from this inequality for any $K > 0$ and $p > 1$

$$\begin{aligned}
\mathbb{E}\big[\big(\sup_n |X_n| \wedge K \big)^p \big] &= \mathbb{E}\Big[\int_0^K p \lambda^{p-1} \mathbf{1}_{\{\sup_n |X_n| \ge \lambda\}} \, d\lambda \Big] \\
&\le \int_0^K p \lambda^{p-2} \, \mathbb{E}[|X_N| \mathbf{1}_{\{\sup_n |X_n| \ge \lambda\}}] \, d\lambda \\
&= p \, \mathbb{E}\Big[|X_N| \int_0^{\sup_n |X_n| \wedge K} \lambda^{p-2} \, d\lambda \Big] \\
&= \frac{p}{p-1} \mathbb{E}\big[|X_N| (\sup_n |X_n| \wedge K)^{p-1} \big].
\end{aligned}$$

By Hölder's inequality,

$$\mathbb{E}\big[\big(\sup_n |X_n| \wedge K\big)^p\big] \le \frac{p}{p-1}\,\mathbb{E}\big[\big(\sup_n |X_n| \wedge K\big)^p\big]^{(p-1)/p}\,\mathbb{E}[|X_N|^p]^{1/p},$$

which after cancellation and taking the limit $K \to \infty$ yields the asserted moment bound. ■

Corollary 8.1.2 (Doob's L^p-inequality) If $(X(t), \mathcal{F}_t)_{t\in I}$ is a right-continuous martingale indexed by a subinterval $I \subset \mathbb{R}$, then for any $p > 1$

$$\mathbb{E}\big[\sup_{t\in I} |X(t)|^p\big]^{1/p} \le \frac{p}{p-1} \sup_{t\in I} \mathbb{E}[|X(t)|^p]^{1/p}.$$

Proof. By the right-continuity of X we can restrict the supremum on the left to a countable subset $D \subset I$. This countable set D can be exhausted by an increasing sequence of finite sets $D_n \subset D$ with $\bigcup_n D_n = D$. Then the supremum over D_n increases monotonically to the supremum over D, the preceding theorem applies for each D_n and the monotone convergence theorem yields the asserted inequality. ■

Be aware that Doob's L^p-inequality is different for $p = 1$ [Revuz and Yor, 1999, p. 55].

Corollary 8.1.3 For any $X \in V$ there exists a version of $\int_0^t X(s)\,dW(s)$ that is continuous in t, i.e. a continuous process $(J(t),\, t \ge 0)$ with

$$\mathbb{P}\Big(J(t) = \int_0^t X(s)\,dW(s)\Big) = 1 \text{ for all } t \ge 0.$$

Proof. Let $(X_n)_{n\ge1}$ be an approximating sequence for X of simple processes in V. Then by definition $I_n(t) := \int_0^t X_n(s)\,dW(s)$ is continuous in t for all ω. Moreover, $I_n(t)$ is an $\mathcal{F}_t$-martingale so that Doob's inequality and the Itô isometry yield the Cauchy property

$$\mathbb{E}\Big[\sup_{t\ge0} |I_m(t) - I_n(t)|^2\Big] \le 4 \sup_{t\ge0} \mathbb{E}[|I_m(t) - I_n(t)|^2] = 4\|X_m - X_n\|_V^2 \to 0$$

for $m, n \to \infty$. By the Chebyshev inequality and the Lemma of Borel-Cantelli there exist a subsequence $(I_{n_l})_{l\ge1}$ and $L(\omega)$ such that $\mathbb{P}$-almost surely

$$\forall\, l \ge L(\omega) \;\; \sup_{t\ge0} |I_{n_{l+1}}(t) - I_{n_l}(t)| \le 2^{-l}.$$

Hence with probability one the sequence $(I_{n_l}(t))_{l\ge1}$ converges uniformly and the limit function $J(t)$ is continuous. Since for all $t \ge 0$ the random variables $(I_{n_l}(t))_{l\ge1}$ converge in probability to the integral $I(t) = \int_0^t X(s)\,dW(s)$, the random variables $I(t)$ and $J(t)$ must coincide for $\mathbb{P}$-almost all ω. ■

In the sequel we shall consider only t-continuous versions of the stochastic integral.

Bibliography

Freddy Delbaen and Walter Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1):463–520, 1994.

R. M. Dudley. *Probabilities and metrics*. Matematisk Institut, Aarhus Universitet, Aarhus, 1976. Convergence of laws on metric spaces, with a view to statistical testing, Lecture Notes Series, No. 45.

D. Duffie. *Dynamic Asset Pricing Theory*. Princeton Series in Finance. Princeton University Press, 2010. ISBN 9781400829200. URL `https://books.google.de/books?id=f2Wv-LDpsoUC`.

Richard Durrett. *Stochastic calculus. A practical introduction.* Probability and Stochastics Series. Boca Raton, FL: CRC Press., 1996.

W. Feller. *An introduction to probability theory and its applications. Vol II. 2nd ed.* Wiley Series in Probability and Mathematical Statistics. New York etc.: John Wiley and Sons., 1971.

Olav Kallenberg. *Foundations of modern probability. 2nd ed.* Probability and Its Applications. New York, Springer., 2002.

Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus. 2nd ed.* Springer, New York, 1991.

Achim Klenke. *Wahrscheinlichkeitstheorie.* Berlin: Springer, 2008.

Peter E. Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations.* Applications of Mathematics. 23. Berlin: Springer-Verlag., 1992.

Qingkai Kong. *A short course in ordinary differential equations.* Cham: Springer, 2014. ISBN 978-3-319-11238-1/hbk; 978-3-319-11239-8/ebook. doi: 10.1007/978-3-319-11239-8.

Robert S. Liptser and Albert N. Shiryaev. *Statistics of random processes. 1: General theory. 2nd ed.* Springer, Berlin, 2001.

Bernt Øksendal. *Stochastic differential equations. An introduction with applications. 5th ed.* Berlin: Springer., 1998.

Philip Protter. *Stochastic integration and differential equations. A new approach. 2nd corr. printing.* Springer, Berlin, 1992.

Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion. 3rd ed.* Springer, Berlin, 1999.

Jiongmin Yong and Xun Yu Zhou. *Stochastic controls. Hamiltonian systems and HJB equations.* Applications of Mathematics. 43. New York, NY: Springer., 1999.