Mathematical Finance in Continuous Time

Content: Martin Keller-Ressel
Typeset: Philipp Lukas Strietzel

Summer term 2018

Contents

Overview

Central topic of this lecture is financial mathematics in continuous time. The tools to work with the topic are mainly probability theory, martingales, stochastic analysis and partial differential equations.
The topics covered in this lecture are

- replication/hedging and arbitrage theory in continuous time,
- local and stochastic volatility models and stochastic differential equations,
- valuation of European and American options in time-continuous models (partial differential equations with free boundary-value condition)
- affine stochastic processes, ...

0 Repetition Stochastic Analysis

In this lecture there will always be a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $(\mathcal{F}_t)_{t\geq 0}$. Sometimes it will be useful to assume that

- the filtration is continuous from the right, i.e., $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$,
- the filtration is complete, i.e., $\mathcal{F}_0$ contains all $\mathbb{P}$-Nullsets.

Definition 0.1. A **financial market modell (FMM)** is a family of $(d+1)$ adapted stochastic processes $(S_t^0, ..., S_t^d)_{t\geq 0}$ such that $S_t^0 > 0$ almost surely for all $t \geq 0$, where

- S^0 is the 'Numeraire', which means it is a locally riskfree asset, e.g. a bank account or an investment in bonds,
- $S^1, ..., S^d$ are assets, e.g. stocks, currencies, commodities,...

Financial market models are mostly given by stochastic differential equations, driven by a Brownian motion (BM).

Example 0.2 (The Black-Scholes-Model)**.** The **Black-Scholes-Model** is a 2-dimensional financial market model $(S_t^0, ..., S_t)$ given by

$$\begin{cases} S_t^0 = e^{rt}, \\ S_t = \exp\left(\left(\mu - \dfrac{\sigma^2}{2}\right) t + \sigma B_t\right), \end{cases}$$

where $r \in \mathbb{R}$ is the interest rate, $\mu \in \mathbb{R}$ the drift and $\sigma > 0$ the volatility. Hence, S_t is a geometric Brownian motion.
In terms of stochastic differential equations the Black-Scholes-Model is described as

$$\begin{cases} \mathrm{d}S_t^0 = S_t^0 r \,\mathrm{d}t, \\ \mathrm{d}S_t = S_t(\mu \,\mathrm{d}t + \sigma \,\mathrm{d}B_t). \end{cases}$$

0.1 The Brownian Motion

Definition 0.3. The **one-dimensional Brownian motion** $(B_t)_{t\geq 0}$ is a stochastic process such that

a) $B_0 = 0$,

b) B_t has independent increments, i.e., for all $0 \leq t_1 \leq ... \leq t_N$ it holds that

$$(B_{t_N} - B_{t_{N-1}}), (B_{t_{N-2}} - B_{t_{N-1}}), ..., (B_{t_2} - B_{t_1}) \text{ are independent,}$$

c) $(B_t - B_s)$ is normally distributed with mean 0 und variance $t - s$ for all $0 \leq s \leq t$, i.e.,

$$(B_t - B_s) \sim \mathcal{N}(0, t - s),$$

d) (B_t) is a continuous stochastic prozess, i.e.,

$$\exists A \in \mathcal{F} \text{ with } \mathbb{P}(A) = 1 \text{ such that } t \mapsto B_t(\omega) \text{ continuous } \forall\, \omega \in A.$$

Remark 0.4. More general an adapted stochastic process $(B_t)_{t\geq 0}$ is a Brownian motion with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$, if

- (a),(c),(d) and

(b') $(B_t - B_s) \perp\!\!\!\perp \mathcal{F}_s \quad \forall\, 0 \leq s \leq t.$

Proposition 0.5 (Properties of the Brownian motion)**.**

- *$(B_t)_{t\geq 0}$ is Brownian motion with respect to $(\mathcal{F}_t)_{t\geq 0}$ $\Longrightarrow$ $(B_t)_{t\geq 0}$ is Brownian motion.*
- *B is a Brownian motion $\Longrightarrow$ B is a Brownian motion with respect to generated filtration $\mathcal{F}_t = \sigma((B_s)_{0\leq s\leq t})$.*
- *The Brownian motion is a **Gaussian process**, i.e., for all $0 \leq t_1 \leq ... \leq t_N$ it holds that*

  $$(B_{t_1}, B_{t_2}, ..., B_{t_N}) \text{ is multivariate normally distributed.}$$

  The law is completely determined by $\mathbb{E}[B_t] = 0$ and $\operatorname{cov}(B_t, B_s) = \min(t, s)$.
- *A Brownian motion with respect to $(\mathcal{F}_t)_{t\geq 0}$ is a **martingale**, i.e., it is adapted and*
  - $\mathbb{E}[|B_t|] < \infty \;\forall\, t \geq 0$,
  - $\mathbb{E}[B_t|\mathcal{F}_s] = B_s \;\forall\, 0 \leq s \leq t$.
- *The Brownian motion has **infinite total variation**, i.e.,*

  $$\mathbb{P}\text{-}\lim_{|P_n|\to 0} \sum_{t_i \in P_n} \left| B_{t_{i+1}} - B_{t_i} \right| = \infty \quad a.s.$$

  for all sequences of partitions $(P_n)_{n\in\mathbb{N}}$ of a given intervall $[0, T]$.

- *The* **quadratic variation** *of the Brownian motion exists, i.e.,*

$$[B,B]_t := \mathbb{P}\text{-}\lim_{|P_n|\to 0} \sum_{t_i \in P_n} (B_{t_{i+1}} - B_{t_i})^2 < \infty$$

  and

$$[B,B]_t = t \qquad \forall\, t \geq 0. \tag{1}$$

  This forms the basis of Itô-calculus. Note that in Itô-terms (1) *can be written as*

$$(\mathrm{d}B_t)^2 = \mathrm{d}t.$$

Definition 0.6. Let $(B^1, B^2, ..., B^k)$ k be independent Brownian motions (with respect to $(\mathcal{F}_t)_{t\geq 0}$). Then $B = (B^1, ..., B^k)$ is called k**-dimensional (multivariate) Brownian motion**.

The **quadratic covariation** of the components, defined as

$$[B^i, B^j]_t := \mathbb{P}\text{-}\lim_{|P_n|\to 0} \sum_{t_l \in P_n} (B^i_{t_{l+1}} - B^i_{t_l})(B^j_{t_{l+1}} - B^j_{t_l})$$

is

$$[B^i, B^j]_t = \delta_{ij} \cdot t = \begin{cases} t, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

In Itô-terms

$$(\mathrm{d}B^i_t)(\mathrm{d}B^j_t) = \delta_{ij}\, \mathrm{d}t.$$

Proof. For $i = j$ see (1). Let $i \neq j$. The idea is to polarize. Set

$$\widetilde{B}_t := (B^i_t + B^j_t)\frac{1}{\sqrt{2}}.$$

By definition $\widetilde{B}$ is a standard Brownian motion, since

- $\widetilde{B}_0 = 0$,
- $\widetilde{B}$ has independent increments,
- $\widetilde{B}$ is continuous,
- $\widetilde{B}_t \sim N(0,t)$ and $\widetilde{B}_t - \widetilde{B}_s \sim N(0, t-s)$.

Now we have

$$t = \left[\widetilde{B}, \widetilde{B}\right]_t = \frac{1}{2}[B^i + B^j, B^i + B^j]_t = \frac{1}{2}\big(\underbrace{[B^i, B^i]_t}_{t} + 2[B^i, B^j]_t + \underbrace{[B^j, B^j]_t}_{t}\big)$$

which implies

$$t = t + [B^i, B^j]_t \qquad \forall t \geq 0.$$

Hence $[B^i, B^j]_t = 0$ for all $t \geq 0$, $i \neq j$. □

0.2 Generation of Stochastic Processes from a Brownian Motion

There are five fundamental possibilities to generate stochastic processes from Brownian motion.

(1) Stochastic Integration (↗**Itô-integral**)
For suitable integrands ϑ let (Itô-notation)

$$X_t := \int_0^t \vartheta_s \, \mathrm{d}B_s. \qquad \Big| \qquad \mathrm{d}X_t = \vartheta_t \, \mathrm{d}B_t$$

In financial mathematics this is important for hedging strategies.

(2) Composition with a $C^{1,2}(\mathbb{R}_{\geq}0 \times \mathbb{R})$-function $f(t,x)$,

$$Y_t := f(t, B_t). \qquad \Big| \qquad \mathrm{d}Y_t = \mathrm{d}f(t, B_t)$$

Y is now calculable by the **Itô-formula** (see Theorem 0.8).

(3) Sum of an initial value, a normal integral and an Itô-integral

$$X_t := X_0 + \int_0^t \mu_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}B_s \qquad (2) \qquad \Big| \qquad \mathrm{d}X_t = \mu_t \, \mathrm{d}t + \sigma_t \, \mathrm{d}B_t$$

The class of these processes is called **Itô-processes**.

(4) Composition of an Itô-process and a $C^{1,2}$-function

$$Z_t := f(t, X_t). \qquad \Big| \qquad \mathrm{d}Z_t = \mathrm{d}f(t, X_t)$$

(5) Stochastic integration with respect to an Itô-process

$$Y_t = \int_0^t \vartheta_s \, \mathrm{d}X_s. \qquad \Big| \qquad \mathrm{d}Y_t = \vartheta_t \, \mathrm{d}X_t$$

In the following will be discussed some properties of these five possibilites.

0.2.1 Stochastic Integration

The Itô-Integral is well-defined for adapted square integrable processes, i.e., for every

$$\vartheta \in \mathcal{H}^2(0,\infty) := \left\{ \vartheta(\omega,t) \text{ adapted and } \underbrace{\mathbb{E}\left[\int_0^t \vartheta(\omega,s)^2 \,\mathrm{d}s\right]}_{=\|\vartheta\|^2_{L_2(\mathrm{d}\mathbb{P}\times\mathrm{d}t)}} < \infty \quad \forall t > 0 \right\}.$$

It is defined as L_2-Limit of 'simple integrands', whose are piecewise constant elements of $\mathcal{H}^2(0,\infty)$. (See [1, Chapter 3] or [2, Chapter 6].)

Lemma 0.7 (Properties of the Itô-Integral)**.** *Let*

$$X_t = \int_0^t \vartheta_s \,\mathrm{d}B_s.$$

Then the following properties hold.

(i) The ***Itô-isometry***

$$\underbrace{\mathbb{E}[X_t^2]}_{=\|X_t\|_{L_2(\mathrm{d}\mathbb{P})}} = \underbrace{\mathbb{E}\left[\int_0^t \vartheta_s^2 \,\mathrm{d}s\right]}_{=\|\vartheta\|^2_{L_2(\mathrm{d}\mathbb{P}\times\mathrm{d}t)}}.$$

(ii) X_t is a ***continuous martingale****, in particular*

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \qquad \text{and} \qquad \mathbb{E}[X_t] = 0.$$

(iii) The Itô-integral is ***linear****, i.e.,*

$$\int_0^t (c\vartheta_s + \xi_s)\,\mathrm{d}B_s = c \cdot \int_0^t \vartheta_s \,\mathrm{d}B_s + \int_0^t \xi_s \,\mathrm{d}B_s \qquad \forall c \in \mathbb{R},\ \vartheta,\xi \in \mathcal{H}^2(0,\infty).$$

(iv) The ***conditional Itô-isometry***

$$\mathbb{E}\left[\left(\int_s^t \vartheta_r \,\mathrm{d}B_r\right)^2 \middle| \mathcal{F}_s\right] = \mathbb{E}\left[\int_s^t \vartheta_r^2 \,\mathrm{d}r \middle| \mathcal{F}_s\right] \qquad \forall 0 \le s \le t.$$

(v) For the ***quadratic variation*** *it holds that*

$$[X,X]_t = \int_0^t \vartheta_s^2 \,\mathrm{d}s \qquad \text{or} \qquad \mathrm{d}[X,X]_t = (\mathrm{d}X_t)^2 = \vartheta_t^2 \,\mathrm{d}t.$$

Proof. See [1, Chapter 3]. □

0.2.2 Composition with a $C^{1,2}$-function

Y_t can be described by Itô's formula:

Theorem 0.8 (Itô's formula)**.** *Let* $f \in C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R})$*. Then*

$$Y_t = f(t, B_t) = f(0,0) + \int_0^t \frac{\partial f}{\partial t}(s, B_s)\,\mathrm{d}s + \int_0^t \frac{\partial f}{\partial x}(s, B_s)\,\mathrm{d}B_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, B_s)\,\mathrm{d}s.$$

In Itô-notation and with

$$\partial_t f := \frac{\partial f}{\partial t}, \quad \partial_x f := \frac{\partial f}{\partial x}, \quad \partial_{xx} f := \frac{\partial^2 f}{\partial x^2}.$$

We can write Itô's formula as

$$\mathrm{d}Y_t = \mathrm{d}f(t, B_t) = \partial_t f(t, B_t)\,\mathrm{d}t + \partial_x f(t, B_t)\,\mathrm{d}B_t + \frac{1}{2}\partial_{xx} f(t, B_t)\,\mathrm{d}t. \tag{Itô}$$

Proof. For a complete proof see [1, Theorem 5.2]. Here we will do a heuristic derivation using Taylor's formula and Itô-calculus.
Applying Taylor's formula to $f(t,x)$ it holds that

$$f(t+\Delta t, x+\Delta x) - f(t,x) = \partial_t f(t,x)\Delta t + \partial_x f(t,x)\Delta x + \frac{1}{2}\partial_{xx} f(t,x)(\Delta x)^2 + \text{'higher order terms'}.$$

Plugging in B_t for x and passing to the limit yields

$$\begin{aligned}\mathrm{d}f(t, B_t) &= \partial_t f(t, B_t)\,\mathrm{d}t + \partial_x f(t, B_t)\,\mathrm{d}B_t + \frac{1}{2}\partial_{xx} f(t, B_t)\underbrace{(\mathrm{d}B_t)^2}_{=\mathrm{d}t} + \underbrace{o(\mathrm{d}t) + o(B_t^2) + \ldots}_{\to 0} \\ &= \partial_t f(t, B_t)\,\mathrm{d}t + \partial_x f(t, B_t)\,\mathrm{d}B_t + \frac{1}{2}\partial_{xx} f(t, B_t)\,\mathrm{d}t.\end{aligned}$$

□

Example 0.9 (Application of the Itô-formula to the Black-Scholes-Model)**.** Let $f(t,x) = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma x\right)$ and define

$$S_t := f(t, B_t) = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right).$$

Taking partial derivatives yields

$$\begin{aligned}\partial_t f(t,x) &= \left(\mu - \frac{\sigma^2}{2}\right) f(t,x), \\ \partial_x f(t,x) &= \sigma f(t,x), \\ \partial_{xx} f(t,x) &= \sigma^2 f(t,x).\end{aligned}$$

From the Itô-formula (Itô) we derive the stochastic differential equation of the Black-Scholes-Model

$$\begin{aligned}\mathrm{d}S_t &= \mathrm{d}f(t, B_t)\left(\left(\mu - \cancel{\frac{\sigma}{2}}\right)\mathrm{d}t + \sigma\,\mathrm{d}B_t + \cancel{\frac{1}{2}\sigma^2\,\mathrm{d}t}\right) \\ &= S_t(\mu\,\mathrm{d}t + \sigma\,\mathrm{d}B_t)\end{aligned}$$

0.2.3 Sum of an Initial Value, a Lebesgue-Integral and an Itô-Integral

If $\sigma \in \mathcal{H}^2(0,\infty)$ and

$$\int_0^t |\mu(\omega,s)|\,\mathrm{d}s < \infty$$

a.s. forall $t \geq 0$, then Itô-process (2) is well defined.
Note the dichotomy [1] between dt- and dB_t-integral:

Lemma 0.10. *a) An Itô-process of the form* (2) *is a martingale if and only if* $\mu = 0$ *a.s.*

b) An Itô-process of the form (2) *has finite total variation if and only if* $\sigma = 0$ *a.s.*
Here $(\mathcal{F}_t)_{t\geq 0}$ *is the augmented natural filtration of* B*, which is left- and right-continuous.*

Proof. a) '$\Leftarrow$:' Follows directly from the martingale property of $\int_0^t \sigma(\omega,s)\,\mathrm{d}B_s$.

'$\Rightarrow$': Suppose X is martingale, then $\mathbb{E}[X_t - X_s|\mathcal{F}_s] = 0$ for all $s \leq t$. On the other hand we have

$$\mathbb{E}[X_t - X_s|\mathcal{F}_s] = \mathbb{E}\left[\int_s^t \mu(w,r)\,\mathrm{d}r\middle|\mathcal{F}_s\right] + \underbrace{\mathbb{E}\left[\int_s^t \sigma(\omega,r)\,\mathrm{d}B_r\middle|\mathcal{F}_s\right]}_{=0}.$$

Hence

$$\mathbb{E}\left[\int_s^t \mu(\omega,r)\,\mathrm{d}r\big|\mathcal{F}_s\right] = 0,$$

which is equivalent to

$$\mathbb{E}\left[\int_s^t \mu(\omega,r)\,\mathrm{d}r \cdot \mathbb{1}_A\right] = 0 \qquad \forall\, A \in \mathcal{F}_s.$$

Applying Fubini's theorem yields for all $A \in \mathcal{F}_s,\ s \leq t$

$$0 = \mathbb{E}\left[\int_s^t \mu(\omega,r)\,\mathrm{d}r \cdot \mathbb{1}_A\right] = \int_s^t \mathbb{E}[\mu(\omega,r)\mathbb{1}_A]\,\mathrm{d}r$$

Taking derivatives with respect to t leads to

$$\mathbb{E}[\mu(w,t)\mathbb{1}_A] = 0$$

forall $A \in \mathcal{F}_s,\ s \leq t$, and thus

$$\mathbb{E}[\mu(\omega,t)|\mathcal{F}_s] = 0 \qquad \forall\, s \leq t$$

Letting $s \uparrow t$ yields $\mathbb{E}[\mu(\omega,t)|\mathcal{F}_{t-}] = 0$, with $\mathcal{F}_{t-} = \bigcup_{s<t}\mathcal{F}_s$. Hence

$$\mathbb{E}[\mu(\omega,t)|\mathcal{F}_t] = 0 \;\Rightarrow\; \mu(\omega,t) = 0 \quad \forall\, t \geq 0$$

almost surely.

[1] Structure composed of two complementary parts (like yin-yang)

b) '$\Leftarrow$': Is clear, since $X_0 + \int_0^t \mu(\omega, s)\,\mathrm{d}s$ has finite total variation.

'$\Rightarrow$': This direction is based on the relation between total variation (TV) and quadratic variation (QV).
The total variation of a function f is

$$\mathrm{TV}_T(f) := \lim_{|P_n|\to 0} \sum_{t_i \in P_n,\ 0 \le t_i \le T} |f(t_{i+1}) - f(t_i)|\,.$$

If this limit exists independently of the sequence (P_n) $\forall\, T > 0$, then f **is of finite total variation**.
A stochastic process X is of finite total variation if

$$\mathrm{TV}_T(X(\omega)) < \infty \text{ a.s. } \forall T > 0.$$

Suppose X is continuous and of finite TV. Then

$$\sum_{t_i \in P_n} (X_{t_{i+1}} - X_{t_i})^2 \le \underbrace{\max_{t_i \in P_n} \left|X_{t_{i+1}} - X_{t_i}\right|}_{\to 0 \text{ by continuity as } |P_n|\downarrow 0} \cdot \underbrace{\sum_{t_i \in P_n} \left|X_{t_{i+1}} - X_{t_i}\right|}_{\to \mathrm{TV}_T(X) < \infty \text{ as } |P_n|\downarrow 0} \to 0.$$

Hence, X is of finite TV implies $[X, X]_t = 0$ for all $t \ge 0$ and vice versa if $[X, X]_t > 0$ for some $t > 0$ (with strictly positive probability) then X has infinite total variation.
Since X is an Itô-process of finite total variation it follows that $[X, X]_t = 0$, but on the other hand

$$[X, X]_t = \int_0^t \sigma^2(\omega, s)\,\mathrm{d}s.$$

Thus $\sigma(t, \omega) = 0$ almost surely for all $t \ge 0$ and the proof is completed. □

Example 0.11. In the Black-Scholes-Model it holds, that $\sigma > 0$ and S is a martingale. Hence S has infinite total variation and $\mu = 0$.

Lemma 0.12 (Quadratic variation of an Itô-process)**.** *Let X be an Itô-process, i.e.,*

$$\mathrm{d}X_t = \mu_t\,\mathrm{d}t + \sigma_t\,\mathrm{d}B_t.$$

Then

$$[X, X]_t = \int_0^t \sigma_s^2\,\mathrm{d}s.$$

In terms of Itô-calculus

$$\mathrm{d}[X, X]_t = (\mu_t\,\mathrm{d}t + \sigma_t\,\mathrm{d}B_t)^2 = \sigma_t^2\,\mathrm{d}t.$$

Proof. Using Itô calculus it follows that

$$\begin{aligned}\mathrm{d}[X,X]_t = (\mathrm{d}X_t)^2 &= (\mu_t\,\mathrm{d}t + \sigma_t\,\mathrm{d}B_t)^2\\ &= \mu_t^2\underbrace{(\mathrm{d}t)^2}_{0} + 2\mu_t\sigma_t\underbrace{\mathrm{d}t\,\mathrm{d}B_t}_{0} + \sigma_t^2\underbrace{(\mathrm{d}B_t)^2}_{\mathrm{d}t}\\ &= \sigma_t^2\,\mathrm{d}t.\end{aligned}$$

□

Proposition 0.13 (Covariation of two Itô-processes). *Let X, Y be Itô-processes, i.e.,*

$$\begin{aligned}\mathrm{d}X_t &= \mu_t^X\,\mathrm{d}t + \sigma_t^X\,\mathrm{d}B_t,\\ \mathrm{d}Y_t &= \mu_t^Y\,\mathrm{d}t + \sigma_t^Y\,\mathrm{d}B_t.\end{aligned}$$

Then

$$\mathrm{d}[X,Y]_t = (\mathrm{d}X_t)(\mathrm{d}Y_t) = \sigma_t^X\sigma_t^Y\,\mathrm{d}t.$$

Proof. By polarization it holds that

$$\begin{aligned}[X,Y]_t &= \frac{1}{2}\left([X+Y,X+Y]_t - [X,X]_t - [Y,Y]_t\right)\\ &= \frac{1}{2}\int_0^t \left((\sigma_s^X+\sigma_s^Y)^2 - (\sigma_s^X)^2 - (\sigma_s^Y)^2\right)\mathrm{d}s\\ &= \int_0^t \sigma_s^X\sigma_s^Y\,\mathrm{d}s.\end{aligned}$$

□

0.2.4 Composition of an Itô-Process and a $C^{1,2}$-Function

Theorem 0.14 (Itô's formula for Itô-processes). *Let $f \in C^{1,2}(\mathbb{R}_{\geq 0}\times\mathbb{R})$ and X be an Itô-process. Then*

$$\mathrm{d}f(t,X_t) = \partial_t f(t,X_t)\,\mathrm{d}t + \partial_x f(t,X_t)\,\mathrm{d}X_t + \frac{1}{2}\partial_{xx} f(t,X_t)\sigma_t^2\,\mathrm{d}t. \qquad \text{(Itô')}$$

Proof. For a complete proof see [1, Theorem 5.6]. Again, we will do a heuristic derivation based on the Taylor expansion and Itô-calculus. It holds that

$$\begin{aligned}f(t+\Delta t, x+\Delta x) - f(t,x) = \partial_t f(t,x)\cdot\Delta t + \partial_x f(t,x)\cdot\Delta x + \frac{1}{2}\partial_{xx} f(t,x)(\Delta x)^2\\ +\ \text{'higher order terms'}.\end{aligned}$$

Since $(\mathrm{d}X_t)^2 = \sigma_t^2\,\mathrm{d}t$, plugging in $\mathrm{d}X_t$ for x and passing to the limit yields the assertion

$$\mathrm{d}f(t,X_t) = \partial_t f(t,X_t)\,\mathrm{d}t + \partial_x f(t,X_t)\,\mathrm{d}X_t + \frac{1}{2}\partial_{xx} f(t,X_t)(\mathrm{d}X_t)^2.$$

□

0.2.5 Stochastic Integration with Respect to an Itô-Process

The stochastic integral

$$\int_0^t \vartheta_s \,\mathrm{d}X_s$$

can be reduced to a Lebesgue- and an Itô-Integral. Hence

$$\begin{aligned}\int_0^t \vartheta_s \,\mathrm{d}X_s &:= \int_0^t \vartheta_s(\mu_s \,\mathrm{d}s + \sigma_s \,\mathrm{d}B_s) \\ &= \int_0^t (\vartheta_s\mu_s)\,\mathrm{d}s + \int_0^t (\vartheta_s\sigma_s)\,\mathrm{d}B_s.\end{aligned}$$

Again in Itô-terms

$$\vartheta_t \,\mathrm{d}X_t = (\vartheta_t\mu_t)\,\mathrm{d}t + (\vartheta_t\sigma_t)\,\mathrm{d}B_t.$$

It is well-defined if

- $\int_0^t |\vartheta_s\mu_s| \,\mathrm{d}s < \infty$ for all $t \geq 0$ and
- $\vartheta \cdot \sigma \in \mathcal{H}^2(0,\infty)$.

0.3 The Multivariate Case

Definition 0.15. Consider a k-dimensional Brownian motion $B = (B^1, ..., B^k)$ and

$$\begin{aligned}&\mu\colon \mathbb{R}_{\geq 0} \times \Omega \to \mathbb{R}^d \text{ adapted,}\\ &\sigma\colon \mathbb{R}_{\geq 0} \times \Omega \to \mathbb{R}^{d\times k} \text{ adapted}\end{aligned}$$

such that

$$\int_0^t |\mu_i(s)| \,\mathrm{d}s < \infty \qquad \forall\, t \geq 0,\ i = 1, ..., d.$$

Define $\varrho(\omega,t) := \sigma(\omega,t)\sigma(\omega,t)^\top$ and note that $\varrho(\omega,t)$ is taking values in $\mathcal{S}_d$ [2]. If $\varrho_{ii} \in \mathcal{H}^2(0,\infty)$ for all $i = 1, ..., d$ then

$$X_t = X_0 + \int_0^t \mu_s \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}B_s, \qquad X_0 \in \mathbb{R}^d.$$

is called **multivariate Itô-process**. Note, the $(\sigma_s \,\mathrm{d}B_s)$-term is a matrix-vector-product.

Lemma 0.16. *Consider a multivariate Itô-process as in Definition 0.15. The* ***quadratic covariation*** *of two components of X, defined as*

$$[X^i, X^j]_t := \mathbb{P}\text{-}\lim_{|P_n|\to 0} \sum_{t_m \in P_n} \big(X^i_{t_m+1} - X^i_{t_m}\big)\big(X^j_{t_m+1} - X^j_{t_m}\big)$$

is

$$[X^i, X^j]_t = \int_0^t \varrho_{ij}(s)\,\mathrm{d}s.$$

[2] $\mathcal{S}_d = \{A \in \mathbb{R}^{d\times d} :\ A \text{ is semidefinite}\}$

Proof. By Itô-calculus it holds that

$$\begin{aligned}
(\mathrm{d}X_t^i)(\mathrm{d}X_t^j) &= \left(\mu_i(t)\,\mathrm{d}t + \sum_m \sigma_{im}(t)\,\mathrm{d}B_t^m\right)\left(\mu_j(t)\,\mathrm{d}t + \sum_l \sigma_{jl}(t)\,\mathrm{d}B_t^l\right) \\
&= \big[(\mathrm{d}t)^2\text{-terms, } (\mathrm{d}t)(\mathrm{d}B_t^m)\text{-terms}\big] + \sum_{m,l=1}^k \sigma_{im}(t)\sigma_{jl}(t)\underbrace{(\mathrm{d}B_t^m)(\mathrm{d}B_t^l)}_{=\delta_{ml}\cdot\mathrm{d}t} \\
&= \sum_{m=1}^k \sigma_{im}(t)\sigma_{jm}(t)\,\mathrm{d}t \\
&= \varrho_{ij}(t)\,\mathrm{d}t. \qquad \square
\end{aligned}$$

Hence

$$[X^i, X^j]_t = \int_0^t \varrho_{ij}(s)\,\mathrm{d}s.$$

Theorem 0.17 (Itô's formula for the multivariate Brownian motion)**.** *Let $f \in C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R}^k)$ and $X_t = f(t, B_t)$. Then*

$$\mathrm{d}f(t, B_t) = \partial_t f(t, B_t)\,\mathrm{d}t + \nabla f(t, B_t)\,\mathrm{d}B_t + \frac{1}{2}\Delta f(t, B_t)\,\mathrm{d}t,$$

where

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_k}\right) \in \mathbb{R}^k$$

*is the **gradient** and*

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \ldots + \frac{\partial^2 f}{\partial x_k^2} \in \mathbb{R}$$

*is the **Laplace operator**.*

Proof. Again by using the Taylor expansion it holds that

$$f(t+\Delta t, x+\Delta x) - f(t,x) = \partial_t f(t,x)\Delta t + \nabla f(t,x)\Delta x + \frac{1}{2}\sum_{i,j=1}^k \frac{\partial^2 f}{\partial x_i \partial x_j}(t,x)\Delta x_i \Delta x_j + \text{'higher order terms'}.$$

Plugging in B_t for x and passing to the limit yields

$$\mathrm{d}f(t, B_t) = \partial_t f(t, B_t)\,\mathrm{d}t + \nabla f(t, B_t)\,\mathrm{d}B_t + \frac{1}{2}\sum_{i,j=1}^k \frac{\partial^2 f}{\partial x_i \partial x_j}(t, B_t)\underbrace{(\mathrm{d}B_t^i)(\mathrm{d}B_t^j)}_{=\delta_{ij}\,\mathrm{d}t}.$$

Hence,

$$\mathrm{d}f(t, B_t) = \partial_t f(t, B_t)\,\mathrm{d}t + \nabla f(t, B_t)\,\mathrm{d}B_t + \frac{1}{2}\Delta f(t, B_t)\,\mathrm{d}t. \qquad \square$$

Theorem 0.18 (Itô's formula for multivariate Itô-processes)**.** *Let $f \in C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R}^k)$ and X_t a multivariate Itô-process. Then*

$$\mathrm{d}f(t, X_t) = \partial_t f(t, X_t)\,\mathrm{d}t + \nabla f(t, X_t)\,\mathrm{d}X_t + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2 f}{\partial x_i \partial x_j}(t, X_t)\varrho_{ij}\,\mathrm{d}t.$$

Proof. Again, by Taylor expansion, plugging in X_t and passing to the limit the assertion follows directly. □

An application of the multivariate Itô formula is the product rule for Itô-processes:

Theorem 0.19 (Product rule)**.** *Let (X, Y) be an Itô-process. Then*

$$\mathrm{d}(X_t Y_t) = X_t\,\mathrm{d}Y_t + Y_t\,\mathrm{d}X_t + \mathrm{d}[X, Y]_t.$$

In addition if X or Y has finite total variation then

$$\mathrm{d}(X_t Y_t) = X_t\,\mathrm{d}Y_t + Y_t\,\mathrm{d}X_t.$$

Proof. Consider the function $f(t, x, y) = x \cdot y$. Hence $\nabla f(x, y) = (y, x)$, $\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial y^2} = 0$ and $\frac{\partial^2 f}{\partial x \partial y} = 1$. Applying the multivariate Itô-formula yields

$$\begin{aligned}\mathrm{d}(X_t, Y_t) = \mathrm{d}f(t, X_t, Y_t) =& Y_t\,\mathrm{d}X_t + X_t\,\mathrm{d}Y_t \\ &+ \frac{1}{2}\big(0 \cdot \mathrm{d}[X, X]_t + 0 \cdot \mathrm{d}[Y, Y]_t + 1 \cdot \mathrm{d}[X, Y]_t + 1 \cdot \mathrm{d}[Y, X]_t\big) \\ =& Y_t\,\mathrm{d}X_t + X_t\,\mathrm{d}Y_t + \mathrm{d}[X, Y]_t.\end{aligned}$$

Since

$$[X, Y]_t = \int_0^t \sigma_s^X \sigma_s^Y\,\mathrm{d}s$$

the second assertion follows, because if X has finite variation then $\sigma^X = 0$ almost surely. □

0.4 Local Martingales ('The Ugly Truth')

Definition 0.20. a) A **stopping time** τ with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$ is a random variable taking values in $[0, \infty]$ with the property

$$\{\tau \leq t\} \in \mathcal{F}_t \qquad \forall\, t \geq 0.$$

b) Let X be an adapted stochastic process. The process X **stopped at** τ is defined as

$$X_t^\tau := X_{t\wedge\tau} = \begin{cases} X_t, & \text{if } t \leq \tau, \\ X_\tau, & \text{if } t > \tau. \end{cases}$$

Remark 0.21. By Doob's optional stopping theorem the martingale property is preserved under stopping, i.e.,

$$X \text{ is a martingale} \quad \Rightarrow \quad X^\tau \text{ is a martingale.}$$

Definition 0.22. A **localizing sequence** $(\tau_n)_{n\in\mathbb{N}}$ is an increasing sequence of stopping times ($\tau_n \leq \tau_{n+1}, ...$) such that

$$\lim_{n\to\infty} \tau_n = \infty \text{ almost surely.}$$

Definition 0.23. Let X be an adapted stochastic process and (τ_n) a localizing sequence. If

$$X^{\tau_n} \text{ are martingales } \forall\, n \in \mathbb{N} \tag{3}$$

then X is called **local martingale**.

Remark 0.24. Hence by Remark 0.21 if X is a martingale, then (3) does hold necessarily. On the other hand (3) does **not** imply that X is a martingale. However $X^{\tau n}$ converges locally to X, i.e.,

$$\lim_{n\to\infty} X_t^{\tau_n} = X_t.$$

0.4.1 Local Martingales and the Itô-Integral

Using a localizing sequence also the Itô-Integral can be generalized.

Lemma 0.25. *Let*

$$\mathcal{L}^2 := \left\{ \vartheta \textit{ adapted:} \quad \int_0^t \vartheta_s^2 \,\mathrm{d}s < \infty \;\; a.s.\; \forall\, t \geq 0 \right\}.$$

Then $\mathcal{H}^2(0,\infty) \subseteq \mathcal{L}^2$ *and each* $\mathcal{L}^2$*-function is locally in* $\mathcal{H}^2$*, i.e.*

$$\forall\, \vartheta \in \mathcal{L}^2 \;\exists\; \textit{a localizing sequence } (\tau_n) \textit{ s.t. } \vartheta(t,\omega)\mathbb{1}_{\{\tau(\omega)\leq t\}} \in \mathcal{H}^2(0,\infty).$$

Defining

$$\int_0^t \vartheta_s \,\mathrm{d}B_s := \lim_{n\to\infty} \int_0^{t\wedge\tau_n} \vartheta_s \,\mathrm{d}B_s = \lim_{n\to\infty} \int_0^t \underbrace{\vartheta_s \mathbb{1}_{\{s\leq\tau_n\}}}_{\in\mathcal{H}^2(0,\infty)} \mathrm{d}B_s$$

the Itô-integral is well-defined for all $\vartheta \in \mathcal{L}^2$.

Lemma 0.26 (Properties of the Itô-Integral)**.** *Let* $f \in \mathcal{L}^2$ *and*

$$X_t = \int_0^t f(\omega, s) \,\mathrm{d}B_s.$$

Then the following properties hold.

*(i) X is a continuous **local martingale**.*

*(ii) The Itô-integral is **linear**, i.e.,*

$$\int_0^t [cf(\omega,s)-g(\omega,s)]\,\mathrm{d}B_s = c\int_0^t f(\omega,s)\,\mathrm{d}B_s + \int_0^t g(\omega,s)\,\mathrm{d}B_s \qquad \forall\, c\in\mathbb{R},\ f,g\in\mathcal{L}^2.$$

*(iii) For the **quadratic variation** it holds that*

$$[X,X]_t = \int_0^t f^2(\omega,s)\,\mathrm{d}s.$$

Note that the Itô-isometry does not necessarily hold!

Also the definition of an Itô-process can be generalized:

Definition 0.27. Let μ such that

$$\int_0^t |\mu(\omega,s)|\,\mathrm{d}s < \infty$$

almost surely for all $t\geq 0$ and $\sigma\in\mathcal{L}^2$. Then the process

$$X_t = X_0 + \int_0^t \mu(\omega,s)\,\mathrm{d}s + \int_0^t \sigma(\omega,s)\,\mathrm{d}B_s$$

is called **Itô-process** and it is well-defined. In the multivariate case we need to require

$$\int_0^t \varrho_{ii}(\omega,s)\,\mathrm{d}s < \infty$$

almost surely for all $t\geq 0,\ i=1,...,d$.
Recall: ϱ is the covariation matrix $\varrho(\omega,t)=\sigma(\omega,t)\sigma(\omega,t)^\top$.

Remark 0.28.

- The Itô-integral with respect to an Itô-process $\int_0^t \vartheta_s\,\mathrm{d}X_s$ is well-defined if

  (1) $\int_0^t |\mu_s\vartheta_s|\,\mathrm{d}s < \infty \qquad \forall\, t\geq 0$ and

  (2) $\int_0^t (\sigma_s\vartheta_s)^2\,\mathrm{d}s < \infty \qquad \forall\, t\geq 0.$

  Hence (2) is equivalent to $\vartheta_t\in\mathcal{L}^2(X)$, where

  $$\mathcal{L}^2(X) := \left\{\vartheta \text{ adapted} : \int_0^t \vartheta_s^2\,\mathrm{d}[X,X]_s < \infty \text{ a.s. } \forall\, t\geq 0\right\}.$$

- The class of Itô-processes is closed under

  - stopping,
  - stochastic integration,
  - $C^{1,2}$-transformations.

- The class of local martingales is closed under
  - stopping,
  - stochastic integration.

Proposition 0.29 (from local to true martingales)**.** *Let M be a local martingale.*

a) If M is bounded, i.e., $|M_t| \le c < \infty$ for all $t \ge 0$. Then M is a true martingale.

b) If M is bounded from below, i.e., $M_t^- \le c < \infty$ for all $t \ge 0$ then M is a supermartingale.
If in addition $t \mapsto \mathbb{E}[M_t]$ is constant, then M is a martingale.

Proof. Exercise. □

1 Financial Markets in Continuous Time

1.1 Basics and Arbitrage

Definition 1.1. A **financial market model** (FMM) $\overline{S} = (S^0, S) = (S^0, S^1, ..., S^d)$ is a multivariate ($d+1$-dimensional) Itô-process of the form

$$\begin{aligned} \mathrm{d}S_t^0 &= S_t^0 r(t,\omega)\,\mathrm{d}t, & S_0^0 &= 1, \\ \mathrm{d}S_t^i &= S^i\big(\mu_i(t,\omega)\,\mathrm{d}t + \sigma_i(t,\omega)\,\mathrm{d}B_t\big), & S_0^i &> 0, \end{aligned} \tag{FMM}$$

where

- $r\colon \mathbb{R}_{\ge 0} \times \Omega \to \mathbb{R}$ is the **interest rate** (short rate),
- $\mu_i\colon \mathbb{R}_{\ge 0} \times \Omega \to \mathbb{R}$ are the **drift coefficients**,
- $\sigma_i\colon \mathbb{R}_{\ge 0} \times \Omega \to \mathbb{R}$ are the **volatility coefficients**.

Note: S^0 is the **locally risk-free asset** ('Numeraire') and it holds that

$$S_t^0 = \exp\left(\int_0^t r(s,\omega)\,\mathrm{d}s\right).$$

S_t^i is the price of **asset** i at time t.

Remark 1.2. Let B be a k-dimensional Brownian motion and

$$\Sigma := \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_d \end{pmatrix} \in \mathbb{R}^{d\times k}, \qquad \mu := \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_d \end{pmatrix} \in \mathbb{R}^{d}, \qquad \operatorname{diag}(S) := \begin{pmatrix} S^1 & & & \\ & S^2 & & \\ & & \ddots & \\ & & & S^d \end{pmatrix}.$$

Then the **matrix-vector-notation** of a financial market model is

$$\mathrm{d}S_t = \operatorname{diag}(S_t) \cdot \big(\mu(t,\omega)\,\mathrm{d}t + \Sigma(t,\omega)\,\mathrm{d}B_t\big).$$

1.2 The Pricing and Hedging Problem

Consider a european derivative $\Phi(S_T^i)$ or a more general claim C that is $\mathcal{F}_T$ measurable. Then the central questions are:

- What ist the fair price of C at time $t \in [0,T]$?
- What strategy replicates C?

To answer these questions the following Definitions are useful.

Definition 1.3. A **portfolio** or **strategy** $\overline{\vartheta} = (\vartheta^0, \vartheta) = (\vartheta^0, \vartheta^1, ..., \vartheta^d)$ is an $\mathbb{R}^{d+1}$-valued adapted process, where ϑ_t^i is the number of units of asset i in the portfolio at time t. ϑ_t^i is real-valued and can be negative ('short position').
The associated **value process**

$$V_t := \sum_{i=0}^{d} \vartheta_t^i \cdot S_t^i = \overline{\vartheta_t} \cdot \overline{S_t}$$

is the value of the portfolio at time t.

Definition 1.4. A portfolio is called **self-financing (SF)** if its value process V is an Itô-process of the form

$$\mathrm{d}V_t = \overline{\vartheta_t}\,\mathrm{d}\overline{S_t}.$$

Remark 1.5. Definition 1.4 implies that no capital is added or withdrawn from the portfolio after time $t = 0$. This is the continuous version of the discrete condition

$$\underbrace{V_{t+1} - V_t}_{\substack{\text{change in} \\ \text{portf. value}}} = \overline{\vartheta_t} \cdot \underbrace{(S_{t+1} - S_t)}_{\substack{\text{trading gains} \\ \text{or losses}}}.$$

Passing to limit leads to the condition above.

Definition 1.6. The **discounted price process** is defined as

$$X_t := \left(\frac{S_t^1}{S_t^0}, ..., \frac{S_t^d}{S_t^0}\right).$$

Analogously call

$$\widetilde{V}_t := \frac{V_t}{S_t^0}.$$

the **discounted value process**.

Lemma 1.7. *The discounted price process is an Itô-process such that*

$$\mathrm{d}X_t^i = X_t^i\big((\mu(t,\omega) - r(t,\omega))\,\mathrm{d}t + \sigma_i(t,\omega)\,\mathrm{d}B_t\big), \qquad i \in 1, ..., d.$$

Proof. Consider the function $f(x) = \frac{1}{x}$. Then $f'(x) = -\frac{1}{x^2}$.
Applying Itô's formula for Itô-processes to f yields

$$\mathrm{d}\left(\frac{1}{S_t^0}\right) = \mathrm{d}f(S_t^0) = -\frac{1}{(S_t^0)^2}\,\mathrm{d}S_t^0 = -\frac{1}{S_t^0} r(t,\omega)\,\mathrm{d}t.$$

The product rule applied to $X_t^i = S_t^i \cdot \frac{1}{S_t^0}$ leads to the assertion, i.e.

$$\begin{aligned}
\mathrm{d}X_t^i &= S_t^i\,\mathrm{d}\left(\frac{1}{S_t^0}\right) + \frac{1}{S_t^0}\,\mathrm{d}S_t^i + \cancel{\mathrm{d}\left[\frac{1}{S^0}, S^i\right]_t} \\
&= -\frac{S_t^i}{S_t^0} r(t,\omega)\,\mathrm{d}t + \frac{S_t^i}{S_t^0}\big(\mu(t,\omega)\,\mathrm{d}t + \sigma(t,\omega)\,\mathrm{d}B_t\big) \\
&= X_t^i\big((\mu(t,\omega) - r(t,\omega)\big)\,\mathrm{d}t + \sigma(t,\omega)\,\mathrm{d}B_t.
\end{aligned}$$

□

Lemma 1.8. *Let* $\vartheta \in \mathcal{L}^2(X)$*. Then* $\overline{\vartheta} = (\vartheta^0, \vartheta)$ *is self-financing if and only if*

$$\vartheta_t^0 = \int_0^t \vartheta_s\,\mathrm{d}X_s - \vartheta_t X_t. \tag{SF}$$

In this case, the discounted value process is given by

$$\widetilde{V}_t = V_0 + \int_0^t \vartheta_s\,\mathrm{d}X_s.$$

Proof. Exercise. □

Note that by Lemma 1.8 it is sufficient to specify ϑ, because ϑ^0 is determined uniquely.

Definition 1.9. A self-financing strategy $\overline{\vartheta} = (\vartheta^0, \vartheta)$ with value process V is called **arbitrage** if there exists $T > 0$, such that

(i) $V_0 = 0$,

(ii) $\mathbb{P}(V \geq 0) = 1$,

(iii) $\mathbb{P}(V > 0) > 0$.

"An arbitrage is too good to be true."

Theorem 1.10 ('All locally riskfree portfolios grow at the same rate')**.** *Let $\overline{\vartheta} = (\vartheta^0, \vartheta)$ be a self-financing portfolio with locally riskfree value processes, i.e., $V_0 > 0$ and*

$$\mathrm{d}V_t = V_t \cdot k(\omega, t)\,\mathrm{d}t.$$

In an arbitrage-free market it holds that

$$k(\omega, t) = r(\omega, t)$$

$\mathrm{d}\mathbb{P} \times \mathrm{d}t$*-almost surely.*

Proof. Let $\widetilde{V}^{\vartheta}$ be the discounted value process of ϑ. Then by Lemma 1.8

$$\mathrm{d}\widetilde{V}^{\vartheta} = \vartheta_t \cdot \mathrm{d}X_t.$$

On the other hand, using Itô's formula and the product rule it follows that

$$\begin{aligned} \mathrm{d}V_t^{\vartheta} = \mathrm{d}\left(\frac{V_t}{S_t^0}\right) &= V_t\,\mathrm{d}\left(\frac{1}{S_t^0}\right) + \frac{1}{S_t^0}\,\mathrm{d}V_t + \mathrm{d}\left[\frac{1}{S^0}, V\right]_t \\ &= -\frac{V_t}{S_t^0} r_t\,\mathrm{d}t + \frac{V_t}{S_t^0} k_t\,\mathrm{d}t + 0 \\ &= \widetilde{V}_t^{\vartheta}(k_t - r_t)\,\mathrm{d}t. \end{aligned} \qquad (\#)$$

Setting $\pi_t := \vartheta_t \cdot \mathbb{1}_{\{k_t > r_t\}}$ we form a self-financing portfolio with initial capital $\widetilde{V}_0^{\pi} = 0$ and discounted value process $\widetilde{V}_t^{\pi}$. Using again Lemma 1.8, (SF) and (#) yields

$$\mathrm{d}\widetilde{V}_t^{\pi} = \pi_t\,\mathrm{d}X_t = \vartheta_t \mathbb{1}_{\{k_t > r_t\}}\,\mathrm{d}X_t = \mathbb{1}_{\{k_t > r_t\}}\,\mathrm{d}\widetilde{V}_t^{\vartheta} = \mathbb{1}_{\{k_t > r_t\}} \widetilde{V}_t^{\vartheta}(k_t - r_t)\,\mathrm{d}t.$$

In other notation and with $N = \{(\omega, t) : \; k(\omega, t) > r(\omega, t)\}$ we have

$$\widetilde{V}_t^{\pi} = \int_0^t \underbrace{\mathbb{1}_{\{k_s > r_s\}}(k_s - r_s)\widetilde{V}_s^{\vartheta}}_{\geq 0, \text{ and } > 0 \text{ for } (s,\omega) \text{ in } N}\,\mathrm{d}s.$$

Assuming that there is no arbitrage we conclude from $\widetilde{V}_0^{\pi} = 0$ and $\mathbb{P}(\widetilde{V}_T^{\pi} \geq 0) = 1$ that $\mathbb{P}(\widetilde{V}_T^{\pi} > 0) = 0$. This implies

$$\begin{aligned} 0 &= \mathbb{E}\big[\widetilde{V}_T^{\pi}\big] \\ &= \int_{\Omega} \int_0^{\pi} \mathbb{1}_{\{k(s,\omega) > r(s,\omega)\}} \cdot (k(s, \omega) - r(s, \omega))\widetilde{V}_s^{\vartheta}(\omega)\,\mathrm{d}s\,\mathrm{d}\mathbb{P}. \end{aligned}$$

Hence N is a $\mathrm{d}\mathbb{P} \times \mathrm{d}t$-Null set and thus $k(t, \omega) \leq r(t, \omega)$ almost everywhere. Repeating the same argument for

$$\pi_t' = -\vartheta_t \mathbb{1}_{\{k_t < r_t\}}$$

yields $k(t, \omega) \geq r(t, \omega)$ $(\mathrm{d}\mathbb{P} \times \mathrm{d}t)$-almost everywhere. Thus the proof is complete. □

1.3 Classification of One-Dimensional Market Models

Specializing to a single asset and a constant interest rate, i.e., $d = 1$ and $r(\omega, t) \equiv r$ we reduce (FMM) to

$$\mathrm{d}S_t^0 = rS_t^0 \,\mathrm{d}t,$$
$$\mathrm{d}S_t = S_t\big(\mu(t,\omega)\,\mathrm{d}t + \Sigma(t,\omega)\,\mathrm{d}B_t\big).$$

The crucial parameter is now the volatility coefficient $\Sigma(t, \omega)$ that can be estimated from historic data by

$$\int_0^t \Sigma^2(s,\omega)\,\mathrm{d}s := [\log S, \log S]_t = \mathbb{P}\text{-}\lim_{|P_n|\downarrow 0} \underbrace{\sum_{t_i \in P_n} (\log S_{t_{i+1}} - \log S_{t_i})^2}_{\text{'realized variance'}}.$$

This definition is reasonable, since for $[S,S]_t = \int_0^t S_r^2 \Sigma(r,\omega)^2\,\mathrm{d}r$ and $f(x) = \log x$[3] follows by Itô's formula

$$\begin{aligned}
\mathrm{d}\log S_t &= \frac{1}{S_t}\,\mathrm{d}S_t - \frac{1}{2S_t^2}\,\mathrm{d}[S,S]_t \\
&= \big(\mu(t,\omega)\,\mathrm{d}t + \Sigma(t,\omega)\,\mathrm{d}B_t\big) - \frac{1}{2}\Sigma(t,\omega)^2\,\mathrm{d}t \\
&= \left(\left(\mu(t,\omega) - \frac{\Sigma^2(t,\omega)}{2}\right)\mathrm{d}t + \Sigma(t,\omega)\,\mathrm{d}B_t\right)
\end{aligned}$$

and thus (see Lemma 0.7 (v))

$$[\log S, \log S]_t = \int_0^t \Sigma^2(r,\omega)\,\mathrm{d}r.$$

Models can be classified according to $\Sigma(t,\omega)$, e.g. we have for

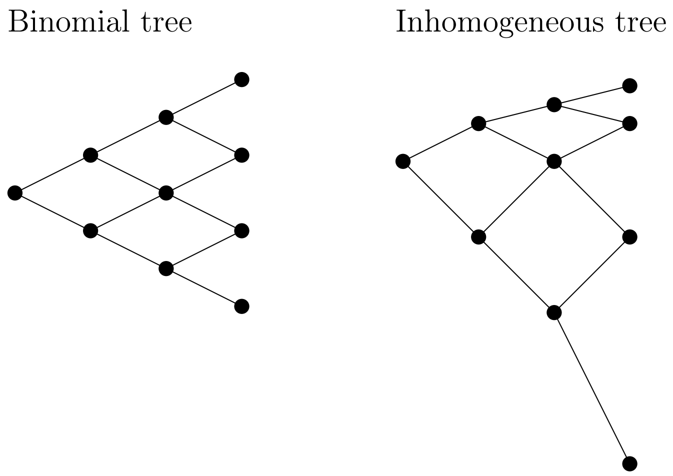
- $\Sigma(t,\omega) = \sigma \in (0,\infty)$ the Black-Scholes model,
- $\Sigma(t,\omega) = \sigma(t, S_t),\ \sigma \in C^{1,2}$ a local volatility model,
- $\Sigma(t,\omega) = \sigma(V_t)$ with another stochastic process V, a stochastic volatility model.

From top to bottom the listed models become more complex but also more realistic.

[3] Hence, $f'(x) = \frac{1}{x}$ and $f''(x) = -\frac{1}{x^2}$.

	pricing of puts/calls	market completeness	discrete analogue	Examples
BS-model	explicit formula	complete	CRR-model, binomial tree	—
local volatility	1-dimensional PDE/Monte-Carlo	complete	inhomogeneous tree	CEV (constant elasticity of variance), Dupire model
stoch. volatility	2-dimensional PDE	incomplete	'stochastic tree'	Heston model, SABR model

The difference between a binomial and an inhomogeneous tree:



2 Local Volatility Models

Definition 2.1. A financial market model of the form

$$\begin{aligned} \mathrm{d}S_t^0 &= rS_t^0\,\mathrm{d}t, \\ \mathrm{d}S_t &= S_t\big(\mu(t,\omega)\,\mathrm{d}t + \sigma(t,S_t)\,\mathrm{d}B_t\big). \end{aligned} \tag{LV}$$

with $\sigma\colon \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is called **local volatility model**. σ is called **local volatility function**.

Remark 2.2. Local volatility models are complete. Hence, perfect replication is possible and derivative prices are unique.

Consider a European Claim with payoff $h(S_T)$. The goal is to derive a replicating portfolio (ϑ^0, ϑ) for the claim. To reach this goal let V_t be a value process, such that

$$V_t = \vartheta^0 S_t^0 + \vartheta S_t. \tag{4}$$

Note that $V_t = f(t, S_t)$ with $f \in C^{1,2}$. The self-financing condition for V_t is

$$\begin{aligned} \mathrm{d}V_t &= \vartheta^0 \,\mathrm{d}S_t^0 + \vartheta \,\mathrm{d}S_t \\ &= \big(\vartheta^0 r S_t^0 + \vartheta \mu(t,\omega) S_t\big)\,\mathrm{d}t + \vartheta \sigma(t, S_t) S_t \,\mathrm{d}B_t. \end{aligned} \tag{5}$$

On the other hand, applying Itô's formula to $V_t = f(t, S_t)$ yields

$$\begin{aligned} \mathrm{d}V_t = \mathrm{d}f(t, S_t) &= \partial_t f \,\mathrm{d}t + \partial_s f \,\mathrm{d}S_t + \frac{1}{2}\partial_{ss} f \,\mathrm{d}[S,S]_t \\ &= \left(\partial_t f + \partial_s f S_t \mu(t,\omega) + \frac{1}{2}\partial_{ss} f S_t^2 \sigma(t, S_t)^2\right)\mathrm{d}t + \partial_s f S_t \sigma(t, S_t)\,\mathrm{d}B_t. \end{aligned} \tag{6}$$

Comparing coefficients of (5) and (6) leads to the following equations

$\mathrm{d}B_t$-terms: $\quad \vartheta_t \sigma(t, S_t) S_t = \partial_s f \sigma(t, S_t) S_t$

$$\implies \vartheta_t = \partial_s f(t, S_t) \qquad \text{'Delta hedging'}$$

$$\mathrm{d}t\text{-terms:} \quad \vartheta_t^0 r S_t^0 + \cancel{\vartheta_t \mu(t,\omega) S_t} = \partial_t f + \cancel{\partial_s f S_t \mu(t,\omega)} + \frac{1}{2}\partial_{ss} f S_t^2 \sigma^2(t, S_t)$$

$$\implies \vartheta_t^0 = \frac{1}{r S_t^0}\left(\partial_t f + \frac{1}{2}\partial_{ss} f S_t^2 \sigma^2(t, S_t)\right)$$

By plugging these results into (4) we obtain

$$f(t, S_t) = V_t = \frac{1}{r}\left(\partial_t f + \frac{1}{2}\partial_{ss} f S_t^2 \sigma^2(t, S_t) + r S_t \partial_s f(t, S_t)\right)$$

which is equivalent to

$$\partial_t f + r S_t \partial_s f + \frac{1}{2} S_t^2 \sigma^2(t, S_t) \partial_{ss} f - r f = 0$$

for all possible points $S_t(\omega) \in \mathbb{R}_{\geq 0}$. Replacing $S_t(\omega)$ by $s \in \mathbb{R}_{\geq 0}$ yields the pricing PDE

$$\partial_t f + r s \partial_s f + \frac{1}{2} s^2 \sigma^2(t, s) \partial_{ss} f - r f = 0 \tag{7}$$

with terminal condition $f(T, s) = h(s)$. This is a parabolic PDE, related to heat equation. We formulate this reasoning as a theorem:

Theorem 2.3. *Consider a European Claim with payoff $h(S_T)$ in a local volatility model. Let $f \in C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0})$ be a solution of the pricing PDE*

$$\begin{cases} \partial_t f + r s \partial_s f + \dfrac{\sigma^2(t,s)}{2} s^2 \partial_{ss} f - r f = 0, \\ \qquad\qquad\qquad\qquad (t, s) \in [0, T] \times \mathbb{R}_{\geq 0} \end{cases}$$

with terminal condition

$$f(T,s) = h(s).$$

Then the portfolio

$$\begin{cases} \vartheta_t = \partial_s f(t, S_t), \\ \vartheta_t^0 = \dfrac{1}{S_t^0}\left(f(t,S_t) - S_t \partial_s f(t,S_t)\right) \end{cases}$$

is a self-financing replicating portfolio for $h(S_t)$ and

$$\Pi_t = f(t, S_t)$$

is the price of the claim $h(S_T)$ at time $t \in [0,T]$.

Remark 2.4.

- The Drift $\mu(t,\omega)$ does not appear in the PDE.
- There is no statement on existence and uniqueness of a solution to the PDE.

Example 2.5.

- In the Black-Scholes-Model the function σ is a constant, i.e.,

$$\sigma(t, S_t) = \sigma > 0.$$

- Let $\sigma(t, S_t) = a \cdot S_t^{\beta}$ with $\beta \in [-1, 0]$. Then the model is called **constant elasticity of variance (CEV) model**. 'Constant elasticity', because

$$\frac{\mathrm{d}\sigma}{\mathrm{d}S} = \beta \cdot \frac{\sigma}{S}.$$

  The CEV model is useful, because for $\beta < 0$ it models the **leverage effect**, i.e., that when prices go up the volatility goes down and vice versa.

- In the **Dupire model** $\sigma(t, S_t)$ is fitted to observed put and call prices.

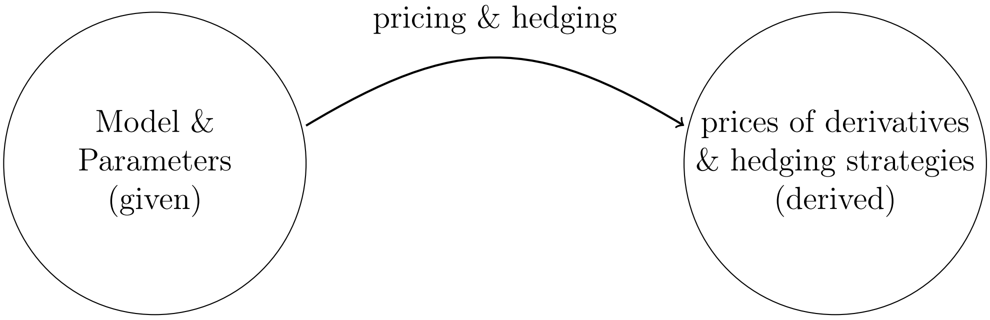
Exercise. Derive the Black-Scholes formula for puts/calls from the pricing PDE.
Hint: Use the transformations

$$\begin{aligned} \tau &= T - t, \\ x &= \log s - r\tau, \\ g(\tau, x) &= e^{r\tau} f(t, s) \end{aligned}$$

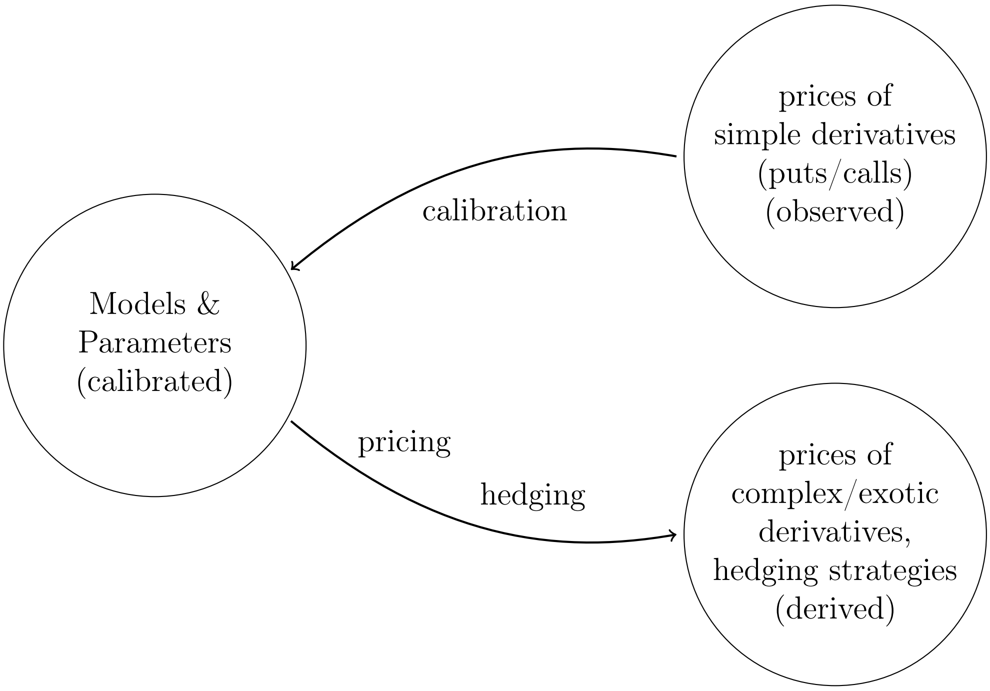
where $f(t,s)$ solves Black-Scholes-PDE.

2.1 Dupire's and Tanaka's Formula

Reason for the Dupire model is that the mathematical point of view is a very simplified one, i.e.,



The Dupire model has a more realistic point of view, i.e.



Note that calibration is the inverse problem of pricing, which - in the context of local volatility models - is solved by Dupire's formula. The formula derives a local volatility function $\sigma(t, S_t)$ from given prices $C(T, K)$ of European calls with maturity $T \in [0, H]$ and strikes $K \in [0, \infty)$.
To derive the formula there is some stochastic calculus needed:

2.1.1 Tanaka's formula

The goal is to derive a generalization of Itô's formula to apply it on $|B_t|$ (or more general on $|X_t|$, where X is an Itô-process). Since $|\cdot|$ is not differentiable in 0 the 'normal' Itô formula is not applicable. Observe that although $f(x) := |x|$ is not differentiable in 0 it is smooth at every other point. Furthermore

$$
\begin{aligned}
f(y) - f(x) &= \int_x^y \operatorname{sign}(\eta)\,\mathrm{d}\eta, \\
\operatorname{sign}(y) - \operatorname{sign}(x) &= \int_x^y 2\delta_0(\eta)\,\mathrm{d}\eta.
\end{aligned}
$$

Thus in a distributional[4] sense it holds that[5]

$$
\begin{aligned}
f'(x) &= \operatorname{sign}(x), \\
f''(x) &= 2\delta_0(x)
\end{aligned}
$$

and the conjecture

$$
|X_t| = |X_0| + \int_0^t \operatorname{sign}(X_s)\,\mathrm{d}X_s + \text{"correction at zero"}
$$

is reasonable. Tanaka's theorem now states that the "correction" $L_t^0(X)$ is well-defined and called **local time at zero of** X.

Theorem 2.6 (Tanaka's formula)**.** *Let X be an Itô-process. Then*

$$
|X_t| = |X_0| + \int_0^t \operatorname{sign}(X_s)\,\mathrm{d}X_s + L_t^0(X),
$$

where the local time at zero $L_t^0(X)$ is given by

$$
L_t^0(X) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{(-\varepsilon,\varepsilon)}(X_s)\,\mathrm{d}[X,X]_s \tag{8}
$$

(in $L_2(\mathrm{d}\mathbb{P})$).

Remark 2.7.

- The local time $L_t^y(X)$ at $y \in \mathbb{R}$ is defined in the same way, replacing $(-\varepsilon,\varepsilon)$ by $(y-\varepsilon, y+\varepsilon)$ in (8).

- In differential notation we have

$$
\mathrm{d}\,|X_t| = \operatorname{sign}(X_t)\,\mathrm{d}X_t + \mathrm{d}L_t.
$$

[4]In german: Distribution, NICHT Verteilung

[5]$\operatorname{sign}(x) := \begin{cases} +1, & x \geq 0, \\ -1, & x < 0, \end{cases}$ δ_0 is the dirac function at zero.

- Let

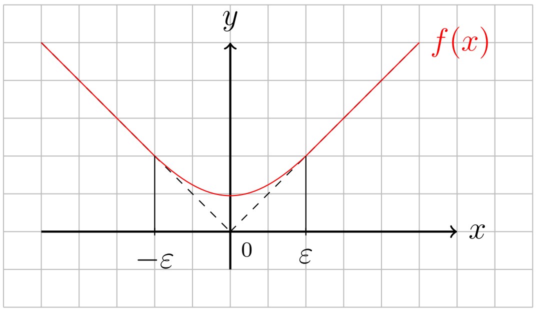
$$X^+ = \frac{1}{2}(|X| + X),$$
$$X^- = \frac{1}{2}(|X| - X).$$

Then

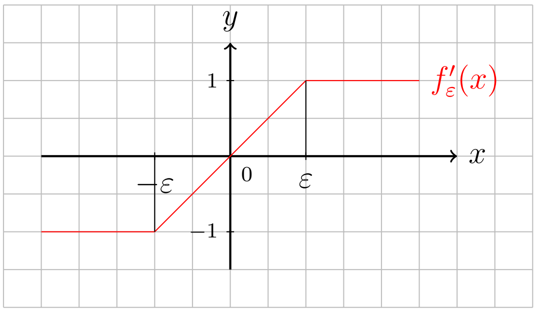
$$\mathrm{d}(X_t^+) = \mathbb{1}_{(X_t \geq 0)} \, \mathrm{d}X_t + \frac{1}{2} L_t^0(X),$$
$$\mathrm{d}(X_t^-) = \mathbb{1}_{(X_t \leq 0)} \, \mathrm{d}X_t + \frac{1}{2} L_t^0(X).$$

Proof. We will only do a sketch of the proof for the case $X_t = B_t$. We try to smoothen $f(x) := |x|$ by setting

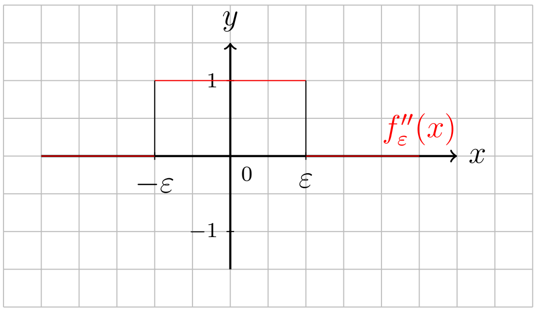
$$f_\varepsilon(x) = \begin{cases} |x|, & |x| \geq \varepsilon, \\ \frac{1}{2}\left(\varepsilon + \frac{x^2}{\varepsilon}\right), & |x| < \varepsilon, \end{cases}$$



$$f'_\varepsilon(x) = \begin{cases} \operatorname{sign}(x), & |x| \geq \varepsilon, \\ \frac{x}{\varepsilon}, & |x| < \varepsilon, \end{cases}$$



$$f''_\varepsilon(x) = \mathbb{1}_{(-\varepsilon,\varepsilon)}(x) \cdot \frac{1}{\varepsilon}.$$



By Itô's formula it holds that

$$f_\varepsilon(B_t) = \frac{\varepsilon}{2} + \underbrace{\int_0^t f'_\varepsilon(B_s) \, \mathrm{d}B_s}_{=I_0} + \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{(-\varepsilon,\varepsilon)}(B_s) \, \mathrm{d}B_s. \tag{9}$$

Note that $I_\varepsilon \mapsto \int_0^t \operatorname{sign}(B_s)\,\mathrm{d}B_s$ is in $\mathcal{L}^2(\mathrm{d}\mathbb{P})$, because the Itô-isometry yields

$$\mathbb{E}\left[\left(\int_0^t f_\varepsilon'(B_s)\,\mathrm{d}B_s - \int_0^t \operatorname{sign}(B_s)\,\mathrm{d}B_s\right)^2\right] = \int_0^t \mathbb{E}\left[\Bigg(\underbrace{\frac{B_s}{\varepsilon} - \operatorname{sign}(B_s)}_{|\cdot|\leq 1}\Bigg)^2 \cdot \mathbb{1}_{(-\varepsilon,\varepsilon)}(B_s)\right] \mathrm{d}s$$
$$\leq t \sup_{0\leq s\leq t} \mathbb{P}(B_s \in (-\varepsilon,\varepsilon))$$
$$\leq t^{\frac{3}{2}} \underbrace{[\Phi(\varepsilon) - \Phi(-\varepsilon)]}_{\to 0 \text{ as } \varepsilon\to 0}.$$

Therefore and by (9) it follows that

$$\begin{aligned} L_t^0(B) &= \lim_{\varepsilon\downarrow 0} \frac{1}{2\varepsilon}\int_0^t \mathbb{1}_{(-\varepsilon,\varepsilon)}(B_s)\,\mathrm{d}s \\ &= \lim_{\varepsilon\downarrow 0}\left(f_\varepsilon(B_t) - \frac{\varepsilon}{2} - \int_0^t f_\varepsilon'(B_s)\,\mathrm{d}B_s\right) \\ &= |B_t| - \int_0^t \operatorname{sign}(B_s)\,\mathrm{d}B_s \end{aligned}$$

in $\mathcal{L}^2(\mathrm{d}\mathbb{P})$. Note that it also follows that $t \mapsto L_t^0(B)$ is continuous up to modification. □

Remark 2.8. Here are some curious properties of $L_t^0(B)$.

- $t \mapsto L_t^0(B)$ is almost surely increasing and continuous.
- $t \mapsto L_t^0(B)$ is constant on the random set

$$K(\omega) = \{t \in \mathbb{R}_{\geq 0} :\ B_t(\omega) \neq 0\}.$$

- $K(\omega)$ has full Lebesgue measure with probability one, i.e.,

$$\mathbb{P}\left(\int_0^t \mathbb{1}_{K(\omega)}(s)\,\mathrm{d}s = t\right) = 1.$$

This implies, that L_t^0 is almost surely a "singular continuous" function, i.e., a continuous function, which is not absolutely continuous. A deterministic example of such a function is given by the "devil's staircase".

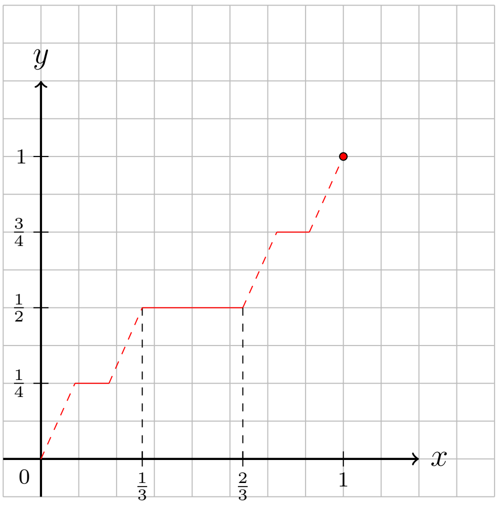


Figure 1: Devil's staircase

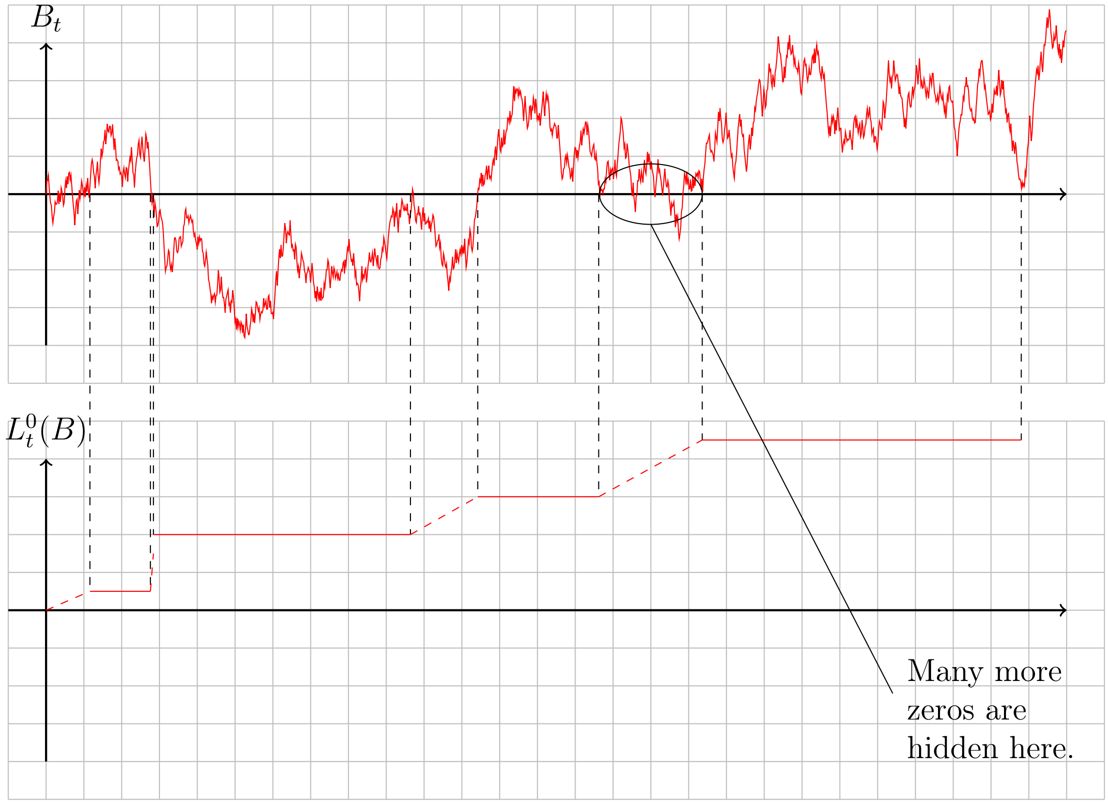


Figure 2: Local time at zero of the Brownian motion

2.1.2 The Dupire Model

Theorem 2.9 (Dupire's formula)**.** *In a local volatility model*

$$\begin{cases} \mathrm{d}S_t^0 = rS_t^0\,\mathrm{d}t, \\ \mathrm{d}S_t = S_t\big(\mu(t,\omega)\,\mathrm{d}t + \sigma(t,S_t)\,\mathrm{d}B_t\big) \end{cases}$$

let $C(T,K)$ be prices of calls with maturity $T \in [0,H)$ and strike $K \in [0,\infty)$. Assume $C(T,K) \in C^{1,2}$ and $\frac{\partial}{\partial T}C(T,K) \geq 0$, $\frac{\partial^2}{\partial K^2}C(T,K) > 0$. Then the local volatility function must be given by

$$\sigma^2(T,K) = -2 \cdot \frac{\frac{\partial}{\partial T}C(T,K) + \sigma K \frac{\partial}{\partial K}C(T,K)}{K^2 \frac{\partial^2}{\partial K^2}C(T,K)}$$

for all $T \in [0,H)$ and $K \in [0,\infty)$.

Proof. We use the discounted pricing formula

$$C(T,K) = \mathbb{E}\left[e^{-rT}(\widetilde{S_T} - K)_+\right]$$

with

$$\mathrm{d}\widetilde{S}_t = \widetilde{S}_t(r\,\mathrm{d}t + \sigma(t,\widetilde{S}_t)\,\mathrm{d}B_t).$$

Hence, $\widetilde{S}_t$ is the price process under the risk free measure. Setting $S_t = \widetilde{S}_t$ yields

$$\begin{aligned}
\frac{\partial}{\partial K}C(T,K) &= -\mathbb{E}\left[e^{-rT}\mathbb{1}_{\{S_t\geq K\}}\right],\\
\frac{\partial^2}{\partial K^2}C(T,K) &= -\frac{\partial}{\partial K}\mathbb{E}\left[e^{-rT}\mathbb{1}_{\{S_t\geq k\}}\right]\\
&= -\lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\mathbb{E}\left[e^{-rT}\mathbb{1}_{\{S_t\in(K-\varepsilon,K+\varepsilon)\}}\right].
\end{aligned}$$

On the other hand it holds that

$$\mathrm{d}(S_t-K)_+ = \mathbb{1}_{\{S_t>k\}}\,\mathrm{d}S_t + \frac{1}{2}L_t^K(S).$$

Set $g(t,K) := e^{-rT}(S_t-K)_+$. From the product rule we obtain

$$\begin{aligned}
\mathrm{d}g(t,K) &= -rg(t,K)\,\mathrm{d}t + e^{-rt}\,\mathrm{d}(S_t-K)_+\\
&= e^{-rt}\left(-r(S_t-K)_+\,\mathrm{d}t + \mathbb{1}_{\{S_t>K\}}\,\mathrm{d}S_t + \frac{1}{2}L_t^K(S)\right)\\
&= e^{-rt}\left(r\cdot K\mathbb{1}_{\{S_t\geq K\}}\,\mathrm{d}t + \mathbb{1}_{\{S_t\geq K\}}\sigma(t,S_t)S_t\,\mathrm{d}B_t + \frac{1}{2}L_t^K(S)\right).
\end{aligned}$$

Now it follows that

$$\begin{aligned}
C(T,K) =& \mathbb{E}\left[e^{-rT}(S_t-K)_+\right]\\
=& \mathbb{E}\left[rK\int_0^T e^{-rt}\mathbb{1}_{\{S_t\geq K\}}\,\mathrm{d}t\right] + \mathbb{E}\left[\int_0^T e^{-rT}\mathbb{1}_{\{S_t\geq K\}}\sigma(t,S_t)S_t\,\mathrm{d}B_t\right]\\
&+ \mathbb{E}\left[\frac{1}{2}\int_0^T e^{-rt}\,\mathrm{d}L_t^K(S)\right]\\
=& rK\int_0^T \underbrace{\mathbb{E}\big[e^{-rt}\mathbb{1}_{\{S_t\geq K\}}\big]}_{=-\frac{\partial}{\partial K}C(t,K)}\mathrm{d}t + \frac{1}{2}\lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\int_0^T \mathbb{E}\big[e^{-rt}\mathbb{1}_{(K-\varepsilon,K+\varepsilon)}(S_t)\big]\,\mathrm{d}[S,S]_t\\
=& rK\int_0^T -\frac{\partial}{\partial K}C(t,K)\,\mathrm{d}t + \frac{1}{2}\int_0^T \underbrace{\lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\mathbb{E}\left[e^{-rt}\mathbb{1}_{(K-\varepsilon,K+\varepsilon)}(S)\right]}_{=-\frac{\partial^2}{\partial K^2}C(t,K)}\cdot\underbrace{S_t^2\sigma^2(t,S_t)}_{=K^2\sigma^2(t,K)}\,\mathrm{d}t\\
=& -\left[rK\int_0^T \frac{\partial}{\partial K}C(t,K)\,\mathrm{d}t + \frac{1}{2}\int_0^T \frac{\partial^2}{\partial K^2}C(t,K)K^2\sigma^2(t,K)\,\mathrm{d}t\right]
\end{aligned}$$

and with

$$\frac{\partial}{\partial T}C(T,K) = -\left[rK\frac{\partial}{\partial K}C(T,K) + \frac{1}{2}\frac{\partial^2}{\partial K^2}C(T,K)K^2\sigma^2(t,K)\right]$$

we obtain the assertion

$$\sigma^2(t,K) = -2\cdot\frac{\frac{\partial}{\partial T}C(T,K) + rK\frac{\partial}{\partial K}C(T,K)}{\frac{\partial^2}{\partial K^2}C(T,K)K^2}. \qquad \square$$

3 Stochastic Volatility Models

Definition 3.1. Let (B^1, B^2) be a 2-dimensional Brownian motion. Choose $\varrho \in [-1, 1]$ and set

$$W_t := \varrho B_t^1 + \sqrt{1 - \varrho^2} B_t^2.$$

Then $\mathrm{d}[B^1, W]_t = \varrho \,\mathrm{d}t$ and B^1 and W are called **correlated Brownian motions** with correlation ϱ.

Definition 3.2. A financial market model of the form

$$\begin{cases} \mathrm{d}S_t^0 = rS_t^0 \,\mathrm{d}t, \\ \mathrm{d}S_t = S_t\big(\mu(V_t, S_t)\,\mathrm{d}t + m(V_t)\,\mathrm{d}B_t\big), \\ \mathrm{d}V_t = a(V_t)\,\mathrm{d}t + b(V_t)\,\mathrm{d}W_t \end{cases}$$

with $\mu\colon \mathbb{R}^2 \to \mathbb{R}$, $m, a, b\colon \mathbb{R} \to \mathbb{R}$ and B, W correlated Brownian motions is called **stochastic volatility model**. The volatility V_t is a non traded quantity. Examples for m are

$$m(V_t) = \sqrt{V_t}, \quad m(V_t) = V_t \quad \text{or} \quad m(V_t) = |V_t|\,.$$

Remark 3.3.

- Volatility is a stochastic process.
- $[S, S]_t$ can not be predicted using S_t alone as in local volatility models.
- The volatility V_t and the stochastic prices S_t are usually negatively correlated, i.e., $\varrho \leq 0$.[6]
- Stochastic volatility models are generally not complete, which implies that there is neither perfect replication nor a unique price.

However, we can make statements about price relations of various derivatives and instead of replication we can construct a locally riskfree portfolio.

Definition 3.4. For notational simplicity we define the generator

$$\mathcal{A} := a(v)\partial_v + \frac{1}{2}m(v)^2 s^2 \partial_{ss} + \varrho m(v) s b(v) \partial_{sv} + \frac{1}{2} b(v)^2 \partial_{vv} + rs\partial_s.$$

Theorem 3.5. *Consider a stochastic volatility model and N derivatives H^i with payoff $\Phi^i(S_T)$ and price processes of the form*

$$\pi_t^i = f^i(t, S_t, V_t),$$

with $f \in C^{1,2,2}$ and $i \in \{2, ..., N+1\}$. Let the extended market $(S, \pi^2, \pi^3, ..., \pi^N)$ be arbitrage free. Then there exists a function $\lambda\colon \mathbb{R}^3 \to \mathbb{R}$ called ***market price of volatility***

[6] Typically $\varrho \approx -0.7$.

risk (MPVR) *and we have the following pricing equations. Namly for all* $i \in \{2, ..., N+1\}$ *it holds that*

$$\partial_t f^i + \mathcal{A} f^i + \lambda(t,s,v) b(v) \partial_v f^i - r f^i = 0 \tag{10}$$

with boundary condition

$$f^i(T,s,v) = \Phi^i(s) \qquad \forall\ (s,v) \in \mathbb{R}^2_{\geq 0},\ i = 2, ..., N+1.$$

Remark 3.6.

- The market price of risk λ is the same for all derivatives. Therefore once λ is fixed the prices of all derivatives are unique.
- In practice λ is obtained from the market via calibration, i.e., if some prices are known we can determine λ to compute other arbitrage-free prices.

Proof. For notational simplicity set $r = 0$ and $\pi^1 = S^1$, i.e., $S^0 \equiv 1$ and

$$\pi = (\pi^1, \pi^2, ..., \pi^{N+1}) = (S^1, \pi^2, ..., \pi^{N+1}).$$

Applying Itô's formula to $\pi^i = f^i(t, S_t, V_t)$ yields[7]

$$\begin{aligned} \mathrm{d}\pi^i_t =& \partial_t f^i \,\mathrm{d}t + \partial_s f^i \,\mathrm{d}S_t + \partial_v f^i \,\mathrm{d}V_t \\ &+ \frac{1}{2}\partial_{ss} f^i \underbrace{m(V_t)^2 S_t^2}_{=[S,S]_t} \mathrm{d}t + \frac{1}{2}\partial_{sv} f^i \underbrace{m(V_t) S_t b(V_t) \varrho}_{=[S,V]_t} \mathrm{d}t + \frac{1}{2}\partial_{vv} f^i \underbrace{b(V_t)^2}_{=[V,V]_t} \\ =& \alpha_t \,\mathrm{d}t + \partial_s f S_t m(V_t) \,\mathrm{d}B_t + \partial_v f^i b(V_t) \,\mathrm{d}W_t \end{aligned} \tag{11}$$

with

$$\alpha_t := \partial_t f^i + S_t \mu_t \partial_s f^i + a(V_t)\partial_v f^i + \frac{1}{2} m(V_t)^2 S_t^2 \partial_{ss} f^i + \varrho m(V_t) S_t b(V_t) \partial_{sv} f^i + \frac{1}{2} b(V_t)^2 \partial_{vv} f^i.$$

In the next step we will construct a self-financing, locally risk-free[8] portfolio

$$\vartheta_t = (\vartheta^1_t, ..., \vartheta^{N+1}_t).$$

Comparing with (11) we need to find a strategy ϑ such that

$$\begin{pmatrix} m(V_t) & \partial_s f^2 m(V_t) S_t & \cdots & \partial_s f^{N+1} m(V_t) S_t \\ 0 & \partial_v f^2 b(V_t) & \cdots & \partial_v f^{N+1} b(V_t) \end{pmatrix} \vartheta_t = 0. \tag{12}$$

[7]Recall that

$$\begin{aligned} \mathrm{d}S_t &= S_t\big(\mu(V_t, S_t)\,\mathrm{d}t + m(V_t)\big)\,\mathrm{d}B_t, \\ \mathrm{d}V_t &= a(V_t)\,\mathrm{d}t + b(V_t)\,\mathrm{d}W_t. \end{aligned}$$

[8]Risk-free means that $\mathrm{d}V_t = V_t K(t,\omega)\,\mathrm{d}t$, i.e., that there are no $(\mathrm{d}B_t)$-terms.

In this case we have

$$\mathrm{d}V_t = \left(\vartheta_t^1 S_t + \sum \vartheta_t^i \alpha_t^i\right) \mathrm{d}t$$

and therefore V_t is locally riskfree. Applying Theorem 1.10 yields

$$\vartheta_t^1 S_t \mu_t + \sum_{i=1}^{N+1} \vartheta_t^i \alpha_t^i = 0.$$

Hence, from (12) we obtain

$$M_t \begin{pmatrix} \vartheta_t^1 \\ \cdots \\ \vartheta_t^{N+1} \end{pmatrix} = 0. \tag{13}$$

where

$$M_t := \begin{pmatrix} Z_t^1 \\ Z_t^2 \\ Z_t^3 \end{pmatrix} := \begin{pmatrix} m(V_t)S_t & \partial_s f^2 m(V_t) S_t & \cdots & \cdots & \partial_s f^{N+1} m(V_t) S_t \\ 0 & \partial_v f^2 b(V_t) & \cdots & \cdots & \partial_v f^{N+1} b(V_t) \\ \mu_t & S_t & \alpha_t^2 & \cdots & \alpha_t^{N+1} \end{pmatrix}.$$

Obviously (12) and (13) imply that $\mathrm{rank}(M_t) \leq 2$. Thus, for all $t \in [0,T]$ there exists a non-trivial linear combination of Z_t^1, Z_t^2, Z_t^3 that equals zero. That means there exists a process λ_t with

$$-\frac{\mu_t}{m(V_t)} Z_t^1 + Z_t^3 + \lambda_t Z_t^2 = 0 \qquad \forall\, t \in [0,T].$$

This implies that for all $i \in \{1, ..., N+1\}$

$$-\mu_t S_t \partial_s f^i + \alpha_t^i + \lambda_t b(V_t) \partial_v f^i = 0. \tag{14}$$

Since all stochastic processes are of the form $f(t, S_t, V_t)$, we also obtain that λ has to be of the form $\lambda(t, S_t, V_t)$. Inserting the explicit form of α into (14) yields the assertion. □

Remark 3.7.

- The solution $f(t, S_t, V_t)$ of (10) is the price at time t for an european option with payoff Φ, given the current price S_t and current volatility V_t.
- The critical part in stochastic volatility models is the specification of the stochastic volatility process using the stochastic differential equation

  $$\mathrm{d}V_t = a(V_t)\,\mathrm{d}t + b(V_t)\,\mathrm{d}W_t.$$

  From stochastic calculus it follows that if a and b are Lipschitz continuous functions then there exists a unique solution V_t.
- The typical features of volatility are:

– **Mean-reversion**, i.e., that V_t tends to a long-term mean γ, more specifically

$$\operatorname{cov}(V_t - V_s, V_s) < 0$$

for $t > s$.

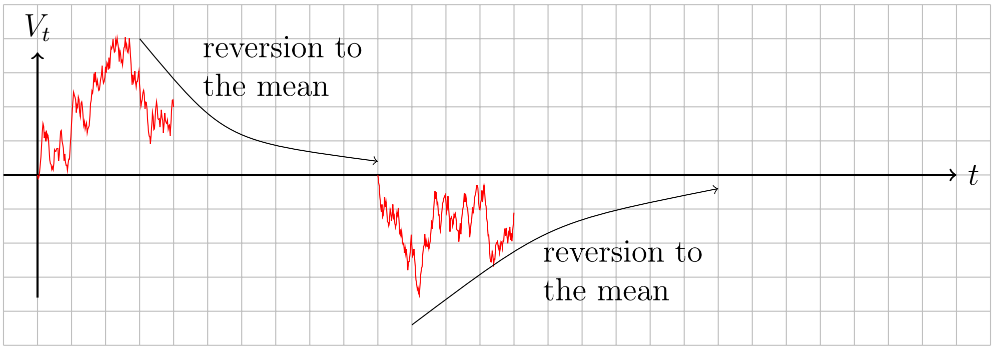


Figure 3: Mean reversion

– **Stationarity**, i.e., a limit in distribution

$$V_\infty := \lim_{t\to 0} V_t$$

exists. Therefore the (geometric) Brownian motion is not a good choice.

3.1 Examples of Stochastic Volatility Models

(A) The **Stein-Stein model** is a stochastic volatility model such that

$$\begin{aligned} a(V_t) &= -\kappa(V_t - \gamma), \\ b(V_t) &= \eta, \\ m(V_t) &= |V_t| \end{aligned}$$

and $\lambda(V_t) = c$. Hence

$$\mathrm{d}V_t = -\kappa(V_t - \gamma)\,\mathrm{d}t + \eta\,\mathrm{d}W_t, \tag{OU}$$

which is called the **Ornstein-Uhlenbeck process** with parameters

- κ, the speed of mean reversion,
- γ, the long-term volatility,
- η, the volatility of the volatility.

(B) The **Hull-White model** is a stochastic volatility model such that

$$\begin{aligned} a(V_t) &= a \cdot V_t, \\ b(V_t) &= b \cdot V_t, \\ m(V_t) &= V_t \end{aligned}$$

and $\lambda(V_t) = c$. Hence

$$\mathrm{d}V_t = V_t(a\,\mathrm{d}t + b\,\mathrm{d}W_t),$$

which is the geometric Brownian motion. The Hull-White-model has neither mean-reversion, nor a stationary distribution, but V_t is log-normal distributed which usually yields a good fit to the statistically observed distribution of V_t.

(C) The **Heston model** is a stochastic volatility model such that

$$\begin{aligned} a(V_t) &= -\kappa(V_t - \gamma), \\ b(V_t) &= \eta\sqrt{V_t}, \\ m(V_t) &= \sqrt{V_t} \end{aligned}$$

and $\lambda(V_t) = c\sqrt{V_t}$. Hence

$$\mathrm{d}V_t = -\kappa(V_t - \gamma)\,\mathrm{d}t + \eta\sqrt{V_t}\,\mathrm{d}W_t,$$

which is the **Cox-Ingersoll-Ross (CIR) process** with parameters as in (A).

The properties are:

- There is mean-reversion and a stationary distribution.
- The process V_t stays positive.
- The pricing PDE can be solved by Fourier transform.

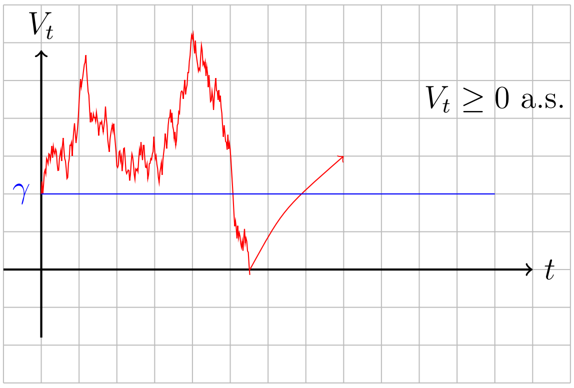


Figure 4: CIR process

(D) The **stochastic alpha-beta-rho (SABR) model** is a financial market model such that

$$\begin{aligned} a(V_t) &= 0, \\ b(V_t) &= \alpha \cdot V_t, \\ m(S_t, V_t) &= S_t^{\beta-1} \cdot V_t. \end{aligned}$$

Strictly speaking it is not in the scope of our definition, because m depends on both S and V. We have

$$\mathrm{d}S_t = S_t\left(\mu(t,\omega)\,\mathrm{d}t + V_t S_t^{\beta-1}\,\mathrm{d}B_t\right),$$

i.e., it is similar to the CEV model and

$$\mathrm{d}V_t = \alpha V_t\,\mathrm{d}W_t,$$

which is the geometric Brownian motion without drift. In the SABR-model there is neither mean reversion, nor a stationary distribution. On the other hand V_t is lognormal distributed, the model extends the CEV-model and it has a closed-form approximation of the implied volatility.

3.1.1 The Stein-Stein Model

First of all we focus on the Ornstein-Uhlenbeck process

$$\mathrm{d}V_t = -\kappa(V_t - \gamma)\,\mathrm{d}t + \eta\,\mathrm{d}W_t. \tag{OU}$$

The goals are

- finding an explicit solution of (OU), the distribution of V_t and its limit as $t \to \infty$,
- answering the question wether there is mean reversion, i.e., wether $\mathrm{cov}(V_t - V_s, V_s) < 0$ or not.

As Ansatz set

$$Z_t := e^{\kappa t}(V_t - \gamma).$$

Hence $Z_0 = V_0 - \gamma$. From the product rule and by (OU) it follows that

$$\begin{aligned}\mathrm{d}Z_t =& \kappa e^{\kappa t}(V_t - \gamma)\,\mathrm{d}t + e^{\kappa t}\,\mathrm{d}V_t\\ =& e^{\kappa t}\big(\kappa(V_t - \gamma)\,\mathrm{d}t + \mathrm{d}V_t\big)\\ =& \eta e^{\kappa t}\,\mathrm{d}W_t.\end{aligned}$$

Thus it holds that

$$Z_t = Z_0 + \eta \int_0^t e^{\kappa s}\,\mathrm{d}W_s$$

and that

$$V_t = \gamma + e^{-\kappa t} Z_t = e^{-\kappa t} V_0 + (1 - e^{-\kappa t})\gamma + \underbrace{\eta e^{-\kappa t} \int_0^t e^{\kappa s}\,\mathrm{d}W_s}_{\mathbb{E}[\cdot]=0}.$$

We can now deduce mean and variance of V_t

$$\begin{aligned}\mathbb{E}[V_t] &= e^{-\kappa t} V_0 + (1 - e^{-\kappa t})\gamma \overset{t\to\infty}{\longrightarrow} \gamma,\\ \mathrm{var}(V_t) &= \mathbb{E}\big[(V_t - \mathbb{E}[V_t])^2\big]\\ &= \eta^2 e^{-2\kappa t}\mathbb{E}\left[\left(\int_0^t e^{\kappa s}\,\mathrm{d}W_s\right)^2\right]\\ &= \eta^2 e^{-2\kappa t}\int_0^t e^{2\kappa s}\,\mathrm{d}s\\ &= \eta^2 e^{-2\kappa t}\left(\frac{e^{2\kappa t} - 1}{2\kappa}\right)\\ &= \frac{\eta^2}{2\kappa}\left(1 - e^{-2\kappa t}\right).\end{aligned}$$

Lemma 3.8. *Let $f \in C[0,T]$ be deterministic. Then the Itô-integral*

$$X_t = \int_0^t f(s)\,\mathrm{d}W_s$$

is a ***Gaussian process*** *with* $\mathbb{E}[X_t] = 0$ *and* $\operatorname{cov}(X_t, X_s) = \int_0^{t\wedge s} f(u)^2\,\mathrm{d}u$.

Proof. See [1, Chapter 4.2]. □

Using the previous Lemma it follows that

- the Ornstein-Uhlenbeck process is a Gaussian process,
- $V_t \sim \mathcal{N}\left(e^{-\kappa t}V_0 + (1-e^{-\kappa t})\gamma, \frac{\eta^2}{2\kappa}(1-e^{-2\kappa t})\right)$,
- there is convergence in distribution, i.e.,

$$V_\infty := \text{d-}\lim_{t\to\infty} V_t \sim N\left(\gamma, \frac{\eta^2}{2\kappa}\right).$$

Focussing on the covariance for $s < t$ it holds that

$$\begin{aligned}
\operatorname{cov}(V_t, V_s) &= \mathbb{E}\big[(V_t - \mathbb{E}[V_t])(V_s - \mathbb{E}[V_s])\big] \\
&= \eta^2 e^{-\kappa(t+s)} + \mathbb{E}\left[\int_0^t e^{\kappa u}\,\mathrm{d}W_u \cdot \int_0^s e^{\kappa u}\,\mathrm{d}W_u\right] \\
&= \eta^2 e^{-\kappa(t+s)} \int_0^s e^{2\kappa u}\,\mathrm{d}u \\
&= \eta^2 e^{-\kappa(t+s)} \frac{e^{2\kappa s}-1}{2\kappa} \\
&= \frac{\eta^2}{\kappa} e^{-\kappa t}\sinh(\kappa s).
\end{aligned}$$

Thus **there is** mean-reversion since

$$\operatorname{cov}(V_t - V_s, V_s) = \operatorname{cov}(V_t, V_s) - \mathbb{E}[V_s^2] = \ldots = \frac{\eta^2}{2\kappa}\underbrace{(e^{2\kappa s}-1)}_{>0}\underbrace{(e^{-\kappa(t-s)}-1)}_{<0} < 0.$$

3.2 The Heston Model and the Cox-Ingersoll-Ross Process

Recall the CIR process

$$\mathrm{d}V_t = \underbrace{-\kappa(V_t - \gamma)}_{a(V_t)}\,\mathrm{d}t + \underbrace{\eta\sqrt{V_t}}_{b(V_t)}\,\mathrm{d}W_t, \qquad \text{(CIR)}$$

where $V_0 = v > 0$ and with parameters

- $\kappa > 0$, the rate of mean reversion,
- $\gamma \geq 0$, the long-term mean,
- $\eta > 0$, the volatility of volatility.

3.2.1 Properties of the Cox-Ingersoll-Ross Process

Although $b(v)$ is not Lipschitz at $v = 0$ the stochastic differential equation (CIR) has a unique solution $(V_t)_{t\geq 0}$ and $V_t \geq 0$ almost surely for $t \geq 0$.

Definition 3.9. Let $\tau_0 := \{t > 0 : \ V_t = 0\}$, i.e., the **first hitting time of zero**.

Lemma 3.10 (Feller)**.** *It holds that $\tau_0 = \infty$ almost surely if and only if Feller's condition*

$$2\kappa\gamma \geq \eta^2$$

is satisfied. Hence if $2\kappa\gamma < \eta^2$ then $\tau_0 < \infty$ almost surely.

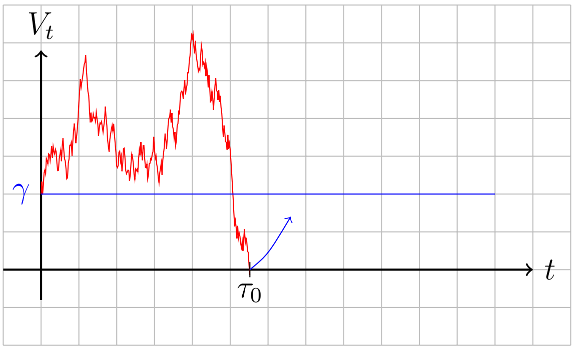


Figure 5: Illustration of τ_0

No proof.

The following Theorem now specifies the law of V_t using the characteristic function $f(u) = \mathbb{E}[e^{iuV_t}]$ for $u \in \mathbb{R}$.

Theorem 3.11. *The characteristic function of the Cox-Ingersoll-Ross process $(V_t)_{t\geq 0}$ is given by*

$$\mathbb{E}[e^{iuV_t}] = \exp\big(\underbrace{p_t(u) + vq_t(u)}_{\text{"affine in } v\text{"}}\big)$$

where

$$q_t(u) = \frac{iue^{-\kappa t}}{1 - \frac{\eta^2}{2}uc_t}, \qquad p_t(u) = -\frac{2\gamma u}{\eta^2}\log\left(1 - \frac{\eta^2}{2}uc_t\right) \tag{15}$$

and $c_t = \frac{1}{\kappa}(1 - e^{-\kappa t})$.

Corollary 3.12 (Properties of the CIR process)**.** *Let V_t be the solution of (CIR). Then the following properties hold.*

a) The fraction $\frac{4V_t}{\eta^2 c_t}$ has a non-central χ^2-distribution[9] *with $\frac{2\gamma\kappa}{\eta^2}$ degrees of freedom and non-centrality $\frac{4e^{-\kappa t}}{\eta^2 c_t}v$.*

b) V_t is stationary and

$$V_t \xrightarrow{d} V_\infty$$

as $t \to \infty$, where

$$\mathbb{E}[e^{iuV_\infty}] = \left(1 - \frac{\eta^2}{2\kappa}u\right)^{-\frac{2\gamma\kappa}{\eta^2}}.$$

Thus V_∞ is Gamma distributed with shape parameter $\frac{2\gamma\kappa}{\eta^2}$ and rate parameter $\frac{2\kappa}{\eta^2}$.

[9] Let X_i be independent with $X_i \sim \mathcal{N}(\mu_i, 1)$. Then $\sum_{i=1}^n X_i^2$ is non-central χ^2-distributed, with n degrees of freedom and non-centrality $\lambda = \sum_{i=1}^n \mu_i^2$.

Proof. Fix $T > 0$ and consider

$$M_t = \mathbb{E}\left[e^{iuV_T}\middle|\mathcal{F}_t\right]$$

for $t \le T$. M_t then is a **conditional characteristic function**. Note that

- M_t is a martingale on $t \in [0, T]$,
- the terminal value satisfies $M_T = e^{iuV_T}$,
- $M_0 = \mathbb{E}[e^{iuV_T}]$ is the characteristic function of V_T.

Now we use the **exponential-affine Ansatz**, i.e., setting

$$M_t = \exp\left(p_{T-t}(u) + V_t \cdot q_{T-t}(u)\right)$$

where $p_t(u),\ q_t(u)$ are deterministic C^2-functions that need to be determined. Note that

$$M_t = f(t, V_t)$$

where

$$f(t,x) = \exp\big(p_{T-t}(u) + x \cdot q_{T-t}(u)\big).$$

The partial derivatives of f are

$$\begin{aligned} \partial_t f(t,x) &= \big(-\dot{p}_{T-t}(u) - x \cdot \dot{q}_{T-t}(u)\big) \cdot f(t,x), \\ \partial_x f(t,x) &= q_{T-t}(u) \cdot f(t,x), \\ \partial_{xx} f(t,x) &= q_{T-t}^2(u) \cdot f(t,x). \end{aligned}$$

Applying Itô's formula to $M_t = f(t, V_t)$ and inserting the definition of V_t yields

$$\begin{aligned} \mathrm{d}M_t =& \big(-\dot{p}_{T-t}(u) - V_t\dot{q}_{T-t}(u)\big)M_t\,\mathrm{d}t + q_{T-t}(u)M_t\,\mathrm{d}V_t + \frac{1}{2}q_{T-t}^2(u)M_t\,\underbrace{\mathrm{d}[V,V]_t}_{=\eta^2 V_t\,\mathrm{d}t} \\ =& M_t\left(-\dot{p}_{T-t}(u) - V_t\dot{q}_{T-t}(u) - \kappa(V_t - \gamma)q_{T-t}(u) + \frac{\eta^2}{2}V_t q_{T-t}^2(u)\right)\mathrm{d}t \\ &+ M_t q_{T-t}(u)\eta\sqrt{V_t}\,\mathrm{d}W_t \end{aligned}$$

where the $\mathrm{d}t$-term must be 0, since M_t is a martingale. Seperating the coefficients of 1 and V_t and setting $\tau = T - t$ we obtain the following system of ordinary differential equations.

$$\begin{aligned} &\text{Coefficients of } 1: && \dot{p}_\tau(u) = \kappa\gamma q_\tau(u), && p_0(u) = 0, \\ &\text{Coefficients of } V_t: && \dot{q}_\tau(u) = -\kappa q_\tau(u) + \frac{\eta^2}{2}q_\tau(u)^2, && q_0(u) = iu. \end{aligned}$$

Here the initial conditions are derived from

$$e^{iuV_T} = M_t = \exp\big(\underbrace{p_0(u)}_{=0} + V_T\,\underbrace{q_0(u)}_{=iu}\big).$$

The proof is concluded noting that the equations in (15) are the solutions of the upper system of ODEs.[10] □

Essentially, processes for which the exponential-affine Ansatz works are called affine processes. To be precise:

Definition 3.13. A $\mathbb{R}^d$-valued stochastic process X_t is an **affine process** if there exist C^1-functions $p_t(u)$, $q_t(u)$ such that

$$\mathbb{E}\left[e^{iu^\top X_T}\middle|\mathcal{F}_t\right] = \exp\left(p_{T-t}(u) + q_{T-t}(u)^\top X_t\right)$$

for all $0 \le t \le T,\ u \in \mathbb{R}^d$.

Examples of affine processes are

- the CIR process, the OU process,
- in the Heston model $(\log S_t, V_t)_{t\ge 0}$ is an affine process,
- Lévy processes,
- Lévy-driven OU processes.

Exercise. Show that the Ornstein-Uhlenbeck process is an affine process. Use this to rederive the results on its law.

3.2.2 Pricing in the Heston Model

Recall the stochastic differential equations in the Heston model

$$\begin{aligned} \mathrm{d}S_t^0 &= rS_t^0\,\mathrm{d}t, \\ \mathrm{d}S_t &= S_t\sqrt{V_t}\,\mathrm{d}B_t, \\ \mathrm{d}V_t &= -\kappa(V_t-\gamma)\,\mathrm{d}t + \eta\sqrt{V_t}\,\mathrm{d}W_t \end{aligned}$$

with $\mathrm{d}[B,W]_t = \varrho\,\mathrm{d}t$ and $\varrho \in [-1,1]$. With $r = 0$ we obtain by Theorem 3.5 the pricing PDE in the Heston model, i.e., let

$$\pi_t = f(t, S_t, V_t)$$

be the fair price of an European option with payoff Φ. Then f satisfies

$$\begin{aligned} 0 =&\partial_t f - \kappa(v-\gamma)\partial_v f + \frac{1}{2}vs^2\partial_{ss} f + \varrho\eta v s\partial_{vs} f + \frac{\eta^2}{2}\partial_{vv} f \\ \Phi(s) =&f(T,s,v). \end{aligned} \tag{HPDE}$$

[10] Obviously it holds that

$$p_\tau(u) = \kappa\gamma\int_0^\tau q_s(u)\,\mathrm{d}s.$$

The ODE for q is a Riccati equation.

It will be convenient to transform to $x = \log s$, i.e., setting

$$h(t, x, v) = f(t, e^x, v).$$

Theorem 3.14 (Fourier pricing)**.** *Consider the Fourier transform of h and Φ in the Heston model, i.e.,*

$$\begin{aligned}\widehat{h}(t, u, v) &= \int_{-\infty}^{\infty} e^{iux} h(t, x, v)\,\mathrm{d}x = \int_{-\infty}^{\infty} e^{iux} f(t, e^x, v)\,\mathrm{d}x,\\ \widehat{\Phi}(u) &= \int_{-\infty}^{\infty} e^{iux} \Phi(e^x)\,\mathrm{d}x.\end{aligned}$$

Then it holds that

$$\widehat{h}(t, u, v) = \exp\big(p_{T-t}(u) + v q_{T-t}(u)\big)\widehat{\Phi}(u),$$

where p, q solve the following ordinary differential equations

$$\begin{cases}\dot{p}_\tau(u) = \kappa\gamma \cdot q_\tau(u), & p_0(u) = 0,\\ \dot{q}_\tau(u) = (iu\varrho\eta - \kappa) q_\tau(u) + \dfrac{\eta^2}{2} q_\tau^2(u) - \dfrac{u^2}{2}, & q_0(u) = 0.\end{cases}$$

Remark 3.15. Note that

$$p_\tau(u) = \kappa\gamma \int_0^\tau q_s(u)\,\mathrm{d}s$$

and that the equation for q is a Riccati ODE.

Proof. We will do the proof in four steps:

Step I: *Rewrite (HPDE) in terms of $x = \log s$*
Consider the partial derivatives of h, i.e.,

$$\begin{aligned}\partial_x h &= e^x \partial_s f = s\partial_s f,\\ \partial_{xx} h &= e^{2x} \partial_s f = s^2 \partial_{ss} f.\end{aligned}$$

Thus (HPDE) becomes

$$\begin{cases}\partial_t h - \kappa(\gamma - v)\partial_v h + \dfrac{1}{2} v \partial_{xx} h + \varrho\eta v \partial_{vx} h + \dfrac{\eta^2}{2} v \partial_{vv} h = 0,\\ h(T, x, v) = \Phi(e^x).\end{cases} \tag{16}$$

Note that the coefficients are affine in v.

Step II: *Apply the Fourier transform*
Consider the Fourier transform of h

$$\widehat{h}(t, u, v) = \mathcal{F}[h](u) = \int_{-\infty}^{\infty} e^{iux} h(t, x, v)\,\mathrm{d}x$$

and calculate the Fourier transforms of h's partial derivatives, i.e.,

$$
\begin{aligned}
\mathcal{F}[\partial_x h](u) &= \int_{-\infty}^{\infty} e^{iux}\partial_x h(t,x,v)\,\mathrm{d}x = iu\int_{-\infty}^{\infty} e^{iux} h(t,x,v)\\
&= iu\mathcal{F}[h](u),\\
\mathcal{F}[\partial_{xx} h](u) &= -u^2\mathcal{F}[h](u),\\
\mathcal{F}[\partial_t h](u) &= \int_{-\infty}^{\infty} e^{iux}\partial_t h(t,x,v)\,\mathrm{d}x\\
&= \partial_t\mathcal{F}[h](u),\\
\mathcal{F}[\partial_v h](u) &= \partial_v\mathcal{F}[h](u),\\
\mathcal{F}[\partial_{vv} h](u) &= \partial_{vv}\mathcal{F}[h](u),\\
\mathcal{F}[\partial_{vx} h](u) &= iu\partial_v\mathcal{F}[h](u).
\end{aligned}
$$

Step III: *Apply the exponential-affine Ansatz*

Set

$$\mathcal{F}[h](u) = \widehat{h}(t,u,v) = \exp\big(p_{T-t}(u) + vq_{T-t}(u)\big)\widehat{\Phi}(u)$$

with p,q to be determined and terminal condition

$$\widehat{h}(T,u,v) = \widehat{\Phi}(u).$$

Thus $p_0(u) = q_0(u) = 0$, which is the initial condition for the Riccati ODE. For the partial derivatives we obtain

$$
\begin{aligned}
\mathcal{F}[\partial_t h](u) &= -\big(\dot{p}_{T-t}(u) + v\dot{q}_{T-t}(u)\big)\cdot\mathcal{F}[h](u),\\
\mathcal{F}[\partial_v h](u) &= q_{T-t}(u)\cdot\mathcal{F}[h](u),\\
\mathcal{F}[\partial_{vx} h](u) &= iuq_{T-t}(u)\cdot\mathcal{F}[h](u),\\
\mathcal{F}[\partial_{vv} h](u) &= q_{T-t}^2(u)\cdot\mathcal{F}[h](u).
\end{aligned}
$$

Step IV: *Apply the Fourier transform to* (16)

Setting $\tau = T - t$ yields

$$0 = \mathcal{F}[(16)](u) = \mathcal{F}[h](u)\cdot\alpha_\tau(v),$$

where

$$\alpha_\tau(v) := \left(-\dot{p}_\tau(u) - v\dot{q}_\tau(u) - \kappa(v-\gamma)q_\tau(u) - \frac{u^2}{2}v + \varrho\eta viuq_\tau(u) + \frac{\eta^2}{2}q_\tau^2(u)\right).$$

Note that α_t is an affine function of v. Since $\mathcal{F}[h](u) \neq 0$ for all $(t,u,v) \in [0,T]\times\mathbb{R}\times\mathbb{R}_{\geq 0}$ we conclude that $\alpha_\tau(v)$ has to be equal to 0 for all $(t,u,v) \in [0,T]\times\mathbb{R}\times\mathbb{R}_{\geq 0}$. Again, by collecting the coefficients we obtain the following equations.

$$
\begin{aligned}
&\text{Coefficients of } 1: && \dot{p}_\tau(u) = \kappa\gamma q_\tau(u), && p_0(u) = 0,\\
&\text{Coefficients of } v: && \dot{q}_\tau(u) = (\varrho\eta iu - \kappa)q_\tau(u) + \frac{\eta^2}{2}q_\tau^2(u) - \frac{u^2}{2}, && q_0(u) = 0.
\end{aligned}
$$

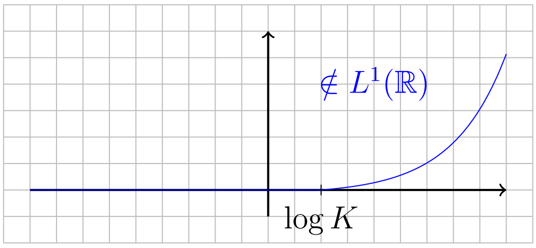
□

Remark 3.16. If we try to price put and call options by using the inverse Fourier transform

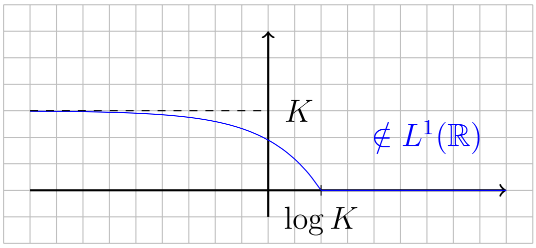
$$h(t,x,v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iux}\widehat{h}(t,u,v)\,\mathrm{d}u$$

we have the problem, that h must be in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ which is not true for the payoffs. Indeed, note that

for a call: $\Phi(x) = (e^x - K)_+$



for a put: $\Phi(x) = (K - e^x)_+$



We solve this by 'dampening' with an exponential function e^{Rx} and get

call:	$\widetilde{\Phi}(x) = e^{-Rx}(e^x - K)_+$	with $R > 1$,
put:	$\widetilde{\Phi}(x) = e^{-Rx}(K - e^x)_+$	with $R < 0$.

Considering the Fourier transform, we obtain

$$\begin{aligned}\mathcal{F}[e^{-Rx}g](u) &= \int_{-\infty}^{\infty} e^{-Rx+iux}g(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} e^{i(u+iR)x}g(x)\,\mathrm{d}x \\ &= \widehat{g}(u+iR)\end{aligned}$$

under suitable integrability conditions. Hence

Exponential dampening in original domain $\iff$ Imaginary shift in Fourier domain

Calculating the Fourier transform of the exponentially dampened put we obtain

$$\begin{aligned}\mathcal{F}\big[e^{-Rx}(K-e^x)_+\big](u) &= \int_{-\infty}^{\infty} e^{(iu-R)x}(K-e^x)_+ \,\mathrm{d}x \\ &= \int_{-\infty}^{\log K} \big(Ke^{(iu-R)x} - e^{(iu-R+1)x} \,\mathrm{d}x \\ &= \kappa \frac{e^{(iu-R)x}}{(iu-R)}\bigg|_{-\infty}^{\log K} - \frac{e^{(iu-R+1)x}}{(iu-R+1)}\bigg|_{-\infty}^{\log K} \\ &= K^{iu-R+1}\left[\frac{1}{(iu-R)} - \frac{1}{(iu-R+1)}\right] \\ &= K^{iu-R+1}\frac{1}{(iu-R)(iu-R+1)}.\end{aligned}$$

We summarize the above calculations in the following corollary:

Corollary 3.17. *The price of a put option with maturity T and strike K in the Heston model is given by the inverse Fourier integral*

$$\Pi_t^{put} = f(t, S_t, V_t) = \frac{\left(\frac{S_t}{K}\right)^R}{2\pi} \cdot K \cdot \int_{-\infty}^{\infty} e^{-iu\log\left(\frac{S_t}{K}\right)} \cdot \frac{\exp\big(p_{T-t}(u+iR) + V_t q_{T-t}(u+iR)\big)}{(iu-R)(iu-R+1)} \,\mathrm{d}u, \tag{17}$$

$R < 0$. For the call the same formular holds with $R > 1$.

Remark 3.18. (17) can be written as a line integral in the complex plane, i.e., for

$$F(\xi) = \exp(p_{T-t}(i\xi)) + V_t q_{T-t}(i\xi))$$

we obtain

$$\frac{1}{2\pi}\int_{-i\infty+R}^{i\infty+R} e^{\xi\log\left(\frac{K}{S_t}\right)} \cdot \frac{F(\xi)}{\xi(\xi+1)} \,\mathrm{d}\xi$$

with $\xi = iu - R$. There are two poles at $\xi = 0$ and $\xi = -1$. In complex analysis it is well-known that the value of the integral is independent under changes of R as long as no pole is crossed.

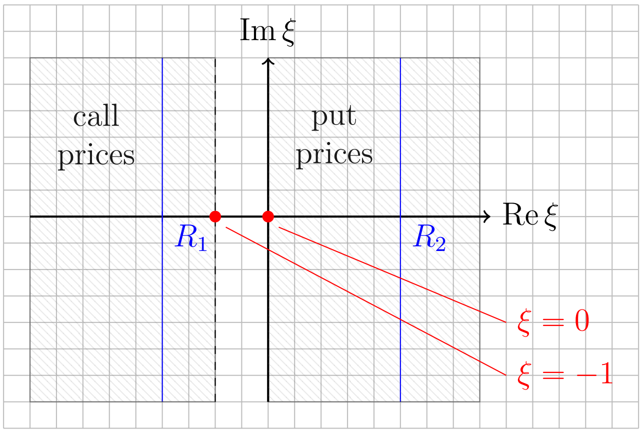


Figure 6: Call and put prices in the Heston model

4 Risk-Neutral Pricing and Arbitrage

In discrete time the first fundamental theorem of asset pricing states that there is no arbitrage if and only if there exists an equivalent martingale measure $\mathbb{Q}$. (See [2, Theorem 2.1]) Furthermore the risk-neutral pricing formula

$$\Pi_t := S_t^0 \cdot \mathbb{E}^{\mathbb{Q}} \left[\frac{C}{S_T^0} \middle| \mathcal{F}_t \right]$$

determines the fair price of a claim C at time t. In order to derive the analogues in continuous time we repeat some basic terminology from probability theory.

4.1 Change of Measure

Definition 4.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathbb{Q}$ be another probability measure on $(\Omega, \mathcal{F})$. We say that

- $\mathbb{Q}$ is **absolutely continuous** with respect to $\mathbb{P}$ (notation: $\mathbb{Q} \ll \mathbb{P}$) if

$$\mathbb{P}(A) = 0 \implies \mathbb{Q}(A) = 0 \qquad \forall\ A \in \mathcal{F}.$$

- $\mathbb{Q}$ is **equivalent** to $\mathbb{P}$ (notation: $\mathbb{Q} \sim \mathbb{P}$) if

$$\mathbb{P}(A) = 0 \iff \mathbb{Q}(A) = 0 \qquad \forall\ A \in \mathcal{F}.$$

- $\mathbb{Q}$ and $\mathbb{P}$ are **mutually singular** (notation: $\mathbb{Q} \perp \mathbb{P}$) if

$$\exists\ A \in \mathcal{F} \text{ such that } \mathbb{Q}(A) = 1 \text{ and } \mathbb{P}(A) = 0.$$

Remark 4.2. There are some simple consequences of the upper definitions, namely

- $\mathbb{Q} \sim \mathbb{P}$ if and only if $(\mathbb{Q} \ll \mathbb{P}) \wedge (\mathbb{P} \ll \mathbb{Q})$,
- $\mathbb{Q} \sim \mathbb{P}$ if and only if
$$\mathbb{P}(A) = 1 \;\Leftrightarrow\; \mathbb{Q}(A) = 1 \quad \forall\, A \in \mathcal{F}.$$

Theorem 4.3 (Radon-Nikodym)**.** *The measure $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ if and only if there exists a random variable $X \in \mathcal{L}(\mathrm{d}\mathbb{P})$ such that $X \geq 0$ $\mathbb{P}$-almost surely, $\mathbb{E}^{\mathbb{P}}[X] = 1$, and*
$$\int_A X \,\mathrm{d}\mathbb{P} = \int_A \mathrm{d}\mathbb{Q} \qquad \forall\, A \in \mathcal{F}. \tag{18}$$
Furthermore $\mathbb{Q} \sim \mathbb{P}$ if and only if $X > 0$ $\mathbb{P}$-almost surely.

Remark 4.4.

- The equation (18) is equivalent to
$$\mathbb{E}^{\mathbb{P}}[X \mathbb{1}_A] = \mathbb{Q}(A).$$
- X is called Radon-Nikodym derivative or Radon-Nikodym density of $\mathbb{Q}$ with respect to $\mathbb{P}$. The notation is
$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} := X.$$
- If $\mathbb{Q} \sim \mathbb{P}$, then it holds that
$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\right)^{-1}.$$
- If $\mathbb{R} \ll \mathbb{Q} \ll \mathbb{P}$, then it holds that
$$\frac{\mathrm{d}\mathbb{R}}{\mathrm{d}\mathbb{P}} = \frac{\mathrm{d}\mathbb{R}}{\mathrm{d}\mathbb{Q}} \cdot \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}.$$
- On a discrete probability space $\Omega = (\omega_1, ..., \omega_N)$ we have
$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}(\omega_i) = \frac{\mathbb{Q}(\omega_i)}{\mathbb{P}(\omega_i)} \quad \text{whenever } \mathbb{P}(\omega_i) > 0.$$
- If $\mathbb{Q}$ and $\mathbb{P}$ are probability measures on $\mathbb{R}$ with densities g and f, then it holds that
$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}(x) = \frac{g(x)}{f(x)} \qquad \forall\, x \in \mathbb{R}.$$

Example 4.5. (a) Let $\mathbb{P} \sim N(0, \sigma^2)$ and $\mathbb{Q} \sim N(-\mu, \sigma^2)$. Then it holds that
$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x+\mu)^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}} = \exp\left(\frac{x^2}{2\sigma^2} - \frac{(x+\mu)^2}{2\sigma^2}\right) = \exp\left(-\frac{x\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right).$$

(b) Consider the multivariate case, i.e., $\mathbb{P}, \mathbb{Q}$ independent with $\mathbb{P} \sim N(0, \operatorname{diag}(\sigma_1, ..., \sigma_d))$ and $\mathbb{Q} \sim N(\mu, \operatorname{diag}(\sigma_1, ..., \sigma_d))$. It holds that

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}(x) = \exp\left(-\sum_{i=1}^{d} \frac{x_i \mu_i}{\sigma_i^2} - \frac{1}{2}\sum_{i=1}^{d} \frac{\mu_i^2}{\sigma_i^2}\right).$$

Corollary 4.6. *Let $Y \in \mathcal{L}(\mathrm{d}\mathbb{Q})$, then it holds that*

$$\mathbb{E}^{\mathbb{Q}}[Y] = \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \cdot Y\right].$$

Proof. The result follows by approximating Y by sums of indicator functions. □

Lemma 4.7 (Change of measure for conditional expectations)**.** *Let $\mathbb{Q} \sim \mathbb{P}$, $Y \in \mathcal{L}(\mathrm{d}\mathbb{Q})$ and $\mathcal{G} \subseteq \mathcal{F}$. Then*

$$\mathbb{E}^{\mathbb{Q}}[Y|\mathcal{G}] = \frac{\mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} Y \middle| \mathcal{G}\right]}{\mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \middle| \mathcal{G}\right]}.$$

This formula is called the ***abstract Bayes formula****.*

Proof. We show that

$$\mathbb{E}^{\mathbb{Q}}[Y|\mathcal{G}] \cdot \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \middle| \mathcal{G}\right] = \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} Y \middle| \mathcal{G}\right].$$

Let $A \in \mathcal{G}$ arbitrary. The tower law and Theorem 4.3 yield

$$\int_A \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} Y \middle| \mathcal{G}\right] \mathrm{d}\mathbb{P} = \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} Y \middle| \mathcal{G}\right] \cdot \mathbb{1}_A\right] = \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} Y \cdot \mathbb{1}_A\right] = \mathbb{E}^{\mathbb{Q}}[Y \mathbb{1}_A].$$

On the other hand it holds that

$$\begin{aligned}\int_A \mathbb{E}^{\mathbb{Q}}[Y|\mathcal{G}] \cdot \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \middle| \mathcal{G}\right] \mathrm{d}\mathbb{P} &= \mathbb{E}^{\mathbb{P}}\left[\mathbb{1}_A \cdot \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \cdot \mathbb{E}^{\mathbb{Q}}[Y|\mathcal{G}] \middle| \mathcal{G}\right]\right] \\ &= \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}} \mathbb{1}_A \cdot \mathbb{E}^{\mathbb{Q}}[Y|\mathcal{G}]\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_A \mathbb{E}^{\mathbb{Q}}[Y|\mathcal{G}]\right] = \mathbb{E}^{\mathbb{Q}}[Y \mathbb{1}_A].\end{aligned}$$

Thus

$$\int_A \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \middle| \mathcal{G}\right] \mathrm{d}\mathbb{P} = \int_A \mathbb{E}^{\mathbb{Q}}[Y|\mathcal{G}] \cdot \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \middle| \mathcal{G}\right] \mathrm{d}\mathbb{P}.$$

Since A was arbitrary the assertion is proven. □

4.1.1 Change of Measure on Filtered Probability Spaces

Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $(\mathcal{F}_t)_{t\geq 0} \subseteq \mathcal{F}$ and a probability measure $\mathbb{Q} \ll \mathbb{P}$. Applying the Radon-Nikodym theorem to $(\Omega, \mathcal{F}_t, \mathbb{P})$ we obtain that for all $t \geq 0$ a **Radon-Nikodym derivative** $X_t \in \mathcal{L}(\Omega, \mathcal{F}_t \mathbb{P})$ such that $X_t \geq 0$ $\mathbb{P}$-almost surely, $\mathbb{E}^{\mathbb{P}}[X_t] = 1$ and

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[\mathbb{1}_A X_t] \qquad \forall\, A \in \mathcal{F}_t.$$

We denote

$$X_t =: \left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t}.$$

Lemma 4.8. *Let X_t be as above. It holds that*

$$X_t = \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\middle|\mathcal{F}_t\right] \qquad \forall\, t \geq 0.$$

In particular, $X := (X_t)_{t\geq 0}$ is a positive $\mathbb{P}$-martingale such that $X_0 = \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right] = 1$.

Remark 4.9. The stochastic process X is called the **density process** of $\mathbb{Q}$ with respect to $\mathbb{P}$.

Proof. Using the Radon-Nikodym theorem we obtain

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}[X_t \cdot \mathbb{1}_A] &= \mathbb{Q}(A) && \forall\, A \in \mathcal{F}_t, \\ \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\mathbb{1}_A\right] &= \mathbb{Q}(A) && \forall\, A \in \mathcal{F}. \end{aligned}$$

The assertion follows immediately. □

Lemma 4.10. *Let X be a local $\mathbb{P}$-martingale on a filtered probability space. Let $\mathbb{Q} \sim \mathbb{P}$ with density process*

$$Z_t = \left.\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\right|_{\mathcal{F}_t}.$$

Then $XZ = (X_t Z_t)_{t\geq 0}$ is a local $\mathbb{Q}$-martingale.

Proof. Let (τ_n) be a localizing sequence for X. Recall the local martingale property

$$\mathbb{E}^{\mathbb{P}}[X_t^{\tau_n} | \mathcal{F}_s] = X_s^{\tau_n} \qquad \forall\, n \in \mathbb{N},\ s \leq t.$$

Applying the Radon-Nikodym theorem for conditional expectations yields

$$\frac{\mathbb{E}^{\mathbb{Q}}[Z_t^{\tau_n} X_t^{\tau_n} | \mathcal{F}_s]}{\underbrace{\mathbb{E}^{\mathbb{Q}}[Z_t^{\tau_n} | \mathcal{F}_s]}_{=Z_s^{\tau_n}}} = X_s^{\tau_n}.$$

Thus

$$\mathbb{E}^{\mathbb{Q}}[Z_t^{\tau_n} X_t^{\tau_n} | \mathcal{F}_s] = Z_s^{\tau_n} X_s^{\tau_n}$$

which again is the local martingale property. □

4.2 Girsanov's Theorem

Definition 4.11. Consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with a Brownian motion $(B_t)_{t\geq 0}$ where $(\mathcal{F}_t)_{t\geq 0}$ the Brownian standard Filtration (BSF). The Itô Process

$$\mathrm{d}X_t = \mu_t \,\mathrm{d}t + \mathrm{d}B_t, \qquad X_0 = 0$$

is called **Brownian motion** with drift.

Clearly if X is a Brownian motion with drift then it is not a Brownian motion, unless $\mu = 0$ $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-almost surely. Roughly said, the Theorem of Girsanov states that under mild conditions we can find a probability measure $\mathbb{Q} \sim \mathbb{P}$ such that X is a Brownian motion under $\mathbb{Q}$.

Theorem 4.12 (Girsanov's theorem). *Let X be a process of the form*

$$\mathrm{d}X_t = \mu_t \,\mathrm{d}t + \mathrm{d}B_t.$$

If

$$M_t = \exp\left(-\int_0^t \mu_s \,\mathrm{d}B_s - \frac{1}{2}\int_0^t \mu_s^2 \,\mathrm{d}s\right), \qquad t \in [0,T]$$

is a martingale under $\mathbb{P}$ then

$$\mathbb{Q}(A) := \mathbb{E}^{\mathbb{P}}\left[\mathbb{1}_A M_T\right] \qquad \forall\ A \in \mathcal{F}_T \tag{19}$$

defines a probability measure $\mathbb{Q} \sim \mathbb{P}$ on $(\Omega, \mathcal{F}_T)$ and X is a Brownian motion on $(\Omega, \mathcal{F}_T, \mathbb{Q})$.

Remark 4.13. There are two claims, namely

(1) $\mathbb{Q}$ defines an equivalent probability measure,

(2) X is a $\mathbb{Q}$-Brownian motion.

Note that M_T is the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ on $(\Omega, \mathcal{F}_T)$, i.e.,

$$M_T = \left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_T}.$$

Proof. The first claim is proven easily, considering that

- $\mathbb{Q}(\Omega) = \mathbb{E}^{\mathbb{P}}[M_T] = M_0 = 1$,
- $\mathbb{Q}(A) \geq 0 \qquad \forall A \in \mathcal{F}_T$.

Furthermore let $A \in \mathcal{F}_T$ be arbitrary and $(A_n)_n$ be a sequence such that $A_j \cap A_i = \emptyset$ for $i \neq j$ and $\bigcup_{n=0}^{\infty} = A$. By monotone convergence it holds that

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[\mathbb{1}_A \cdot M_T] = \mathbb{E}^{\mathbb{P}}\left[\sum_{n=0}^{\infty} \mathbb{1}_{A_n} \cdot M_T\right] = \sum_{n=0}^{\infty} \mathbb{E}^{\mathbb{P}}[\mathbb{1}_A \cdot M_T] = \sum_{n=0}^{\infty} \mathbb{Q}(A_n).$$

Thus $\mathbb{Q}$ is also σ-additive and therefore a probability measure. To proof the second claim we need to show that

a) X is continuous $\mathbb{Q}$-almost surely,

b) $(X_{t_2} - X_{t_1}), ..., (X_{t_n} - X_{t_n - 1})$ are independent on $(\Omega, \mathcal{F}, \mathbb{Q})$ and

c) $(X_t - X_s) \overset{\mathbb{Q}}{\sim} N(0, t-s)$.

Since $\mathbb{Q} \sim \mathbb{P}$ the claim a) follows immediately from the continuity of X with respect to $\mathbb{P}$. To show (b) and (c) we use characteristic functions. We do the proof in two steps.

Step I: Let f be a deterministic, measurable, bounded function $\mathbb{R} \to \mathbb{R}$. We try to show

$$\mathbb{E}^{\mathbb{Q}}\left[\exp\left(i\int_0^T f(s)\,\mathrm{d}X_s\right)\right] = \exp\left(-\frac{1}{2}\int_0^T f(s)^2\,\mathrm{d}s\right). \tag{20}$$

Considering the definitions of X_s and M_T we obtain

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[\exp\left(i\int_0^T f(s)\,\mathrm{d}X_s\right)\right] &= \mathbb{E}^{\mathbb{P}}\left[\exp\left(i\int_0^T f(s)\,\mathrm{d}X_s\right)M_T\right]\\ &= \mathbb{E}^{\mathbb{P}}\left[\exp\left(i\int_0^T f(s)\,\mathrm{d}X_s - \int_0^T \mu_s\,\mathrm{d}B_s - \frac{1}{2}\int_0^T \mu_s^2\,\mathrm{d}s\right)\right]\\ &= \exp\left(-\frac{1}{2}\int_0^T f(s)^2\,\mathrm{d}s\right)\cdot \mathbb{E}^{\mathbb{P}}\Big[\exp(L_T)\Big],\end{aligned}$$

where

$$L_T := \int_0^T (if(s) - \mu_s)\,\mathrm{d}B_s - \frac{1}{2}\int_0^T (if(s)-\mu_s)^2\,\mathrm{d}s.$$

Setting $N_T = \exp(L_T)$ follows that (20) holds if $\mathbb{E}^{\mathbb{P}}[N_t] = 1$. Applying Itô's formula to N_t yields

$$\begin{aligned}\mathrm{d}N_t &= N_t\,\mathrm{d}L_t + \frac{1}{2}N_t\,\mathrm{d}[L,L]_t\\ &= N_t\left((if(t)-\mu_t)\,\mathrm{d}B_t - \frac{1}{2}\cancel{(if(t)-\mu_t)^2\,\mathrm{d}t} + \frac{1}{2}\cancel{(if(t)-\mu(t))^2\,\mathrm{d}t}\right)\\ &= N_t\big(if(s)-\mu_s\big)\,\mathrm{d}B_s.\end{aligned}$$

Thus N_t is a local martingale. Furthermore it holds that

$$|N_t| \leq \exp\left(\frac{1}{2}\int_0^t f(s)^2\,\mathrm{d}s\right)\cdot M_T.$$

Therefore N_t is a Martingale. Let (τ_n) be a localizing sequence for N. Then the theorem of dominated convergence yields

$$\mathbb{E}^{\mathbb{P}}[N_T] = \mathbb{E}^{\mathbb{P}}\left[\lim_{n\to\infty} N_{T\wedge\tau_n}\right] = \lim_{n\to\infty}\mathbb{E}^{\mathbb{P}}[N_{T\wedge\tau_n}] = N_0 = 1.$$

and (20) follows.

Step II: Choose a step function

$$f(t) = \sum_{k=0}^{n-1} \vartheta_k \cdot \mathbb{1}_{(t_k,t_{k+1}]}(t), \qquad (\vartheta_k) \subseteq \mathbb{R}$$

and insert into (20). Then it holds that

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\left[\exp\left(i\sum_{k=0}^{n-1}\vartheta_k(X_{t_{k+1}} - X_{t_k})\right)\right] &= \exp\left(-\frac{1}{2}\sum_{k=0}^{n-1}\vartheta_k^2(t_{k+1}-t_k)\right) \\ &= \prod_{k=0}^{n-1}\exp\left(-\frac{1}{2}\vartheta_k^2(t_{k+1}-t_k)\right). \end{aligned} \tag{21}$$

Hence the left hand side of (21) is the characteristic function of $(X_{t_2} - X_{t_1}), ..., (X_{t_k} - X_{t_{k-1}})$ under $\mathbb{Q}$. The right hand side is the product of the characteristic functions of $\mathcal{N}(0, t_{k+1} - t_k)$, which implies c). Furthermore this factorization yields independence and thus b). □

Corollary 4.14 (Changing the drift)**.** *Let X be an Itô-process of the form*

$$\mathrm{d}X_t = \mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}B_t, \qquad X_0 = x$$

and let $\nu_t \in \mathcal{L}^2_{loc}$.[11] *Set $\vartheta_t := \frac{\mu_t - \nu_t}{\sigma_t}$. If*

$$M_t = \exp\left(-\int_0^t \vartheta_s \,\mathrm{d}B_s - \frac{1}{2}\int_0^t \vartheta_s^2 \,\mathrm{d}s\right)$$

is a $\mathbb{P}$-martingale, then

(a) $\mathbb{Q}(A) := \mathbb{E}^{\mathbb{P}}[\mathbb{1}_A \cdot M_T]$ defines a probability measure $\mathbb{Q} \sim \mathbb{P}$ on $(\Omega, \mathcal{F}_T)$,

(b) $B_t^{\mathbb{Q}} := B_t + \int_0^t \vartheta_s \,\mathrm{d}s$ is a Brownian motion on $(\Omega, \mathcal{F}_T, \mathbb{Q})$,

(c) X takes the form

$$\mathrm{d}X_t = \nu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}B_t^{\mathbb{Q}}, \qquad X_0 = x,$$

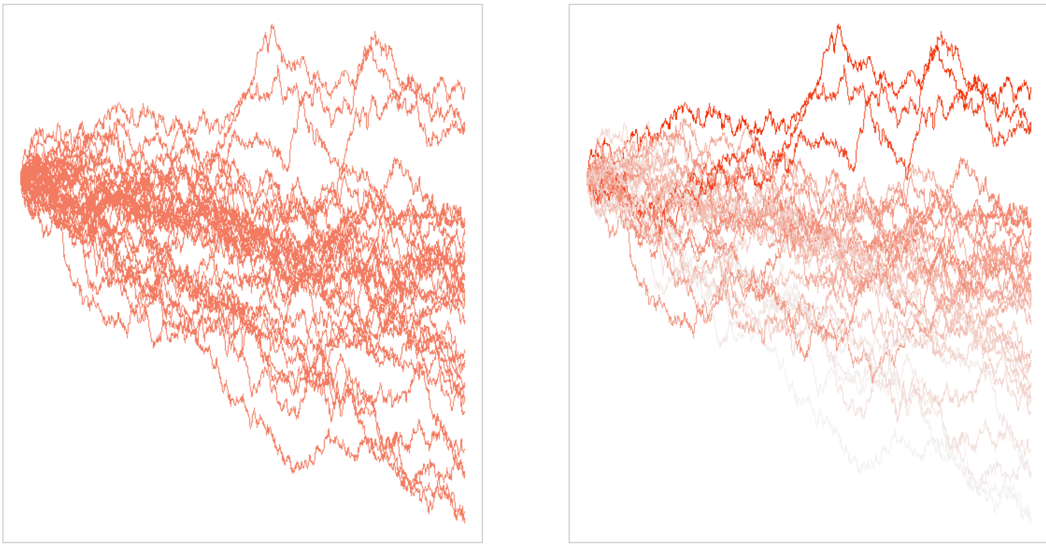
where ν_t is the new drift and $B_t^{\mathbb{Q}}$ a $\mathbb{Q}$ Brownian motion.

Proof. (a) and (b) follow immediately from Theorem 4.12. To proof (c) consider

$$\begin{aligned} \mathrm{d}X_t &= \mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}B_t \\ &= \mu_t \,\mathrm{d}t + \sigma_t \left(\mathrm{d}B_t^{\mathbb{Q}} - \vartheta_t \,\mathrm{d}t\right) \\ &= \cancel{\mu_t \,\mathrm{d}t} + \sigma_t \,\mathrm{d}B_t^{\mathbb{Q}} + (\nu_t - \cancel{\mu_t})\,\mathrm{d}t \\ &= \nu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}B_t^{\mathbb{Q}}. \end{aligned}$$

□

[11] $\mathcal{L}^2_{\text{loc}} = \left\{ f \text{ adapted: } \int_0^T f(s)^2 \,\mathrm{d}s < \infty \text{ almost surely} \right\}$



30 paths of a Brownian motion with negative drift — The same paths reweighted according to the Girsanov formula

Figure 7: Changing the drift according to Girsanov

Source: Martin Keller-Ressel - created with GNU R, from
`https://upload.wikimedia.org/wikipedia/commons/b/b3/Girsanov.png`

Novikov's condition now provides a useful tool to check wether M is a martingale.

Theorem 4.15 (Novikov's condition)**.** *Let $\mu \in \mathcal{L}^2_{loc}$ and set*

$$M_t = \exp\left(-\int_0^t \mu_s \, \mathrm{d}B_s - \frac{1}{2}\int_0^t \mu_s^2 \, \mathrm{d}s \right).$$

If

$$\mathbb{E}^{\mathbb{P}}\left[\exp\left(\frac{1}{2}\int_0^T \mu_s^2 \, \mathrm{d}s \right) \right] < \infty \tag{NOV}$$

then M is a martingale on $[0, T]$. (NOV) *is called **Novikov's condition**.*

Proof. See [1, Theorem 7.7]. □

4.3 Martingale Representation

Theorem 4.16 (Martingale representation on $\mathcal{L}^2$)**.** *Let X be a martingale with respect to the Brownian standard filtration generated by a one-dimensional Brownian motion $(B_t)_{t\geq 0}$ and assume that $\mathbb{E}[X_T^2] < \infty$. Then exists $\varphi \in \mathcal{H}^2[0, T]$ such that*

$$X_t = X_0 + \int_0^t \varphi_s \, \mathrm{d}B_s \qquad \forall\, t \in [0, T]. \tag{REP}$$

The integrand φ is unique up to $\mathrm{d}t \times \mathrm{d}\mathbb{P}$*-Nullsets.*

Proof. First, we show the uniqueness. Assume (REP) holds for another integrand $\psi \in \mathcal{H}^2[0,T]$. Then $\psi - \varphi \in \mathcal{H}^2[0,T]$. It holds that

$$0 = \int_0^t (\psi_s - \varphi_s)\, \mathrm{d}B_s$$

and thus by Itô-isometry

$$0 = \int_0^T \mathbb{E}\big[(\varphi_s - \psi_s)^2\big]\, \mathrm{d}s = \int_{[0,T]\times\Omega} \big[\varphi_s(\omega) - \psi_s(\omega)\big]^2 (\mathrm{d}s \otimes \mathrm{d}\mathbb{P}).$$

Therefore $\psi_s(\omega) = \varphi_s(\omega)$ $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$-almost everywhere. It is now enough to show (REP) for $t = T$, because then we can take $\mathbb{E}[\cdot|\mathcal{F}_t]$ on both sides. To do so the strategy is

Step 1: Showing (REP) for a large class of $X \in \mathcal{L}^2(\mathcal{F}_T)$,

Step 2: Using a density argument.

Step 2 will be omitted here. Let $0 \le t_0 \le t_1 \le ... \le t_N = T$ be a partition of $[0,T]$ and let $(\vartheta_j)_{j=1}^N$. It will be shown that the random variable

$$Z = \prod_{j=1}^N \underbrace{\exp\left(i\vartheta_j (B_{t_j+1} - B_{t_j})\right)}_{=:Z_j} = \prod_{j=1}^N Z_j$$

has a representation (REP). Consider Z_j first. Therefore let

$$X_t := \exp\left(i\vartheta_j B_t + \frac{\vartheta_j^2}{2} t\right).$$

Using Itô's formula yields

$$\begin{aligned} \mathrm{d}X_t &= X_t \left(i\vartheta_j\, \mathrm{d}B_t + \xcancel{\frac{\vartheta_j^2}{2} \mathrm{d}t}\right) + \xcancel{\frac{1}{2} X_t (-\vartheta_j)^2\, \mathrm{d}t} \\ &= i\vartheta_j X_t\, \mathrm{d}B_t. \end{aligned}$$

Thus X is a local martingale. Since $|X_t| \le \exp\left(\frac{\vartheta_j^2}{2} t\right)^2$ follows that X is a martingale. It holds that

$$X_{t+s} = X_s + \int_s^{t+s} i\vartheta_j X_u\, \mathrm{d}B_u.$$

Inserting the definition of X we obtain

$$\exp\left(i\vartheta_j B_{t+s} + \frac{\vartheta_j^2}{2}(t+s)\right) = \exp\left(i\vartheta_j B_s + \frac{\vartheta_j^2}{2} s\right) + i\vartheta_j \int_s^{t+s} \exp\left(i\vartheta_j B_u + \frac{\vartheta_j^2}{2} u\right) \mathrm{d}B_u$$

and

$$Z_j = \exp\big(i\vartheta_j(B_{t+s} - B_s)\big) = e^{-\frac{\vartheta_j^2}{2}t} + i\vartheta_j \int_s^{t+s} \exp\left(i\vartheta_j(B_u - B_s) + \frac{\vartheta_j^2}{2}(u - (t+s))\right) \mathrm{d}B_u.$$

In other words

$$Z_j = \mathbb{E}[Z_j] + \int_0^T \psi_j(s)\,\mathrm{d}B_s$$

with $\psi_j \in \mathcal{H}^2[0,T]$ and $\psi_j(s) = 0$ for $s \notin [t_j, t_{j+1}]$.
Defining now

$$Z_j(t) := \mathbb{E}\big[Z_j \big| \mathcal{F}_t\big] = \mathbb{E}[Z_j] + \int_0^t \psi_j(s)\,\mathrm{d}B_s$$

and using the product rule we obtain for $j \neq k$ that

$$\begin{aligned} \mathrm{d}(Z_j Z_k)_t &= Z_j\,\mathrm{d}Z_k + Z_k\,\mathrm{d}Z_j + \mathrm{d}[Z_j, Z_k]_t \\ &= Z_j\psi_k\,\mathrm{d}B_t + Z_k\psi_j\,\mathrm{d}B_t + \underbrace{\psi_j\psi_k}_{=0}\,\mathrm{d}t. \end{aligned}$$

Thus

$$Z_j Z_k = Z_j(0)Z_k(0) + \int_0^t \big(Z_j(s)\psi_k(s) + Z_k(s)\psi_j(s)\big)\,\mathrm{d}B_s.$$

Iterating N times yields (REP) for $Z = \prod_{j=1}^N Z_j$. □

Corollary 4.17. *Let X be a continuous local martingale with respect to the Brownian standard filtration generated by a one dimensional Brownian motion $(B_t)_{t\geq 0}$. Then exists $\varphi \in \mathcal{L}^2_{loc}[0,T]$, such that*

$$X_t = X_0 + \int_0^t \varphi_s\,\mathrm{d}B_s \qquad \forall\ t \in [0,T].$$

Remark 4.18. Note that there is no uniqueness.

Proof. Consider the localizing sequence $(\tau_n)_{n\in\mathbb{N}}$ defined by

$$\tau_n := \inf\big\{t \geq 0 :\ |X_t| \geq n\big\}.$$

Obviously $X_{t\wedge\tau_n}$ are **bounded** martingales and we can apply Theorem 4.16 and obtain (REP) with φ_n. Setting

$$\varphi(s) := \lim_{n\to\infty} \varphi_n(s)$$

yields the assertion. □

4.4 Risk-Neutral Valuation in Complete Market Models

We consider a financial market model with a single asset S driven by a one-dimensional Brownian motion, i.e.,

$$\begin{cases} \mathrm{d}S_t^0 = S_t^0 r_t \,\mathrm{d}t, \\ \mathrm{d}S_t = S_t(\mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}B_t). \end{cases}$$

Let $\mathcal{F}_t$ be the Brownian standard filtration of $(B_t)_{t\geq 0}$ and r_t, μ_t, σ_t be $(\mathcal{F}_t)$-adapted stochastic processes. We assume that $\sigma_t \neq 0$ holds $(\mathrm{d}t \times \mathrm{d}\mathbb{P})$-almost everywhere. The discounted price process $X_t = \frac{S_t}{S_t^0}$ satisfies

$$\mathrm{d}X_t = X_t\big((\mu_t - r_t)\,\mathrm{d}t + \sigma_t \,\mathrm{d}B_t\big).$$

Example 4.19. Examples for the upper kind of model are

- the Black-Scholes model, where σ_t is constant,
- local volatility models with $\sigma_t = \sigma(S_t)$.

Stochastic volatility models are not of this kind, because $\sigma_t = \sigma(S_t, V_t)$ is not necessarily $(\mathcal{F}_t^B)$-measurable, because another Brownian motion drives V_t.

Definition 4.20. The Process

$$\lambda_t := \frac{\mu_t - r_t}{\sigma_t}$$

is called **market price of risk** (MPR).

Definition 4.21.

a) A probability measure $\mathbb{Q}$ is called **equivalent local martingale measure** (ELMM) if $\mathbb{Q} \sim \mathbb{P}$ and $X = \frac{S}{S_0}$ is a $\mathbb{Q}$-local martingale.

b) A probability measure $\mathbb{Q}$ is called **equivalent martingale measure** (EMM) if $\mathbb{Q} \sim \mathbb{P}$ and $X = \frac{S}{S^0}$ is a $\mathbb{Q}$-martingale.

Theorem 4.22 (Existence and Uniqueness of E(L)MMs)**.**

a) If the market price of risk $\lambda_t = \frac{\mu_t - r_t}{\sigma_t}$ satisfies the $\mathbb{P}$-Novikov condition

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \lambda_t^2 \,\mathrm{d}t\right)\right] < \infty,$$

then a unique equivalent local martingale measure $\mathbb{Q}$ exists and

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left(-\int_0^t \lambda_s \,\mathrm{d}B_s - \frac{1}{2}\int_0^t \lambda_s^2 \,\mathrm{d}s\right).$$

b) If also the $\mathbb{Q}$-Novikov condition

$$\mathbb{E}^{\mathbb{Q}}\left[\exp\left(\frac{1}{2}\int_0^T \sigma_s^2 \,\mathrm{d}s\right)\right] < \infty$$

holds, then $\mathbb{Q}$ is an equivalent martingale measure.

Proof. Set

$$M_t = \exp\left(-\int_0^t \lambda_s \,\mathrm{d}B_s - \frac{1}{2}\int_0^t \lambda_s^2 \,\mathrm{d}s\right).$$

By Novikov's condition M is a $\mathbb{P}$-martingale. By Girsanov's theorem there exists a probability measure $\mathbb{Q}$ such that $\mathbb{Q} \sim \mathbb{P}$ with

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = M_t$$

and

$$B_t^{\mathbb{Q}} = B_t + \int_0^t \lambda_s \,\mathrm{d}s$$

is a $\mathbb{Q}$-Brownian motion. Considering X we obtain

$$\begin{aligned}
\mathrm{d}X_t &= X_t\big((\mu_t - r_t)\,\mathrm{d}t + \sigma_t \,\mathrm{d}B_t\big) \\
&= X_t\left(\cancel{(\mu_t - r_t)\,\mathrm{d}t} + \sigma_t\left((\mathrm{d}B_t^{\mathbb{Q}} - \cancel{\frac{\mu_t - r_t}{\sigma_t}\,\mathrm{d}t}\right)\right) \\
&= X_t\sigma_t \,\mathrm{d}B_t^{\mathbb{Q}}.
\end{aligned}$$

Thus X is a $\mathbb{Q}$-local martingale and therefore $\mathbb{Q}$ is an equivalent local martingale measure. Note that we have proven the existence in a). Furthermore it holds that

$$X_t = \exp\left(\int_0^t \sigma_s \,\mathrm{d}B_s^{\mathbb{Q}} - \frac{1}{2}\int_0^t \sigma_s^2 \,\mathrm{d}s\right).$$

The fact that the $\mathbb{Q}$-Novikov condition is satisfied implies directly that X is a martingale and thus that $\mathbb{Q}$ is an equivalent martingale measure, hence b). We still need to show the uniqueness. Therefore let $\mathbb{Q}'$ be another equivalent local martingale measure. It follows directly that $\mathbb{Q} \sim \mathbb{Q}'$. Hence there exists a Radon-Nikodym derivative

$$Z := \left.\frac{\mathrm{d}\mathbb{Q}'}{\mathrm{d}\mathbb{Q}}\right|_{\mathcal{F}_T}$$

with density process

$$Z_t := \mathbb{E}^{\mathbb{Q}}[Z|\mathcal{F}_t],$$

which is a $\mathbb{Q}$-martingale with $Z_0 = \mathbb{E}^{\mathbb{Q}}[Z] = 1$. The martingale representation theorem yields that there exists a $\gamma \in \mathcal{H}^2[0,T]$ such that

$$Z_t = 1 + \int_0^t \gamma_s \,\mathrm{d}B_s^{\mathbb{Q}}. \tag{22}$$

Since $\mathbb{Q}'$ is an equivalent local martingale measure there exists a localizing sequence (τ_n) such that $X_t^{\tau_n}$ is a $\mathbb{Q}'$-martingale. Hence

$$X_t^{\tau_n} = \mathbb{E}^{\mathbb{Q}'}\left[X_T^{\tau_n} | \mathcal{F}_t\right]$$

and therefore

$$X_t^{\tau_n} = \frac{1}{Z_t^{\tau_n}} \mathbb{E}^{\mathbb{Q}}\left[X_T^{\tau_n} Z_T^{\tau_n} \middle| \mathcal{F}_t\right].$$

Multiplying by $Z_t^{\tau_n}$ we obtain

$$Z_t^{\tau_n} X_t^{\tau_n} = \mathbb{E}^{\mathbb{Q}}\left[X_T^{\tau_n} Z_T^{\tau_n} \middle| \mathcal{F}_t\right].$$

Therefore $Z_t X_t$ is a $\mathbb{Q}$ local martingale. On the other hand we have

$$\begin{aligned} \mathrm{d}(XZ)_t &= X_t \,\mathrm{d}Z_t + Z_t \,\mathrm{d}X_t + \mathrm{d}[X, Z]_t \\ &= X_t \gamma_t \,\mathrm{d}B_t^{\mathbb{Q}} + Z_t X_t \sigma_t \,\mathrm{d}B_t^{\mathbb{Q}} + \underbrace{X_t \sigma_t \gamma_t \,\mathrm{d}t}_{\text{must be 0!}} \end{aligned}$$

Thus $\sigma_t \gamma_t X_t = 0$ $(\mathrm{d}t \times \mathrm{d}\mathbb{P})$-almost everywhere. Obviously it must be γ_t to be zero. With (22) we obtain that $Z = 1$. Hence

$$\left.\frac{\mathrm{d}\mathbb{Q}'}{\mathrm{d}\mathbb{Q}}\right|_{\mathcal{F}_T} = 1,$$

which implies that the measures are equal. □

Example 4.23. By theorem 4.22 the Black-Scholes Model where μ, r, σ are constants has an unique equivalent martingale measure.

Theorem 4.24 (risk-neutral pricing)**.** *Consider an arbitrage-free, one-dimensional, complete financial market model driven by a one-dimensional Brownian motion, i.e.,*

$$\begin{cases} \mathrm{d}S_t^0 = r_t S_t^0 \,\mathrm{d}t, \\ \mathrm{d}S_t = S_t(\mu_t \,\mathrm{d}t + \sigma t \,\mathrm{d}B_t), \end{cases}$$

with $\sigma_t \neq 0$ $(\mathrm{d}t \times \mathrm{d}\mathbb{P})$*-almost everywhere. Assume that an equivalent local martingale measure* $\mathbb{Q}$ *exists and consider a claim* $C \in \mathcal{L}^1(\Omega, \mathcal{F}_T, \mathbb{Q})$*. Then the discounted fair price of* C *is given by the* $\mathbb{Q}$*-martingale*

$$\widetilde{\Pi}_t^C = \mathbb{E}^{\mathbb{Q}}\left[\frac{C}{S_T^0}\middle| \mathcal{F}_t\right].$$

This is the risk-neutral-pricing-formula. Write the martingale representation of $\widetilde{\Pi}_t^C$ *as*

$$\widetilde{\Pi}_t = \mathbb{E}^{\mathbb{Q}}\left[\frac{C}{S_T^0}\right] + \int_0^t \varphi_s \,\mathrm{d}B_s^{\mathbb{Q}}.$$

Then

$$\vartheta_t = \frac{S_T^0 \varphi_t}{S_t \sigma_t}, \qquad \vartheta_t^0 = \left(\widetilde{\Pi}_t^C - \frac{\varphi_t}{\sigma_t}\right)$$

is a replication/hedging strategy for C*. In addition,* $\mathbb{Q}$ *and* Π_t^C *are unique.*

Remark 4.25.

- ⊕ The risk-neutral pricing approach is more general than the PDE-approach, because
    - · there is no Markovian assumption,
    - · C can be any (non-European) claim.
- ⊕ It allows the computation of Π_t^C by Monte-Carlo simulation.
- ⊖ The hedging strategy is not explicit, since φ is not known explicitly.
- ⊖ It only works for models driven by a single Brownian motion and therefore not for stochastic volatility models.

Proof. Set $\widetilde{Y}_t = \mathbb{E}^{\mathbb{Q}}\left[\frac{C}{S_T^0}\middle|\mathcal{F}_t\right]$ as a candidate for the fair price. Then $Y_t = S_t^0\widetilde{Y}_t$ is the undiscounted price. By the martingale representation theorem there exists $\varphi \in \mathcal{H}^2[0,T]$ such that

$$\widetilde{Y}_t = \widetilde{Y}_0 + \int_0^t \varphi_s \,\mathrm{d}B_s^{\mathbb{Q}}.$$

In Itô-notation

$$\mathrm{d}\widetilde{Y}_t = \varphi_t \,\mathrm{d}B_t^{\mathbb{Q}}.$$

Thus

$$\mathrm{d}Y_t = r_t S_t^0 Y_t \,\mathrm{d}t + \varphi_t \,\mathrm{d}B_t^{\mathbb{Q}}.$$

We define a strategy

$$\vartheta_t = \frac{S_t^0\varphi_t}{S_t\sigma_t} \qquad \vartheta_t^0 = \left(\widetilde{Y}_t - \frac{\varphi_t}{\sigma_t}\right)$$

and have to show that this strategy replicates the claim. Consider the value process V_t. It holds that

$$V_t^\vartheta = \vartheta_t^0 \cdot S_t^0 + \vartheta_t S_t = \left(Y_t - \frac{\varphi_t S_t^0}{\sigma_t}\right) + \frac{S_t^0\varphi_t}{\sigma_t} = Y_t.$$

Therefore $V_T^\vartheta = Y_T = C$ which means that the strategy replicates C. We still have to show that the strategy is self-financing. It holds that

$$\begin{aligned}
\vartheta_t^0 \,\mathrm{d}S_t^0 + \vartheta_t \,\mathrm{d}S_t &= \left(\widetilde{Y}_t - \frac{\varphi_t}{\sigma_t}\right) r_t S_t^0 \,\mathrm{d}t + \frac{S_t^0\varphi_t}{S_t\sigma_t} S_t\Big(r_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}B_t^{\mathbb{Q}}\Big) \\
&= \left(Y_t r_t S_t^0 - \frac{\varphi_t r_t}{\sigma_t} S_t^0\right) \mathrm{d}t + \frac{S_t^0\varphi_t}{\sigma_t} r_t \,\mathrm{d}t + S_t^0\varphi_t \,\mathrm{d}B_t^{\mathbb{Q}} \\
&= Y_t r_t S_t^0 \,\mathrm{d}t + S_t^0 \varphi_t \,\mathrm{d}B_t^{\mathbb{Q}} \\
&= \mathrm{d}Y_t,
\end{aligned}$$

which is the self-financing condition. By the replication principle the assertion

$$\widetilde{\Pi}_t^C = \widetilde{V}_t^\vartheta = \mathbb{E}^{\mathbb{Q}}\left[\frac{C}{S_T^0}\middle|\mathcal{F}_t\right]$$

is shown. For the uniqueness consider another equivalent local martingale measure $\mathbb{Q}' \sim \mathbb{Q}$ on $(\Omega, \mathcal{F}_T)$ with density process

$$Z_t = \frac{\mathrm{d}\mathbb{Q}}{\mathbb{Q}'}\bigg|_{\mathcal{F}_t}$$

such that

$$\begin{aligned} Z_t &= \mathbb{E}^{\mathbb{Q}}[Z_T | \mathcal{F}_t], \\ Z_0 &= 1. \end{aligned}$$

By the martingale representation theorem there exists a $\gamma \in \mathcal{H}^2[0, T]$

$$Z_t = 1 + \int_0^t \gamma_s \, \mathrm{d}B_s^{\mathbb{Q}}.$$

Since $\mathbb{Q}'$ is an equivalent local martingale measure it follows that $X_t = \frac{S_t}{S_t^0}$ is a local $\mathbb{Q}'$-martingale and from Lemma 4.10 we conclude that XZ is a local $\mathbb{Q}$-martingale. On the other hand it holds that

$$\begin{aligned} \mathrm{d}(XZ)_t &= X_t \, \mathrm{d}Z_t + Z_t \, \mathrm{d}X_t + \mathrm{d}[X, Z]_t \\ &= X_t \gamma_t \, \mathrm{d}B_t^{\mathbb{Q}} + Z_t \sigma_t X_t \, \mathrm{d}B_t^{\mathbb{Q}} + \sigma_t \gamma_t X_t \, \mathrm{d}t \end{aligned}$$

Again, the $(\mathrm{d}t)$-term must be 0 and since $X_t \sigma_t \neq 0$ it has to hold that $\gamma_t = 0$ $(\mathrm{d}t \times \mathrm{d}\mathbb{P})$-almost everywhere. Therefore

$$Z_T = 1 + \int_0^T 0 \, \mathrm{d}B_t^{\mathbb{Q}} = 1$$

and thus $\mathbb{Q}' = \mathbb{Q}$. □

4.5 Fundamental Theorem of Asset Pricing I

In this section we have a multivariate setup with one numeraire, d risky assets and a k-dimensional Brownian motion $(B_t)_{t \geq 0}$ with the Brownian standard filtration $(\mathcal{F}_t)$. This includes models for many assets as well as stochastic volatility models. In mathematical notation we have for $i = 1, ..., d$

$$\begin{cases} \mathrm{d}S_t^0 = r_t S_t^0 \, \mathrm{d}t, \\ \mathrm{d}S_t^i = S_t^i \big(\mu_t^i \, \mathrm{d}t + \sigma_t^i \, \mathrm{d}B_t \big), \end{cases} \tag{FMM}$$

where

- $B_t \in \mathbb{R}^{k \times 1}$ is a k-dimensional Brownian motion,
- $\mu_t = (\mu_t^1, ..., \mu_t^d)^\top$ is an $\mathbb{R}^d$-valued, $(\mathcal{F}_t)$-adapted process,

- $\Sigma_t = (\sigma_t^1, ..., \sigma_t^d)^\top$ is an $\mathbb{R}^{d\times k}$-valued, $(\mathcal{F}_t)$-adapted process and
- $I = (1, ..., 1)^\top \in \mathbb{R}^d$ is a vector of ones.

There can be additional volatility processes

$$\mathrm{d}V_t^j = a(V_t^j)\,\mathrm{d}t + b(V_t^j)\,\mathrm{d}B_t, \qquad j = d+1, ..., d+n,$$

but they only appear indirectly through $(\mathcal{F}_t)_{t\geq 0}$ and $(B_t)_{t\geq 0}$. We need some basic definitions.

Definition 4.26. A probability measure $\mathbb{Q}$ with $\mathbb{Q} \sim \mathbb{P}$ such that all discounted assets $X^i = \frac{S^i}{S^0}$ are local $\mathbb{Q}$-martingales is called **equivalent local martingale measure**.

Definition 4.27. An $\mathbb{R}^k$-valued, $(\mathcal{F}_t)$-adapted process λ is **called market price of risk** if

$$(\mu_t - r_t I) = \Sigma_t \cdot \lambda_t \tag{MPR}$$

holds $(\mathrm{d}t \times \mathrm{d}\mathbb{P})$-almost surely. Hence that this is the multivariate generalization of $\lambda_t = \frac{\mu_t - r_t}{\sigma_t}$.

Definition 4.28. Let ϑ b an arbitrage strategy. The strategy is called **admissible** if its discounted value process $\widetilde{V}_t^\vartheta$ is bounded from below, i.e.,

$$\widetilde{V}_t^\vartheta \geq -B$$

for some $B > 0$. Hence, B can be interpreted as some finite limit on the credit line.

We write (NA) for "there is no arbitrage". Then the basic statement of the first fundamental theorem of asset pricing is essentially

$$\text{(NA)} \quad \Leftrightarrow \quad \exists \text{ ELMM } \mathbb{Q}.$$

"Essentially" means that "$\Leftarrow$" holds without further conditions and "$\Rightarrow$" holds under mild conditions. We will use the market price of risk λ and show that the existence of an equivalent local martingale measure $\mathbb{Q}$ implies that

- there is no arbitrage,
- the existence of a market price of risk λ.

On the other hand we show that the abscence of arbitrage implies that there exists a market price of risk λ. If this λ satisfies Novikov's condition then there exists an equivalent local martingale measure $\mathbb{Q}$.

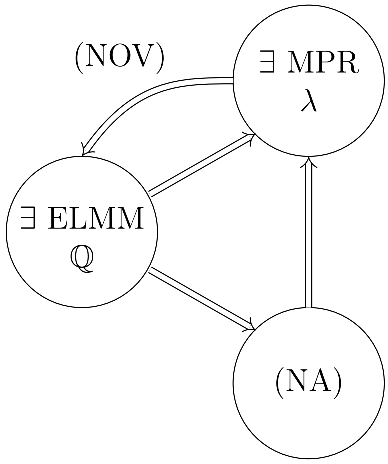


Figure 8: Sketch of the first fundamental theorem of asset pricing

Theorem 4.29 (Fundamental theorem of asset pricing I, "$\Leftarrow$")**.** *If an equivalent local martingale measure* $\mathbb{Q}$ *exists for (FMM), then there is no arbitrage.*

Proof. We will do the proof by contradiction. Let $X = \frac{S}{S^0}$ be the d-dimensional, discounted price and $T > 0$ be the time horizon. Let ϑ be an $\mathbb{R}^d$-valued, admissible arbitrage strategy with discounted value process $\tilde{V}_t^\vartheta$. Thus

$$\mathrm{d}\widetilde{V}_t^\vartheta = \vartheta_t \cdot \mathrm{d}X_t.$$

By assumption X is a local $\mathbb{Q}$-martingale and therefore also $\tilde{V}_t^\vartheta$ is a local $\mathbb{Q}$-martingale with lower bound $-B$. The arbitrage property of ϑ implies that

$$\begin{aligned}\widetilde{V}_0^\vartheta &= 0,\\ \mathbb{P}(\widetilde{V}_T^\vartheta \geq 0) &= 1,\\ \mathbb{P}(\widetilde{V}_T^\vartheta > 0) &> 0.\end{aligned}$$

Since $\mathbb{Q} \sim \mathbb{P}$ it also holds that

$$\begin{aligned}\mathbb{Q}(\widetilde{V}_T^\vartheta \geq 0) &= 1,\\ \mathbb{Q}(\widetilde{V}_T^\vartheta > 0) &> 0.\end{aligned}$$

Therefore $\mathbb{E}^{\mathbb{Q}}[\widetilde{V}_T^\vartheta] > 0$. Now, let (τ_n) be a localizing sequence for $\widetilde{V}^\theta$. By Fatou's Lemma it holds that

$$\mathbb{E}^{\mathbb{Q}}[\widetilde{V}_T^\vartheta] = \mathbb{E}^{\mathbb{Q}}\left[\lim_{n\to\infty} \widetilde{V}_{T\wedge\tau_n}^\vartheta\right] \leq \lim_{n\to\infty} \mathbb{E}^{\mathbb{Q}}[\widetilde{V}_{T\wedge\tau_n}^\vartheta] = \lim_{n\to\infty} \widetilde{V}_0^\vartheta = 0,$$

which is a contradiction. Note that the lower bound $-B$ was needed to apply Fatou's lemma. □

Recall that the market price of risk λ is defined as the solution of

$$\underbrace{\mu_t}_{\substack{\text{growth}\\\text{rate}}} - \underbrace{r_t}_{\substack{\text{interest}\\\text{rate}}} I = \underbrace{\Sigma_t}_{\substack{\text{volatility}\\\text{matrix}\\d\times k}} \underbrace{\lambda_t}_{\text{MPR}}.$$

Corollary 4.30. *If an equivalent local martingale measure $\mathbb{Q}$ exists, then (MPR) has a solution λ.*

Proof. Consider a density process $Z_t := \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\big|_{\mathcal{F}_t}$ for $t \in [0,T]$. Then Z_t is a martingale with $Z_0 = 1$. By the martingale representation theorem there exists $\gamma \in \mathcal{H}^2[0,T]$ such that

$$Z_t = 1 + \int_0^t \gamma_s^\top \,\mathrm{d}B_s.$$

Since $\mathbb{Q}$ is an equivalent local martingale measure it holds that every $X^i = \frac{S^i}{S^0}$ is a local $\mathbb{Q}$-martingale. Therefore XZ is a local $\mathbb{P}$-martingale. Set now $\lambda := -\frac{\gamma_t}{Z_t}$. It holds that

$$\begin{aligned}
\mathrm{d}(X^i Z)_t &= X_t^i \,\mathrm{d}Z_t + Z_t \,\mathrm{d}X_t^i + \mathrm{d}[X^i, Z]_t \\
&= X_t^i \gamma_t \,\mathrm{d}B_t + Z_t X_t^i \big[(\mu_t - r_t)\,\mathrm{d}t + \sigma_t^i \,\mathrm{d}B_t\big] + X_t^i \sigma_t^i \gamma_t \,\mathrm{d}t \\
&= [\cdots]\,\mathrm{d}B_t + \underbrace{Z_t X_t^i}_{>0} \underbrace{\big[(\mu_t^i - r_t) - \sigma_t^i \lambda_t\big]}_{\overset{!}{=}0} \mathrm{d}t.
\end{aligned}$$

We conclude for all $i = 1,..,d$ that

$$\mu_t^i - r_t = \sigma_t^i \cdot \lambda_t$$

$(\mathrm{d}t \times \mathrm{d}\mathbb{P})$-almost everywhere. In matrix-vector notation we obtain

$$\mu_t - r_t \cdot I = \Sigma_t \lambda_t$$

and therefore λ is the market price of risk. □

Theorem 4.31 (Fundamental theorem of asset pricing I, "$\Rightarrow$")**.** *Consider (FMM).*

(a) If (FMM) is free of arbitrage, then a market price of risk λ exists.

(b) If in addition Novikov's condition is satisfied, then an equivalent local martingale measure $\mathbb{Q}$ exists.

Proof. (a) We do this part of the proof by contraposition, assuming

$$\mu_t - r_t I = \Sigma_t \lambda_t$$

has no solution λ_t on a set $A \subseteq [0,T] \times \Omega$ of strictly positive $(\mathrm{d}t \times \mathrm{d}\mathbb{P})$-measure. This means that

$$\mu_t(\omega) - r_t(\omega) I \notin \mathrm{Im}(\Sigma_t(\omega)).$$

From linear algebra[12] we have

$$\operatorname{Im}(\Sigma_t(\omega)) = \ker(\Sigma_t^\top(\omega))^\perp$$

and thus

$$\big(\mu_t(\omega) - r_t(\omega)I\big) \notin \ker(\Sigma_t^\top(\omega))^\perp,$$

i.e.,

$$\exists b_t \in \ker(\Sigma_t^\top(\omega)) \text{ such that } b_t^\top(\mu_t - r_t(\omega)I) \neq 0 \qquad \forall\ (t,\omega) \in A.$$

Therefore we can split $A = A_+ \mathbin{\dot\cup} A_-$ such that

$$\begin{aligned} b_t^\top(\mu_t - r_t I) &> 0 && \forall\ (t,\omega) \in A_+, \\ b_t^\top(\mu_t - r_t I) &< 0 && \forall\ (t,\omega) \in A_-. \end{aligned}$$

Note that at least one of A_+, A_- has strictly positive measure. Now consider first the case that A_+ has strictly positive measure. Define the self-financing strategy

$$\vartheta^i = \frac{b_t^i}{X_t^i}\mathbb{1}_{A_+}, \qquad i = 1, ..., d$$

and let the initial capital be 0. Then for the discounted value process it holds that

$$\begin{aligned} \widetilde{V}_t = 0 + \int_0^t \vartheta_s \,\mathrm{d}X_s &= \int_0^t \sum_{i=1}^d \mathbb{1}_{A_+} \frac{b_t^i}{\cancel{X_t^i}} \cancel{X_t^i} \big[(\mu_t^i - r_t^i)\,\mathrm{d}t + \sigma_t^i \,\mathrm{d}B_t\big] \\ &= \int_0^t \mathbb{1}_{A_+} \big[b_t^\top(\mu_t - r_t I)\,\mathrm{d}t + \underbrace{b_t^\top \Sigma_t}_{=0} \,\mathrm{d}B_t\big] \\ &= \int_0^t \mathbb{1}_{A_+}(s,\omega)\big(\underbrace{b_s^\top(\mu_s - r_s I)}_{>0}\,\mathrm{d}s\big) > 0. \end{aligned}$$

Thus ϑ is an arbitrage, leading to a contradiction. In case A_- using

$$\vartheta^i = -\frac{b_t^i}{X_t^i}\mathbb{1}_{A_-}(t,\omega)$$

we obtain an arbitrage as well.

(b) Assume that the market price of risk λ exists and the Novikov condition holds for λ. Then

$$M_t =: \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\bigg|_{\mathcal{F}_t}$$

defines an equivalent martingale measure $\mathbb{Q} \sim \mathbb{P}$, where

$$M_t = \exp\left(-\int_0^t \lambda_s \,\mathrm{d}B_s - \frac{1}{2}\int_0^t \|\lambda_s\|^2 \,\mathrm{d}s\right)$$

[12] $\mathbb{R}^d = \operatorname{Im}(\Sigma) \oplus \ker(\Sigma^\top)$

as in Girsanov's theorem. Hence,

$$B_t^{\mathbb{Q}} = B_t + \int_0^t \lambda_s \,\mathrm{d}s$$

is a k-dimensional $\mathbb{Q}$-Brownian motion. For the discounted price process it holds that

$$\begin{aligned} \mathrm{d}X_t^i &= X_t^i\big[(\mu_t^i - r_t)\,\mathrm{d}t + \sigma_t^i \,\mathrm{d}B_t\big] \\ &= X_t^i\big[\cancel{(\mu_t^i - r_t)\,\mathrm{d}t} + \sigma_t^i(\mathrm{d}B_t^{\mathbb{Q}} - \cancel{\lambda_t \,\mathrm{d}t})\big] \\ &= X_t^i \sigma_t^i \,\mathrm{d}B_t^{\mathbb{Q}} \end{aligned}$$

and thus the X_t^i are $\mathbb{Q}$-local martingales. Therefore $\mathbb{Q}$ is an equivalent local martingale measure.

□

4.6 Fundamental Theorem of Asset Pricing II

Definition 4.32. A claim C is called **attainable** if there exist a (self-financing) replication/hedging strategy for C.

Definition 4.33. A financial market model is called **complete** if every claim C with bounded discounted payoff $\frac{C}{S_T^0}$ is attainable.

Theorem 4.34 (Fundamental theorem of asset pricing II)**.** *Consider a financial market model with an equivalent local martigale measure $\mathbb{Q}$. Then the following are equivalent.*

a) The financial market model is complete.

b) The equivalet local martingale measure $\mathbb{Q}$ is unique.

c) There is a unique market price of risk λ_t.

Proof. "a)⇒b)": Set $C := S_T^0 \cdot \mathbb{1}_A$ with $A \in \mathcal{F}_T$. By completeness there exists a replication stragtegy ϑ with a discounted value process $\widetilde{V}_t^{\vartheta}$, i.e.,

$$\widetilde{V}_t^{\vartheta} = V_0^{\vartheta} + \int_0^t \vartheta_s \,\mathrm{d}X_s.$$

Since X is a local $\mathbb{Q}$-martingale it follows that $\widetilde{V}_t^{\vartheta}$ is $\mathbb{Q}$-martingale an thus

$$\widetilde{V}_t^{\vartheta} = \mathbb{E}^{\mathbb{Q}}[\widetilde{V}_T^{\vartheta}|\mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_A|\mathcal{F}_t].$$

In particular for $t = 0$ it holds that

$$\widetilde{V}_0^{\vartheta} = \mathbb{Q}(A).$$

Assume now that there exists another equivalent local martingale measure $\mathbb{Q}' \neq \mathbb{Q}$. It holds that

$$\widetilde{V}_t^{\vartheta} = \mathbb{E}^{\mathbb{Q}'}[\mathbb{1}_A|\mathcal{F}_t]$$

and therefore

$$\mathbb{Q}(A)\widetilde{V}_0^{\vartheta} = \mathbb{Q}'(A) \qquad \forall\, A \in \mathcal{F}_T.$$

Thus $\mathbb{Q} = \mathbb{Q}'$ on $(\Omega, \mathcal{F}_T)$.

"b)⇒c)": We show this part by proving "¬c) ⇒ ¬b)". Assume that λ and λ' are two different solutions of (MPR), i.e.,

$$(\mu_t - r_t I) = \Sigma_t \lambda_t = \Sigma_t \lambda'_t.$$

Then it holds that

$$\Sigma_t(\lambda_t - \lambda'_t) = 0.$$

Setting $\gamma_t = \frac{\lambda_t - \lambda'_t}{\|\lambda_t - \lambda'_t\|}$ we obtain $\Sigma_t \gamma_t = 0$, but $\gamma_t \neq 0$ with strictly positive $(\mathrm{d}t \times \mathrm{d}\mathbb{P})$-measure. Hence γ is bounded. Define $\mathbb{Q}'$ by

$$\left.\frac{\mathrm{d}\mathbb{Q}'}{\mathrm{d}\mathbb{Q}}\right|_{\mathcal{F}_t} = \exp\left(-\int_0^t \gamma_s \,\mathrm{d}B_s^{\mathbb{Q}} - \frac{1}{2}\int_0^t \|\gamma_s\|^2 \,\mathrm{d}s\right).$$

Hence

$$B_t^{\mathbb{Q}'} = B_t^{\mathbb{Q}} + \int_0^t \gamma_s \,\mathrm{d}s$$

is a $\mathbb{Q}'$-Brownian motion and it holds that

$$\begin{aligned} \mathrm{d}X_t^i = X_t^i \sigma_t^i \,\mathrm{d}B_t^{\mathbb{Q}} &= X_t^i \sigma_t^i (\mathrm{d}B_t^{\mathbb{Q}'} - \gamma_t \,\mathrm{d}t) \\ &= X_t^i \sigma_t^i B_t^{\mathbb{Q}'} - X_t \underbrace{\sigma_t^i \gamma_t}_{=0} \,\mathrm{d}t \\ &= X_t^i \sigma_t^i \,\mathrm{d}B_t^{\mathbb{Q}'}. \end{aligned}$$

Therefore X^i is a local $\mathbb{Q}'$-martingale. It follows that both $\mathbb{Q}$ and $\mathbb{Q}'$ are equivalent local martingale measures. But, $\mathbb{Q}' \neq \mathbb{Q}$ and thus $\mathbb{Q}$ is not unique.

"c)⇒b)": Again we show the claim by contraposition, proving that "¬b) ⇒ ¬c)". Let

$$Z_t := \left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} \qquad \text{and} \qquad Z'_t := \left.\frac{\mathrm{d}\mathbb{Q}'}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t}.$$

By the martingale representation theorem there exist $\gamma, \gamma' \in \mathcal{H}^2(0,T)$ such that

$$Z_t = 1 + \int_0^t \gamma_s \,\mathrm{d}B_s, \qquad Z'_t = 1 + \int_0^t \gamma'_s \,\mathrm{d}B_s.$$

Set
$$\lambda_t := -\frac{\gamma_t}{Z_t}, \qquad \lambda'_t := -\frac{\gamma'_t}{Z'_t}.$$
Analogously to the proof of Corollary 4.30 we obtain that λ and λ' are both solutions of (MPR).

"b)⇒a)": We will omit this step here.

□

4.7 Summary

Some summarizing remarks:

- In order to avoid arbitrage, every claim C has to be priced by
$$\Pi_t^C = S_t^0 \mathbb{E}^{\mathbb{Q}}\left[\frac{C}{S_T^0}\Big|\mathcal{F}_t\right]$$
where $\mathbb{Q}$ is an equivalent local martingale measure.
- Since
$$\mathrm{d}S_t^0 = r_t S_t^0 \,\mathrm{d}t \qquad \Longrightarrow \qquad S_t^0 = \exp\left(\int_0^t r_s \,\mathrm{d}s\right)$$
we get the equation
$$\Pi_t^C = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s \,\mathrm{d}s\right) C \Big| \mathcal{F}_t\right].$$
- Different choices of $\mathbb{Q}$ lead to different prices Π_t^C. However if C is attainable then all choices of $\mathbb{Q}$ lead to the same price and
$$\Pi_t^C = V_t^\vartheta$$
where V^ϑ is the value process of the replication strategy.

Who chooses the equivalent local martingale measure $\mathbb{Q}$? The market!
How can we find $\mathbb{Q}$? Calibration!

5 American Options and Optimal Stopping

Recall the definition of an **American put** or **call option**. It is the right to sell or buy one unit of an underlying asset S at **any time** $\tau \in [0,T]$ for a fixed price K. Thus the mathematical approach is an optimal stopping problem, i.e., to find the right τ to exercise **optimally**. The optimal time τ has to be modelled as a stopping time. Hence, the decision to stop can only be based on **past** observations.

5.1 The Optimal Stopping Problem

The setup is the following.

- We have a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and
- a **Reward process** $(Z_t)_{t\in[0,T]}$, that is continuous and adapted such that

$$\sup_{0\leq\tau\leq T} \mathbb{E}[|Z_\tau|] < \infty.$$

Definition 5.1. The maximization problem

$$\max \left\{ \mathbb{E}[Z_\tau] : \ \tau \text{ stopping time, } 0 \leq \tau \leq T \right\} \tag{OSP}$$

is called **optimal stopping problem** for Z. A stopping time $\widehat{\tau}$ is called **optimal** for (OSP) if

$$\mathbb{E}[Z_{\widehat{\tau}}] = \sup_{0\leq\tau\leq T} \mathbb{E}[Z_\tau].$$

Example 5.2. If Z is a

- martingale, then any stopping time $0 \leq \tau \leq T$ is optimal,
- supermartingale, then $\tau = 0$ is optimal,
- submartingale, then $\tau = T$ is optimal.

Remark 5.3. A solution η of (OSP) is called **minimal** if $\eta \leq \widehat{\tau}$ for any other solution $\widehat{\tau}$ of (OSP).

Definition 5.4. To (OSP) we associate the **value process**

$$V_t := \sup_{t\leq\tau\leq T} \mathbb{E}[Z_\tau|\mathcal{F}_t],$$

which is the expected reward behaving optimally after t. Any stopping time $\widehat{\tau}_t$ attaining the supremum is called **optimal after** t.

Definition 5.5. The **Snell envelope** of Z is the smallest, right-continuous supermartingale, which dominates Z.

Remark 5.6. A process Y is called **modification of** X if

$$\mathbb{P}(X_t = Y_t) = 1 \qquad \forall\ t \geq 0.$$

Theorem 5.7.

a) The value process V is a modification of the Snell envelope of Z.

b) A stopping time $t \leq \tau_t \leq T$ is optimal after t, if and only if *(i)* $V_{\tau_t} = Z_{\tau_t}$ *almost surely and*

(ii) the stopped value process $(V_{s\wedge\tau_t})_{t\le s\le T}$ *is a martingale*

Proof.

a) Consider the value process

$$V_t = \sup_{t\le\tau\le T} \mathbb{E}[Z_\tau|\mathcal{F}_t].$$

Then it holds that

- $V_T = Z_T$,
- $V_t \ge Z_t$ almost surely for all $t \in [0,T]$, (V dominates Z.)
- for all $t \in [0,T]$

$$V_t \ge \mathbb{E}[Z_T|\mathcal{F}_t] = \mathbb{E}[V_T|\mathcal{F}_t].$$

  (V is a supermartingale.)

From stochastic calculus we know that the above implies that V has a right-continuous modification V^*. We show that V^* is the Snell envelope. Therefore let D be another right-continuous supermartingale dominating Z. Since D dominates Z and because of the supermartingale property we obtain

$$\mathbb{E}[Z_\tau|\mathcal{F}_t] \le \mathbb{E}[D_\tau|\mathcal{F}_t] \le D_t$$

for all stopping times $t \le \tau \le T$. Thus also

$$\underbrace{\sup_{t\le\tau\le T} \mathbb{E}[Z_\tau|\mathcal{F}_t]}_{=V_t=V_t^*} \le D_t.$$

Obviously $V_t^* \le D$ holds for all $t \ge 0$. Since D was an abitrary supermartingale we have shown that V^* is the Snell envelope.

b) "$\Rightarrow$": Let $\widehat{\tau}_t$ be optimal after t for (OSP). Then

$$V_t = \sup_{t\le\tau\le T} \mathbb{E}[Z_\tau|\mathcal{F}_t] = \mathbb{E}[Z_{\widehat{\tau}_t}|\mathcal{F}_t]. \tag{23}$$

In addition, by stopping the supermartingale V, we obtain

$$V_t \le \mathbb{E}[V_{\widehat{\tau}_t\wedge\sigma}|\mathcal{F}_t] \tag{24}$$

for **any** stopping time $t \le \sigma \le T$. We try to prove (i) setting $\sigma = t$. By (23) and (24) it holds that

$$\mathbb{E}[Z_{\widehat{\tau}_t}|\mathcal{F}_t] \ge \mathbb{E}[V_{\widehat{\tau}_t}|\mathcal{F}_t]$$

and thus

$$\mathbb{E}[Z_{\widehat{\tau}_t}] \ge \mathbb{E}[V_{\widehat{\tau}_t}].$$

But since V dominates Z, i.e., $V_s \geq Z_s$ for all $s \in [0,T]$, this can only happen if they are equal almost surely. Hence

$$Z_{\widehat{\tau}_t} = V_{\widehat{\tau}_t}.$$

For (ii) we obtain by (23), (24) and (i) that

$$\begin{aligned}\mathbb{E}[Z_{\widehat{\tau}_t}|\mathcal{F}_t] \geq \mathbb{E}[V_{\widehat{\tau}_t\wedge\sigma}|\mathcal{F}_t] &\geq \mathbb{E}\big[\mathbb{E}[V_{\widehat{\tau}_t}|\mathcal{F}_{\sigma\wedge\widehat{\tau}_t}]\big|\mathcal{F}_t\big] \\ &= \mathbb{E}[V_{\widehat{\tau}_t}|\mathcal{F}_t] \\ &= \mathbb{E}[Z_{\widehat{\tau}_t}|\mathcal{F}_t].\end{aligned}$$

Therefore it holds that

$$\mathbb{E}[V_{\widehat{\tau}_t\wedge\sigma}|\mathcal{F}_t] = \mathbb{E}[V_{\widehat{\tau}_t}|\mathcal{F}_t]$$

for all $t \geq 0$ such that $t \leq \sigma \leq T$. From the converse to Doob's theorem we get that $V_{\widehat{\tau}_t\wedge s}$ is a martingale.

"$\Leftarrow$": By (i), (ii) and since $\widehat{\tau}_t \leq T$ it holds that

$$V_t = V_{t\wedge\widehat{\tau}_t} = \mathbb{E}[V_{T\wedge\widehat{\tau}_t}|\mathcal{F}_t] = \mathbb{E}[V_{\widehat{\tau}_t}|\mathcal{F}_t] = \mathbb{E}[Z_{\widehat{\tau}_t}|\mathcal{F}_t].$$

Thus $\widehat{\tau}_t$ is optimal after t. □

Proposition 5.8. *Let* $\mathbb{E}\left[\sup_{0\leq\tau\leq T}|Z_\tau|\right] < \infty$. *Then an optimal solution* $\widehat{\tau}_t$ *for (OSP) exists for all* $t \in [0,T]$.

Proof. See e.g. Karatzas, I., Shreve, S., *Brownian Motion and Stochastic Calculus*, Springer, 1991. □

Theorem 5.9. *Under the assumptions of Proposition 5.8 the stopping time*

$$\eta_t := \inf\{s > t : \; V_s = Z_s\} \tag{25}$$

is optimal for (OSP) after t. *Hence* η_t *is the first time that* V *and* Z *coincide after* t. *Any other solution* $\widehat{\tau}_t$ *satisfies*

$$\eta_t \leq \widehat{\tau}_t \text{ almost surely,} \tag{26}$$

i.e., η_t *is the minimal solution.*

Proof. We have to proof minimality and optimality. For the minimality let η_t be given by (25) and let $\widehat{\tau}_t$ be optimal. Suppose that $\widehat{\tau}_t(\omega) < \eta_t(\omega)$ for $\omega \in B$ with $\mathbb{P}(B) > 0$. Then it holds that

$$Z_{\widehat{\tau}_t}(\omega) \neq V_{\widehat{\tau}_t}(\omega) \text{ for } \omega \in B.$$

Thus, by Theorem 5.7 $\widehat{\tau}_t$ is not optimal, which is a contradiction. Therefore (26) holds. For the optimality consider

$$V_t = \sup_{t\leq\tau\leq T} \mathbb{E}[Z_\tau|\mathcal{F}_t] \geq \mathbb{E}[Z_{\eta_t}|\mathcal{F}_t].$$

By (25), (26) and because V is a supermartingale we conclude from Theorem 5.7 b) (ii) that

$$\mathbb{E}[Z_{\eta_t}|\mathcal{F}_t] = \mathbb{E}[V_{\eta_t}|\mathcal{F}_t] \geq \mathbb{E}[V_{\widehat{\tau}_t}|\mathcal{F}_t] = V_t.$$

Therefore it holds that

$$\mathbb{E}[Z_{\eta_t}|\mathcal{F}_t] = \sup_{t\leq\tau\leq T} \mathbb{E}[Z_\tau|\mathcal{F}_t]$$

and thus η_t is optimal. □

5.2 Pricing of American Options in Continuous Time

Consider the model

$$\begin{cases} \mathrm{d}S_t = S_t\big(\mu(t,S_t)\,\mathrm{d}t + \sigma(t,S_t)\,\mathrm{d}B_t\big), \\ S_t^0 \equiv 1. \end{cases} \tag{AM-FMM}$$

with $\sigma(t,S_t) \neq 0$ $(\mathrm{d}t \times \mathrm{d}\mathbb{P})$-almost surely. Note that the interest rate is zero here. By definition of the model we know that an unique equivalent local martingale measure $\mathbb{Q}$ exists, hence that the market is complete. Now we consider an American option with payoff $\Phi(\tau,S_\tau)$ at $\tau \in [0,T]$, e.g. $\Phi(t,x) = (K-x)_+$ for a put option. The value, when exercised at τ is

$$\Pi_t^{\mathrm{AM}}(\tau) = \mathbb{E}^{\mathbb{Q}}[\Phi(\tau,S_\tau)|\mathcal{F}_t].$$

The value when exercised optimally is

$$\Pi_t^{\mathrm{AM}} := \sup_{t\leq\tau\leq T} \mathbb{E}^{\mathbb{Q}}[\Phi(\tau,S_\tau)|\mathcal{F}_t],$$

which is exactly the value process of (OSP) with reward $Z_t = \Phi(t,S_t)$. The dynamics of (AM-FMM) under the equivalent local martingale measure $\mathbb{Q}$ is

$$\mathrm{d}S_t = S_t\sigma(t,S_t)\,\mathrm{d}B_t^{\mathbb{Q}}.$$

Note that the solution S_t is a Markov process, which means the law of $(S_r)_{r\in[t,T]}$ conditionally on $\mathcal{F}_t$ only depends on (t,S_t). Set

$$\begin{aligned} J(t,x,\tau) &:= \mathbb{E}^{\mathbb{Q}}[\Phi(\tau,S_\tau)|S_t = x], \\ f(t,x) &:= \sup_{t\leq\tau\leq T} J(t,x,\tau). \end{aligned} \tag{27}$$

Note that (27) is equivalent to (OSP) under the Markov property. Now it follows that

$$\Pi_t^{\mathrm{AM}} = f(t,S_t).$$

Obviously

- $f(t,S_t)$ is the value process of (OSP),
- $\Phi(t,S_t)$ is the reward process of (OSP).

Definition 5.10. The set

$$\mathcal{C} := \{(t,x) \in [0,T] \times \mathbb{R}_{\geq 0} : \ f(t,x) > \Phi(t,x)\} \subseteq [0,T] \times \mathbb{R}_{\geq 0}$$

is called **continuation region**.

From Theorem 5.9 we know that the following stopping time is optimal after t

$$\begin{aligned}\widehat{\eta}_t &:= \inf\{r > t : \ V_r = Z_r\} \\ &= \inf\{r > t : \ f(t, S_t) = \Phi(t, S_t)\} \\ &= \inf\{r > t : \ (t, S_t) \notin \mathcal{C}\}\end{aligned}$$

Note that $\widehat{\eta}_t$ is the first exit time of $\mathcal{C}$.

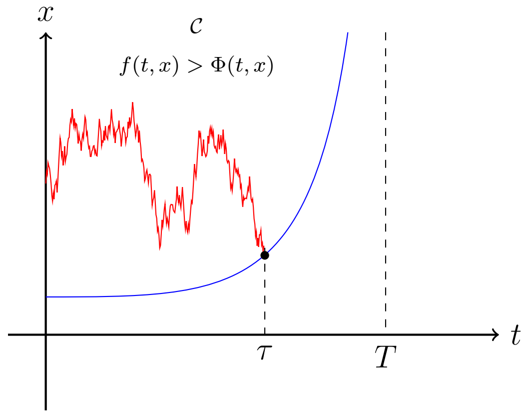


Figure 9: Illustration of the continuation region

Assume $f \in C^{1,2}$ and write

$$\mathcal{A} = \frac{x^2}{2}\sigma(t,x)^2\frac{\partial^2}{\partial x^2},$$

which is the differential operator, generator of S under $\mathbb{Q}$. Applying Itô's formula to f yields

$$\begin{aligned}f(t+h, S_{t+h}) = f(t, S_t) &+ \int_t^{t+h}\left(\frac{\partial}{\partial t}f(r, S_r) + \frac{S_r^2}{2}\sigma(r, S_r)^2\frac{\partial^2}{\partial x^2}f(r, S_r)\right)\mathrm{d}r \\ &+ \underbrace{\int_t^{t+h} S_r\sigma(r, S_r)\frac{\partial}{\partial x}f(r, S_r)\,\mathrm{d}B_r^{\mathbb{Q}}}_{\text{(local) martingale under }\mathbb{Q}}.\end{aligned}$$

We take $\mathbb{E}[\cdot|\mathcal{F}]$ on both sides and obtain

$$\mathbb{E}[f(t+h, S_{t+h}|\mathcal{F}_t] = f(t, S_t) + \mathbb{E}^{\mathbb{Q}}\left[\int_t^{t+h}\left(\frac{\partial}{\partial t} + \mathcal{A}\right)f(r, S_r)\,\mathrm{d}r\middle|\mathcal{F}_t\right]. \tag{28}$$

Now we have to distinguish two cases.

a) Let $(t, S_t) \in \mathcal{C}$. Since $\mathcal{C}$ is open we have that $(t+h, S_{t+h}) \in \mathcal{C}$ for h small enough. Therefore

$$\widehat{\tau}_t > t + h$$

and we have not stopped yet. By Theorem 5.7 the value process $V_r = f(r, S_r)$ is a martingale for $r \in [t, t+h]$. From (28) we obtain that

$$\underbrace{\mathbb{E}^{\mathbb{Q}}\left[\int_t^{t+h}\left(\frac{\partial}{\partial t}+\mathcal{A}\right)f(r,S_r)\,\mathrm{d}r\middle|\mathcal{F}_t\right]}_{=:M_{t,h}} = 0 \qquad \text{for } h \text{ small.}$$

Thus it holds that

$$\lim_{h\to 0}\frac{1}{h}M_{t,h} = 0,$$

which implies

$$\left(\frac{\partial}{\partial t}+\mathcal{A}\right)f(t,S_t) = 0 \qquad \text{whenever } (t,S_t)\in\mathcal{C}.$$

Therefore the assertion

$$\left(\frac{\partial}{\partial t}+\mathcal{A}\right)f(t,x) = 0 \qquad \forall\ (t,x)\in\mathcal{C}$$

holds.

b) Let $(t, S_t) \notin \mathcal{C}$. We stop immediately, i.e., $\widehat{\tau}_t = t$ and obtain

$$f(t,S_t) = \Phi(t,S_t).$$

Moreover by Theorem 5.7 it holds that $V_r = f(r, S_r)$ is a supermartingale. From (28) we know

$$\mathbb{E}^{\mathbb{Q}}\left[\int_t^{t+h}\left(\frac{\partial}{\partial t}+\mathcal{A}\right)f(r,S_r)\,\mathrm{d}r\middle|\mathcal{F}_t\right] \le 0.$$

Thus

$$\left(\frac{\partial}{\partial t}+\mathcal{A}\right)f(t,x) \le 0 \qquad \forall\ (t,x)\notin\mathcal{C}.$$

Summarizing now the above stated we have

$(t,x)\in\mathcal{C}$	$f(t,x) > \Phi(t,x)$	$\left(\frac{\partial}{\partial t}+\mathcal{A}\right)f(t,x) = 0$
$(t,x)\notin\mathcal{C}$	$f(t,x) = \Phi(t,x)$	$\left(\frac{\partial}{\partial t}+\mathcal{A}\right)f(t,x) \le 0$

with terminal condition $f(T,X) = \Phi(T,x)$. We can formulate this as a **free boundary problem**, i.e., given Φ try to find the solutions $f, \mathcal{C}$ for

$$\begin{cases} f(T,x) = \Phi(T,x), & \\ \left(\frac{\partial}{\partial t}+\mathcal{A}\right)f(t,x) = 0 & \forall\ (t,x)\in\mathcal{C}, \\ f(t,x) = \Phi(t,x) & \forall\ (t,x)\in\delta\mathcal{C}. \end{cases}$$

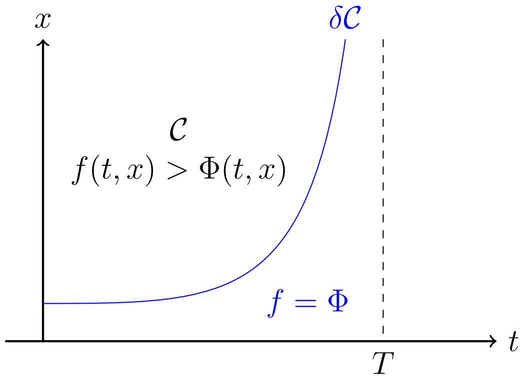


Figure 10: Illustration of the exercise region

We can also formulate it as a **variational problem**, i.e., given Φ find a solution f for

$$\begin{cases} f(T,x) = \Phi(T,x), & \\ \left(\frac{\partial}{\partial t}\right) f(t,x) \le 0 & \forall\ (t,x), \\ \max\left\{\Phi(t,x) - f(t,x), \left(\frac{\partial}{\partial t} + \mathcal{A}\right) f(t,x)\right\} = 0 & \forall\ (t,x), \\ f(t,x) \ge \Phi(t,x) & \forall\ (t,x). \end{cases}$$

Remark 5.11. All the results remain valid for a non-zero interest rate $r \in \mathbb{R}$, when $\mathcal{A}$ is replaced by

$$\mathcal{A} = rx\frac{\partial}{\partial x} + \frac{x^2}{2}\sigma^2(t,x)\frac{\partial^2}{\partial x^2} - r.$$

Corollary 5.12. *It holds that*

$$\left\{(t,x) : \left(\frac{\partial}{\partial t} + \mathcal{A}\right) \Phi(t,x) > 0\right\} \subseteq \mathcal{C},$$

i.e., it is never optimal to stop if

$$\left(\frac{\partial}{\partial t} + \mathcal{A}\right) \Phi(t, S_t) > 0.$$

Proof. Suppose $(t, S_t) \notin \mathcal{C}$ and $\left(\frac{\partial}{\partial t} + \mathcal{A}\right) \Phi(t, S_t) > 0$. Then for small $h \ge 0$ it holds that

$$\left(\frac{\partial}{\partial t} + \mathcal{A}\right) f(t+h, S_{t+h}) > 0.$$

Therefore

$$\mathbb{E}\left[\int_t^{t+h} \left(\frac{\partial}{\partial t} + \mathcal{A}\right)(r, S_r)\,\mathrm{d}r \middle| \mathcal{F}_t\right] > 0.$$

Thus the value process V_r is not a supermartingale on $[t, t+h]$ which is a contradiction. □

Remark 5.13. As in discrete time, it is

- never optimal to exercise the American call option early if $r \geq 0$,
- sometimes optimal to exercise the American put option early if $r < 0$.

6 Numerical Methods in Mathematical Finance

The motivation for the usage of numerical methods is to find a solution to the pricing problem for a claim C:

$$\Pi_0 = \mathbb{E}^{\mathbb{Q}}\left[\frac{C}{S_T^0}\right],$$

where $\mathbb{Q}$ is an equivalent local martingale measure. Examples for claims are

(a) the European call

$$C = (S_T - K)_+,$$

(b) the Asian call option (e.g. in commodity markets)

$$C = \left(\frac{1}{T}\int_0^T S_t \, \mathrm{d}t - K\right)_+,$$

(c) the look-back option

$$C = \left(S_T - \min_{0\leq t\leq T} S_t\right),$$

(d) the basket option

$$C = \left(\sum_{i=1}^N w_i S_t^i - K\right)_+,$$

where $S^1, ..., S^N$is a basket of assets and w_i are weights.

Note that b) and c) have a **path dependent payoff** and d) has a high-dimensional payoff. Here are some methods to determine the fair price.

- In the Black-Scholes Model we have explicit prices for European put/call options.
- The **PDE-method** is used for
  - European options in local/stochastic volatility models,
  - American options in local/stochastic volatility models as free boundary problem.
- The **Fourier method** is used for European options in special models like the Heston model.

- **Monte-Carlo simulation methods** apply in very general circumstances.

Lets have a closer look on the basic idea of the Monte-Carlo simulation. We use the law of large numbers, i.e., for $X \in \mathcal{L}^1$ it holds that

$$\frac{1}{M}\sum_{i=1}^{M} X_i \underset{M\to\infty}{\longrightarrow} \mathbb{E}[X]$$

$\mathbb{P}$-almost surely, where X_i are iid copies of X. This methods has also many other applications e.g. in particle physics, molecular dynamics, computational biology or chemisty. The main questions answerend in this chapter are:

- How do we generate a random sample with an arbitrary distribution?
- What is the convergence rate of Monte-Carlo simulation?
- How to improve convergence of Monte-Carlo simulation?

Furthermore, X may be given only as solution of a SDE, thus:

- How can we approximate this solution?

6.1 Generation of Random Samples

6.1.1 Uniform Random Number Generation

First of all there are no truly random numbers generated on a computer. Therefore we have the following definition.

Definition 6.1. Numbers generated by a computer are called **pseudorandom numbers** if they are generated by deterministic algorithm but statistically indistinguishable from truly random numbers in feasible time.

Note that the set of floating point numbers on a given computer is **finite**, e.g. 2^{36} for single precision or 2^{72} for double precision. Thus the problem reduces to generating a uniform distribution on a **finite set** $\{0, ..., m-1\}$.

Definition 6.2. A **random number generation (RNG)** constists of a finite set X which is called the **state space**, an element $x_0 \in X$ called the **seed**, a **transition function** $T\colon X \to X$ and a function $G\colon X \to \{0, ..., m-1\}$. Pseudorandom numbers i_k are calculated by

$$\begin{aligned} x_{k+1} &= T(x_k), \\ i_k &= G(x_k). \end{aligned}$$

Corollary 6.3. *There is a* ***finite period*** $p \in \mathbb{N}$ *such that*

$$x_{p+k} = x_k \qquad \forall\ k \in \mathbb{N}$$

and therefore $i_{k+p} = i_k$ *for all* $k \in \mathbb{N}$.

Desirable properties of a random number generator are

- statistical uniformity,
- speed,
- large period length,
- reproducibility and
- the possibility of 'jumping ahead', i.e., going quickly from x_k to x_{k+n}.

Definition 6.4. The typical implementation of a random number generator is the **linear congruential generator** defined by

$$x_{k+1} = (ax_k + c) \mod m. \tag{LCG}$$

Proposition 6.5. *The linear congruential generator has full period lenght* $p = m$ *if the following three conditions are satisfied.*

- c *and* a *are relatively prime, i.e., they have no common prime divisors.*
- *Every prime dividing* m *also divides* $a - 1$.
- *If* m *is divisible by* 4*, so is* $a - 1$.

Proof. See number theory. □

6.1.2 Non-Uniform Random Number Generation

Theorem 6.6. *Let* F *be a cumulative distribution function. Then*

$$F^{-1}(u) := \inf\{x :\ F(x) > u\}$$

is called the ***generalized inverse/quantile function****. If* $U \sim \text{Unif}[0,1]$*, then* $X := F^{-1}(U)$ *has distribution* F.

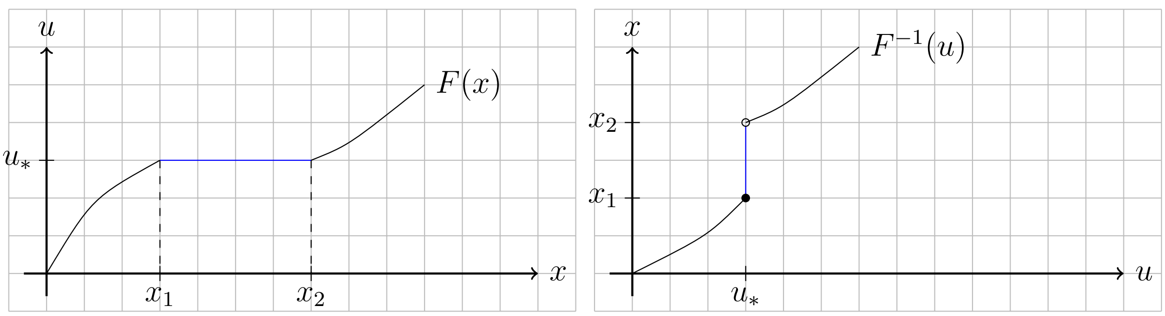


Figure 11: Illustration of $F^{-1}(u)$.

Proof. First we show that

$$F(x) \geq u \;\Leftrightarrow\; F^{-1}(u) \leq x.$$

"$\Rightarrow$": Follows directly by definition.

"$\Leftarrow$": $F^{-1}(u) \leq x$ implies that it exists a sequence $x_n \downarrow x$ such that $F(x_n) \geq u$. By the right-continuity it holds that

$$F(x) = \lim_{n\to\infty} F(x_n) \geq u.$$

Using now the proven equivalence we obtain

$$\mathbb{P}(X \leq x) = \mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(u \leq F(x)) = F(x). \qquad \square$$

Example 6.7. Consider the exponential distribution, i.e.,

$$F(x) = 1 - e^{-\lambda x}.$$

We obtain the inverse function

$$F^{-1}(u) = -\frac{1}{\lambda}\log(1-u).$$

Therefore if $U \sim \text{Unif}[0,1]$ then $X = -\frac{1}{\lambda}\log(1-U)$ is again $\text{Exp}(\lambda)$ distributed.

6.1.3 The Acceptance-Rejection Method

Let f and g be density functions on $\mathbb{R}^d$, where g is the function from which we can sample efficiently and f is the **target density** on $\mathbb{R}^d$, from which we want to sample. The Acceptence-Rejection method can be applied if there exists a $c > 1$ such that

$$f(x) \leq cg(x) \qquad \forall x \in \mathbb{R}^d.$$

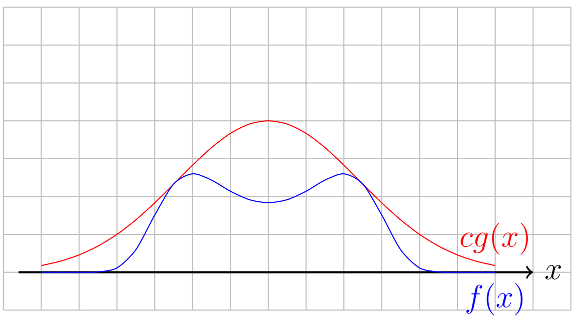


Figure 12: Illustration of the Acceptance-Rejection condition

Algorithm 6.8.

a) Generate a sample X from g and $U \sim \text{Unif}[0,1]$, $U \perp\!\!\!\perp X$.

b) If $U \leq \frac{f(X)}{cg(X)}$ return X, else goto a).

Theorem 6.9. *Let Y be the output of the Acceptance-Rejection method. Then*

a) Y has distribution given by f,

b) the loop in (ARM) is repeated c times on average.

Proof. By construction Y has distribution of X, conditioned on the set

$$\left\{ u \leq \frac{f(x)}{g(x)} \right\},$$

i.e., for all $A \in \mathcal{B}(\mathbb{R}^d)$

$$\begin{aligned} \mathbb{P}(Y \in A) &= \mathbb{P}\left(X \in A \middle| u \leq \frac{f(x)}{cg(x)} \right) \\ &= \frac{\mathbb{P}\left(X \in A \wedge u \leq \frac{f(x)}{cg(x)} \right)}{\mathbb{P}\left(u \leq \frac{f(x)}{cg(x)} \right)} \end{aligned}$$

Moreover it holds that

$$\begin{aligned} \mathbb{P}\left(X \in A \wedge U \leq \frac{f(x)}{cg(x)} \right) &= \int_A \mathbb{P}\left(u \leq \frac{f(x)}{cg(x)} \right) g(x)\,\mathrm{d}x \\ &= \int_A \frac{f(x)}{cg(x)} g(x)\,\mathrm{d}x \\ &= \frac{1}{c} \int_A f(x)\,\mathrm{d}x. \end{aligned}$$

Choosing $A = \mathbb{R}^d$

$$\mathbb{P}\left(u \leq \frac{f(x)}{cg(x)} \right) \overset{A=\mathbb{R}^d}{=} \frac{1}{c}$$

is probability of accepting X or exiting the loop. It holds that

$$\mathbb{P}(Y \in A) = \int_A f(x)\,\mathrm{d}x$$

for all $A \in \mathcal{B}(\mathbb{R}^d)$ and therefore Y has density f. The geometric distribution is averaging the waiting time until acceptance c. □

Remark 6.10. The smaller c is, the more efficient is the Acceptance-Rejection method.

Example 6.11. We try to sample a standard normal distribution from a double exponential distribution, i.e.,

$$g(x) = \frac{1}{2}e^{-|x|}$$

and

$$f(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

We obtain

$$\frac{f(x)}{g(x)} = \frac{2}{\sqrt{\pi}}e^{-\frac{x^2}{2}+|x|} \leq \sqrt{\frac{2e}{\pi}} \approx 1.315 =: c.$$

Thus the Acceptance-Rejection method is efficient.

6.2 Monte-Carlo Simulation

6.2.1 Basics

The goal of the Monte-Carlo simulation is computing

$$I(f, X) := \mathbb{E}[f(X)]$$

using the law of large numbers, i.e.,

$$I(f, X) = \lim_{M\to\infty} \underbrace{\frac{1}{M}\sum_{i=1}^{M} f(X_i)}_{=:I_M(f,X)}$$

where $(X_i)_{i\in 1,\ldots,M}$ are iid copies of X. $I_M(f, X)$ is called **Monte-Carlo estimate**.

Definition 6.12. The **Monte-Carlo error** is defined as

$$\varepsilon_M := \varepsilon_M(f, X) = I(f, X) - I_M(f, X).$$

Note that the Monte-Carlo estimate is unbiased, i.e.,

$$\mathbb{E}[I_M(f, X)] = I(f, X) \qquad \text{or} \qquad \mathbb{E}[\varepsilon_M(f, X)] = 0.$$

The **mean-square error (MSE)** is

$$\text{var}(\varepsilon_M(f, X))$$

and the **root mean-square error (RMSE)** is

$$\sqrt{\text{var}(\varepsilon_M(f, X))}.$$

Theorem 6.13. *Let $\sigma^2 := \operatorname{var}(f(X)) < \infty$. Then the root mean-square error satisfies*

$$\sqrt{\operatorname{var}(\varepsilon_M(f,X))} = \frac{\sigma}{\sqrt{M}}.$$

This means it has order $\frac{1}{2}$.[13] *Moreover $\sqrt{M} \cdot \varepsilon_M(f,X)$ is asymptotically normal with standard deviation σ, i.e.,*

$$\lim_{M\to\infty} \mathbb{P}\left(\frac{a\sigma}{\sqrt{M}} < \varepsilon_M < \frac{b\sigma}{\sqrt{M}}\right) = \Phi(b) - \Phi(a) \qquad \forall\ a < b.$$

Proof. It holds that

$$\operatorname{var}(\varepsilon_M(f,X)) = \operatorname{var}\left(\frac{1}{M}\sum_{i=1}^{M} f(X_i)\right) = \frac{1}{M^2} M \cdot \operatorname{var}(f(X)) = \frac{1}{M} \cdot \underbrace{\operatorname{var}(f(X))}_{=\sigma^2}.$$

The asymptotic normality follows from the central limit theorem. □

Remark 6.14.

(a) There is no deterministic error bound, but we can bound the probability of large errors.

(b) The 'typical error' (RMSE) decreases like $\frac{1}{\sqrt{M}}$, i.e., order $\frac{1}{2}$.

(c) Suppose that we want to control the error like in

$$\mathbb{P}(|\varepsilon_M(f,X)| > \varepsilon) \leq \delta$$

for some $\varepsilon > 0,\ \delta > 0$. Let $\varepsilon' = \frac{\sqrt{M}}{\sigma}\varepsilon$. It holds that

$$\begin{aligned}\mathbb{P}(|\varepsilon_M| > \varepsilon) &= 1 - \mathbb{P}\left(-\frac{\varepsilon'\sigma}{\sqrt{M}} < \varepsilon_M < \frac{\varepsilon'\sigma}{\sqrt{M}}\right) \\ &\approx 1 - (\Phi(\varepsilon') - \Phi(-\varepsilon') = 2 - 2\Phi(\varepsilon') =: \delta.\end{aligned}$$

Thus

$$\Phi\left(\frac{\sqrt{M}}{\sigma}\varepsilon\right) = \frac{1}{2}(1-\delta)$$

Converting the formula yields

$$\begin{aligned}\sqrt{M} &= \frac{\sigma}{\varepsilon}\Phi^{-1}\left(\frac{1}{2}(1-\delta)\right) \\ M &= \frac{\sigma^2}{\varepsilon^2}\left(\Phi^{-1}\left(\frac{1}{2}(1-\delta)\right)\right)^2.\end{aligned}$$

Hence M is proportional to $\frac{1}{\varepsilon^2}$.

(d) In practice, also $\sigma^2 = \operatorname{var}(f(X))$ is unknown.

[13]Because $\sqrt{\cdot} = \cdot^{\frac{1}{2}}$.

6.2.2 Comparison with deterministic integration

Suppose X to be uniformly distributed on $[0,1]^d$. Then

$$\mathbb{E}[f(X)] = \int_{[0,1]^d} f(x)\,\mathrm{d}x.$$

We discretize $[0,1]^d$ on uniform grid points $\{X_1, ..., X_N\}^d$ with spacing $\frac{1}{N}$ and use a deterministic integration method of order k. Thus

$$\text{error} = c \cdot \left(\frac{1}{N}\right)^k$$

where the number of evaluations of f is N^d. Therefore

$$\frac{\text{error}}{\#\text{evaluations}} = c \cdot N^{-\frac{k}{d}},$$

where $N^{-\frac{k}{d}}$ deteriorates with d. This is called the **curse of dimensionality**. The Monte-Carlo method has order $\frac{1}{2}$ independently of the dimension d.

6.2.3 Variance Reduction

We cannot reduce the order of convergence of the Monte-Carlo simulation, but we can try to reduce the variance σ^2. Therefore we try to find a random variable Y and a function g such that

$$\mathbb{E}[g(Y)] = \mathbb{E}[f(X)],$$

but $\mathrm{var}(g(Y)) < \mathrm{var}(f(X))$.

Antithetic Variables

To motivate this method we start with an example.

Example 6.15.

- Let $U \sim \mathrm{Unif}[0,1]$. Then $1-U \sim \mathrm{Unif}[0,1]$ aswell. Therefore

$$\mathbb{E}[f(U)] = \mathbb{E}[f(1-U)].$$

- Let $B \sim N(0, I_d)$. Then $-B \sim N(0, I_d)$ aswell. Therefore

$$\mathbb{E}[f(B)] = \mathbb{E}[f(-B)].$$

The general approach is taking a simple transformation $\widetilde{X}$ of the random variable X such that $\widetilde{X} \stackrel{d}{=} X$. The antithetic Monte-Carlo estimate is then defined as

$$I_M^A(f, X) = \frac{1}{M} \sum_{i=1}^{M} \frac{f(X_i) + f(\widetilde{X}_i)}{2}$$

with a computing time of $o(2M)$. We compare $I_M^A(f, X)$ with $I_{2M}(f, X) = \frac{1}{2M} \sum_{i=1}^{2M} f(X_i)$. The antithetic Monte-Carlo estimate is more efficient, if

$$\operatorname{var}\left(\frac{1}{M} \sum_{i=1}^{M} \frac{f(X_i) + f(\widetilde{X}_i)}{2} \right) < \underbrace{\operatorname{var}\left(\frac{1}{2M} \sum_{i=1}^{2} M f(X_i) \right)}_{=\frac{1}{2M} \operatorname{var}(f(X))}.$$

It holds that

$$\begin{aligned} \operatorname{var}\left(\frac{1}{M} \sum_{i=1}^{M} \frac{f(X_i) + f(\widetilde{X}_i)}{2} \right) &= \frac{1}{M^2} \frac{M}{4} \cdot \operatorname{var}(f(X) + f(\widetilde{X})) \\ &= \frac{1}{4M} \left[2 \operatorname{var}(f(X)) + 2 \operatorname{cov}(f(X), f(\widetilde{X}) \right] \\ &= \frac{1}{2M} \operatorname{var}(f(X)) + \frac{1}{2M} \operatorname{cov}(f(X), f(\widetilde{X}). \end{aligned}$$

Thus the antithetic simulation is more efficient if

$$\operatorname{cov}(f(X), f(\widetilde{X})) < 0.$$

Example 6.16.

- In the Black-Scholes model let

  $$\begin{aligned} S_T &= S_0 \cdot \exp\left(-\frac{\sigma^2 T}{2} + \sigma B_T \right), \\ \widetilde{S}_T &= S_0 \exp\left(-\frac{\sigma^2 T}{2} - \sigma B_T \right). \end{aligned}$$

  Then S_T and $\widetilde{S}_T$ are antithetic random variables.

- Consider the Brownian motion and replace $\Delta B_{t_i} = B_{t_{i+1}} - B_{t_i}$ by

  $$\widetilde{\Delta B_{t_i}} = -\Delta B_{t_i}.$$

  Then we obtain an antithetic copy of B.

Control Variates

Assume that there exists a random variable Y and function g, such that

$$I(g,Y) = \mathbb{E}[g(Y)]$$

is exactly known without Monte-Carlo simulation. Then $g(Y)$ can be used as control variate to compute

$$I(f,X) = \mathbb{E}[f(X) - \lambda \cdot (g(Y) - I(g,Y))]$$

with the Monte-Carlo estimate

$$I_M^{\lambda,C}(f,X) = \frac{1}{M}\sum_{i=1}^{M}\big(f(X_i) - \lambda[g(Y_i) - I(g,y)]\big),$$

where (X_i, Y_i) are iid copies of (X,Y). The variance of the Monte-Carlo error is proportional to

$$\operatorname{var}(f(X_i) - \lambda g(Y_i)) = \operatorname{var}(f(X)) - 2\lambda \operatorname{cov}(f(X), g(Y)) + \lambda^2 \operatorname{var}(g(Y)).$$

This formula is minimized by

$$\lambda_* = \frac{\operatorname{cov}(f(X), g(Y))}{\operatorname{var}(g(Y))}.$$

Inserting yields

$$\begin{aligned}\operatorname{var}(f(X_i) - \lambda_* g(Y_i)) &= \operatorname{var}(f(X)) - \frac{\operatorname{cov}(f(X), g(Y))^2}{\operatorname{var}(g(Y))}\\ &= \operatorname{var}(f(X))(1-\varrho_{XY}^2),\end{aligned}$$

where $\varrho := \operatorname{Corr}(f(X), g(Y))$. Thus the error reduction is proportional to ϱ_{XY}^2 and the control variate $g(Y)$ should be strongly correlated (positively or negatively) to $g(X)$.

Example 6.17. Consider a discretized Asian option with payoff

$$\left(\frac{1}{N}\sum_{i=1}^{N} S_{t_i} - K\right)_+ \tag{AA}$$

in the BS-Model. Replacing the arithmetic average by a geometric average we obtain

$$\left(\left(\prod_{i=1}^{N} S_{t_i}\right)^{\frac{1}{N}} - K\right)_+ . \tag{GA}$$

Note that

$$\left(\prod_{i=1}^{N} S_{t_i}\right)^{\frac{1}{N}} = \exp\left(\frac{1}{N}\sum_{i=1}^{N}\log S_{t_i}\right)$$

is normally distributed. Thus we have the explicit price for the geometric Asian option in the Black-Scholes model. The payoffs (GA) and (AA) are strongly correlated and therefore (GA) is an efficient control variate for (AA).

Importance Sampling

Importance sampling is related to the Acceptance-Rejection Sampling, i.e., we sample more often in regions where variance is high. We assume that

- X has density $p\colon \mathbb{R}^d \to \mathbb{R}$, the target density,
- Y has density $q\colon \mathbb{R}^d \to \mathbb{R}$, the auxilliary density.

It holds that

$$\begin{aligned} I(f, X) &= \mathbb{E}[f(X)] \\ &= \int_{\mathbb{R}^d} f(x)p(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^d} f(x)\frac{p(x)}{q(x)}q(x)\,\mathrm{d}x \\ &= \mathbb{E}\left[f(Y)\frac{p(Y)}{q(Y)}\right] = I\left(f \cdot \frac{p}{q}, Y\right). \end{aligned}$$

This can be interpreted as change of measure with Radon-Nikodym derivative

$$\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}} = \frac{p(x)}{q(x)}.$$

The Monte-Carlo estimate is now

$$I_M\left(f\frac{p}{q}, Y\right) = \frac{1}{M}\sum_{i=1}^{M} f(Y_i)\frac{p(Y_i)}{q(Y_i)},$$

where Y_i are iid copies of Y. The speed-up is governed by the variance

$$\operatorname{var}\left(f(Y)\frac{p(Y)}{q(Y)}\right).$$

Thus it would be Ideal to choose

$$q(x) = p(x)f(x),$$

where the variance of a constant c is zero. But since q is a density it holds that

$$1 = \int q(x)\,\mathrm{d}x = \frac{1}{c}\int p(x)f(x)\,\mathrm{d}x.$$

Therefore $c = \mathbb{E}[f(X)]$, which is unknown. The **guideline** is to choose q approximately proportional to $p \cdot f$, i.e., choose

- q high where $p \cdot f$ is high,

- q low where $p \cdot f$ is low.

Example 6.18. Consider a call option $(S_T - K)_+$ in the Black-Scholes model with $K \gg S_0$ e.g. $K = 150,\ S_0 = 100$. Let

$$S_T = S_0 \exp\left(-\frac{\sigma^2 T}{2} + \sigma B_T\right),$$
$$f(x) = \left(S_0 \exp\left(-\frac{\sigma^2 T}{2} + x\right) - K\right)_+,$$

and $X \sim N(0, \sigma^2 T)$. Then $\mathbb{E}[f(X)]$ is the price of the call option. Let

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp\left(-\frac{x^2}{2\sigma\sqrt{T}}\right),$$
$$q(x) = \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp\left(-\frac{(x - x_T)^2}{2\sigma\sqrt{T}}\right).$$

Let $x_* = \log\left(\frac{K}{S_0}\right) + \frac{\sigma^2 T}{2}$ and $Y \sim N(x_*, \sigma^2 T)$. It holds that

$$\frac{p(x)}{q(x)} = \exp\left(-\frac{1}{2\sigma\sqrt{T}}(x^2 - (x - x_*)^2)\right)$$
$$= \exp\left(-\frac{1}{2\sigma\sqrt{T}} 2x x_* - x_*^2)\right)$$

Thus the improved Monte-Carlo estimate is

$$\frac{1}{M}\sum_{i=1}^{M} f(Y_i)\frac{p(Y_i)}{q(Y_i)}.$$

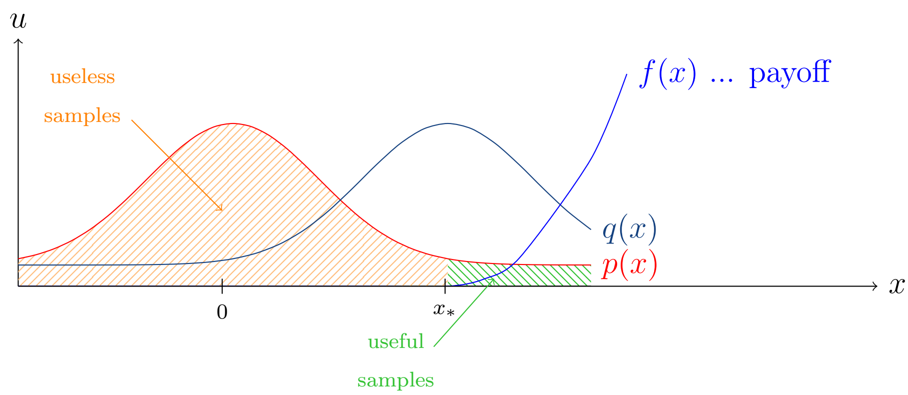


Figure 13: Illustration of importance sampling

6.3 Numerical Methods for SDEs

We try to find a numerical approximation to the solution of the stochastic differential equation

$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + b(X_t)\,\mathrm{d}B_t, \qquad X_0 = x_0 \tag{SDE}$$

e.g. to a local volatility model. The methods that will be discussed can be extended to

- time-dependent coefficients $a(t, X_t),\ b(t, X_t)$,
- the multivariate case, i.e.

$$\begin{aligned} &X\ldots\ \mathbb{R}^k\text{-valued},\\ &B\ldots\ \mathbb{R}^d\text{-valued}. \end{aligned}$$

The approximation $\overline{X}^{\mathcal{D}}$ of X will be given on a time grid

$$\mathcal{D} := \{0 = t_0 < t_1 < \ldots t_N = T\},$$

with mesh of $\mathcal{D}$

$$|D| := \max_{0 \le i \le N-1} |t_{i+1} - t_i| \, .$$

The central question we are asking is the convergence of $\overline{X}^{\mathcal{D}}$ to X as $|\mathcal{D}| \to 0$.

Definition 6.19. A **numerical scheme** is a method, which given x_0, a, b and grid $\mathcal{D}$ returns an approximation $\overline{X}^{\mathcal{D}}$ for X.

The notation is

$$\begin{aligned} \Delta t_i :=& t_{i+1} - t_i, && \\ \Delta Y_i :=& Y_{t_{i+1}} - Y_{t_i} && \text{for any stoch. process } Y,\\ \overline{X}_i^{\mathcal{D}} :=& \overline{X}_{t_i}^{\mathcal{D}} && \text{for the approximation along the grid,}\\ \lfloor t \rfloor & && \text{for the largest gridpoint } \le t. \end{aligned}$$

Definition 6.20.

a) A scheme **converges strongly** to X if

$$\lim_{|\mathcal{D}| \to 0} \mathbb{E}\left[\left|\overline{X}_T^{\mathcal{D}} - X_T\right|\right] = 0.$$

b) The scheme **has strong order** $\gamma > 0$ if

$$\mathbb{E}\left[\left|\overline{X}_T^{\mathcal{D}} - X_T\right|\right] \le C \, |\mathcal{D}|^{\gamma} \, ,$$

where C may depend on x_0, a, b and T.

Definition 6.21. a) Given a family $\mathcal{G}$ of functions $f\colon \mathbb{R} \to \mathbb{R}$, we say $\overline{X}^{\mathcal{D}}$ **converges weakly** to X with respect to $\mathcal{G}$) if

$$\left|\mathbb{E}\left[f\left(\overline{X}_T^{\mathcal{D}}\right)\right] - \mathbb{E}[f(X_T)]\right| = 0 \qquad \forall\, f \in \mathcal{G}.$$

b) We say $\overline{X}^{\mathcal{D}}$ is of **weak order** $\gamma > 0$, if

$$\left|\mathbb{E}\left[f\left(\overline{X}_T^{\mathcal{D}}\right)\right] - \mathbb{E}[f(X_T)]\right| \le C \cdot |\mathcal{D}|^{\gamma} \qquad \forall\, f \in \mathcal{G},$$

where C may depend on x_0, a, b and T.

Remark 6.22.

- Most approximation problems in mathematical finance are of **weak type**, e.g. option pricing.
- The family $\mathcal{G}$ should reflect the relevant application, e.g. it should contain all option payoffs.

Lemma 6.23. *Suppose all functions in $\mathcal{G}$ are Lipschitz with a uniform constant L. Then strong convergence of order γ implies weak convergence of order γ.*

Proof. The Lipschitz continuity and the strong convergence yield

$$\begin{aligned}\left|\mathbb{E}\left[f\left(\overline{X}_T^{\mathcal{D}}\right)\right] - \mathbb{E}[f(X_T)]\right| &\le \mathbb{E}\left[\left|f\left(\overline{X}_T^{\mathcal{D}}\right) - f(X_T)\right|\right] \\ &\le L \cdot \mathbb{E}\left[\left|\overline{X}_T^{\mathcal{D}} - X_T\right|\right] \\ &\le L \cdot C\,|\mathcal{D}|^{\gamma}.\end{aligned}$$

□

6.3.1 The Euler-Maruyama Scheme

The Euler-Maryama scheme is the simplest scheme. It is analogous to the Euler scheme for ODEs. The idea is replacing $\mathrm{d}t$ and $\mathrm{d}B_t$ in (SDE) by discrete forward increments Δt_i and ΔB_i. Thus we consider

$$\overline{X}_{i+1} = \overline{X}_i + a(\overline{X}_i)\Delta t_i + b(\overline{X}_i)\Delta B_i, \qquad \overline{X}_0 = x_0. \tag{EM}$$

Obviously the random variables (ΔB_i) are iid normally distributed. (EM) can be extended between gridpoints by setting

$$\begin{aligned}\overline{X}_t &= \overline{X}_{\lfloor t \rfloor} + a\left(\overline{X}_{\lfloor t \rfloor}\right)(t - \lfloor t \rfloor) + b\left(\overline{X}_{\lfloor t \rfloor}\right)\left(B_t - B_{\lfloor t \rfloor}\right) \\ &= \overline{X}_{\lfloor t \rfloor} + \int_{\lfloor t \rfloor}^{t} a\left(\overline{X}_{\lfloor t \rfloor}\right)\mathrm{d}u + \int_{\lfloor t \rfloor}^{t} b\left(\overline{X}_{\lfloor t \rfloor}\right)\mathrm{d}B_u.\end{aligned} \tag{EM'}$$

Thus

$$\overline{X}_t = X_0 + \int_0^t a\left(\overline{X}_{\lfloor u \rfloor}\right) \mathrm{d}u + \int_0^t b\left(\overline{X}_{\lfloor u \rfloor}\right) \mathrm{d}B_u. \tag{I-EM}$$

We need the following assumptions on a and b.

$$\begin{aligned} \text{Lipschitz continuity:} \quad & |a(x)-a(y)| \leq L \cdot |x-y|\,, \\ & |b(x)-b(y)| \leq L \cdot |x-y|\,, \\ \text{Linear growth:} \quad & |a(x)| \leq C(1+|x|), \\ & |b(X)| \leq C(1+|x|). \end{aligned} \tag{L}$$

Theorem 6.24. *Suppose that the coefficients a, b satisfy the assumptions (L). Then the Euler-Maruyama scheme has strong order $\frac{1}{2}$.*

To prepare the proof we need the following inequalities.

(i) From Jensen's inequality it follows that

$$(a_1, ..., a_n)^2 \leq n \cdot (a_1^2 + ... + a_n^2)$$

for any $a_i \in \mathbb{R}$.

(ii) [**Doob's $\mathcal{L}^2$-inequality**] For any martingale M it holds that

$$\mathbb{E}\left[\sup_{0\leq s\leq t} M_s^2\right] \leq 4 \cdot \mathbb{E}[M_t^2].$$

(iii) [**Hölder's inequality**] Let $\frac{1}{p} + \frac{1}{q} = 1$. Then it holds that

$$\|X \cdot Y\|_1 \leq \|X\|_p \cdot \|Y\|_q\,.$$

(iv) [**Gronwall's inequality**] Let $f \in C([a,b])$ and $f \geq 0$ with

$$f(t) \leq C + D \int_a^t f(u)\,\mathrm{d}u \qquad \forall\; t \in [a,b],$$

for $C, D > 0$. Then it holds that

$$f(t) \leq Ce^{(t-a)D} \qquad \forall\; t \in [a,b].$$

Lemma 6.25. *Let X be the solution of (SDE), where a and b satisfy (L). Then*

$$\mathbb{E}\left[\sup_{0\leq s\leq t} X_s^2\right] \leq C(1+X_0^2) < \infty \qquad \forall\; t \in [0,T],$$

where C may depent on a, b, T.

Proof. Set $\delta_t := \mathbb{E}\left[\sup_{0\le s\le t} X_s^2\right]$. Then by the Definition of (SDE) and (i) it holds that

$$\begin{aligned}
\delta_t &= \mathbb{E}\left[\sup_{0\le s\le t}\left(X_0 + \int_0^s a(X_u)\,\mathrm{d}u + \int_0^s b(X_u)\,\mathrm{d}B_u\right)^2\right] \\
&\le C_1\left(X_0^2 + \underbrace{\mathbb{E}\left[\sup_{0\le s\le t}\left(\int_0^s a(X_u)\,\mathrm{d}u\right)^2\right]}_{=:\alpha_t} + \underbrace{\mathbb{E}\left[\sup_{0\le s\le t}\left(\int_0^s b(X_u)\,\mathrm{d}B_u\right)^2\right]}_{=:\beta_t}\right)
\end{aligned}$$

for some $C_1 > 0$. By (ii), (L) and the Itô-isometry it holds that

$$\begin{aligned}
\beta_t &\le C_2 \cdot \mathbb{E}\left[\left(\int_0^t b(X_u)\,\mathrm{d}B_u\right)^2\right] \\
&= C_2 \int_0^t \mathbb{E}[b(X_u)^2]\,\mathrm{d}u \\
&\le C_3 \cdot \int_0^t (1 + \mathbb{E}[X_u^2])\,\mathrm{d}u \\
&\le C_4 \cdot \int_0^t \delta_u\,\mathrm{d}u
\end{aligned}$$

for some constants $C_2, C_3, C_4 > 0$. For α_t we get a similar estimate using Hölder instead of Doob and Itô. Thus for $C_5 > 0$ it holds that

$$\delta_t \le C_1 X_0^2 + C_5 \cdot \int_0^t \delta_s\,\mathrm{d}s.$$

Now we are able to apply Gronwall's inequality and obtain

$$\begin{aligned}
\delta_t &\le C_1(1 + X_0^2)\exp(tC_5) \\
&\le C(1 + X_0^2)
\end{aligned}$$

for some constant $C := C_1 \exp(TC_5) > 0$. □

Proof of Theorem 6.24. Set

$$\varepsilon_t := \mathbb{E}\left[\sup_{0\le s\le t}(X_s - \overline{X}_s^{\mathcal{D}})^2\right].$$

Using (I-EM) and (SDE) we obtain

$$\begin{aligned}
X_s - \overline{X}_s &= \int_0^s \big(a(X_u) - a(\overline{X}_{\lfloor u\rfloor})\big)\,\mathrm{d}u + \int_0^s \big(b(X_u) - b(\overline{X}_{\lfloor u\rfloor})\big)\,\mathrm{d}B_u \\
&= \alpha_s + \widetilde{\alpha}_s + \beta_s + \widetilde{\beta}_s,
\end{aligned}$$

where

$$\beta_s = \int_0^s \big(b(X_u) - b(X_{\lfloor u \rfloor})\big)\,\mathrm{d}B_u \qquad \text{is the discretization error of exact solution,}$$

$$\widetilde{\beta}_s = \int_0^s \big(b(X_{\lfloor u \rfloor}) - b(\overline{X}_{\lfloor u \rfloor})\big)\,\mathrm{d}B_u \quad \text{is the exact solution-approximation on gridpoints}$$

and similarily for $\alpha, \widetilde{\alpha}$. We try now to derive an estimate for $\widetilde{\beta}$. By Doob's inequality, the Itô-isometry and (L) it holds that

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\le s\le t} \widetilde{\beta}_s^2\right] &\le C \cdot \mathbb{E}[\widetilde{\beta}_t^2] \\
&= C \cdot \int_0^t \mathbb{E}\left[\big(b(X_{\lfloor u \rfloor}) - b(\overline{X}_{\lfloor u \rfloor})\big)^2\right] \mathrm{d}u \\
&\le C \cdot \int_0^t \mathbb{E}\left[\big(X_{\lfloor u \rfloor} - \overline{X}_{\lfloor u \rfloor}\big)^2\right] \mathrm{d}u \\
&\le C \cdot \int_0^t \varepsilon_u \,\mathrm{d}u.
\end{aligned}$$

Again we obtain a similar estimate for $\widetilde{\alpha}$. Looking at an estimate for β we obtain as above

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\le s\le t} \beta_s^2\right] &\le C \cdot \int_0^t \mathbb{E}\left[\big(X_u - X_{\lfloor u \rfloor}\big)^2\right] \mathrm{d}u \\
&= C \cdot \int_0^t \mathbb{E}\left[\left(\int_{\lfloor u \rfloor}^u a(X_s)\,\mathrm{d}s + \int_{\lfloor u \rfloor}^u b(X_s)\,\mathrm{d}B_s\right)^2\right] \\
&\le C \cdot \mathbb{E}\left[\sup_{0\le s\le t} X_s^2\right] \cdot \int_0^t (u - \lfloor u \rfloor)\,\mathrm{d}u.
\end{aligned}$$

By Lemma 6.25 it holds that

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\le s\le t} \beta_s^2\right] &\le C(1 + X_0^2) \cdot T \cdot |\mathcal{D}| \\
&\le C \cdot (1 + X_0^2)\, |\mathcal{D}|\,.
\end{aligned}$$

Inserting the estimates yields

$$\varepsilon_t \le C(1 + X_0^2)\,|\mathcal{D}| + D \int_0^t \varepsilon_s \,\mathrm{d}s.$$

Again Gronwall's inequality is applicable and we obtain

$$\varepsilon t \le C(1 + X_0^2)\,|\mathcal{D}|\,.$$

Therefore and by Hölder's inequality we obtain order $\frac{1}{2}$, because

$$\mathbb{E}\left[\left|X_T - \overline{X}_T^{\mathcal{D}}\right|\right] \le \sqrt{\mathbb{E}\left[\left(X_T - \overline{X}_T^{\mathcal{D}}\right)^2\right]} \le \sqrt{\varepsilon_T} \le C \cdot |\mathcal{D}|^{\frac{1}{2}}\,. \qquad \square$$

Theorem 6.26. *Assume all $f \in \mathcal{G}$ and the coefficients a, b are C^4 with polynomially bounded derivatives. Then the Euler-Maruyama scheme has weak order* 1.

No proof. □

6.3.2 The Milstein Scheme

To improve order of convergence, we focus on the $(b(X_t)\,\mathrm{d}B_t)$-term in (SDE). The Euler-Maruyama scheme approximates

$$b(X_t)\,\mathrm{d}B_t \approx b(X_t)(B_{t+h} - B_t)$$

on the time increment $[t, t+h]$. As integral we write

$$\int_t^{t+h} b(X_s)\,\mathrm{d}B_s \approx b(X_t)(B_{t+h} - B_t). \tag{EM}$$

For the Milstein scheme we will need the **iterated Itô-integral**

$$\int_0^t B_s\,\mathrm{d}B_s.$$

Note that by Itô's formula it holds that

$$B_t^2 = 2 \cdot \int_0^t B_s\,\mathrm{d}B_s + t,$$

which implies

$$\int_0^t B_s\,\mathrm{d}B_s = \frac{1}{2}(B_t^2 - t).$$

Applying Itô's formula to $b(X_s)$ yields

$$\begin{aligned} b(X_s) &= b(X_t) + \int_t^s b'(X_u)\,\mathrm{d}X_u + \frac{1}{2}\int_t^s b''(X_u)\,\mathrm{d}[X,X]_u \\ &= b(X_t) + \int_t^s (...)\,\mathrm{d}u + \int_t^s b'(X_u)b(X_u)\,\mathrm{d}B_u. \end{aligned}$$

Neglecting the $(\mathrm{d}u)$-terms we obtain

$$b(X_s) \approx b(X_t) + b'(X_t)b(X_t)(B_s - B_t).$$

We plug this into (EM) and have

$$\int_t^{t+h} b(X_s)\,\mathrm{d}B_s \approx b(X_t)(B_{t+h} - B_t) + b'(X_t)b(X_t)\int_t^{t+h} (B_s - B_t)\,\mathrm{d}B_s.$$

If we apply the formula for the iterated integral it follows that

$$\int_t^{t+h} (B_s - B_t)\,\mathrm{d}B_s = \frac{1}{2}(B_{t+h}^2 - B_t^2) - \frac{h}{2} - B_t(B_{t+h} - B_t) = \frac{1}{2}(B_{t+h} - B_t)^2 - \frac{h}{2}.$$

Putting all together we obtain

$$\int_t^{t+h} b(X_s)\,\mathrm{d}B_s \approx \underbrace{X_t(B_{t+h} - B_t)}_{\text{EM-scheme}} + \underbrace{\frac{1}{2}b'(X_t)b(X_t)\left((B_{t+h} - B_t)^2 - \frac{h}{2}\right)}_{\text{Milstein correction}}.$$

Finally the Milstein scheme is given by

$$\overline{X}_{i+1} = \overline{X}_i + a(\overline{X}_i)\Delta t_i + b(\overline{X}_i)\Delta B_i + \frac{1}{2}b'(\overline{X}_i)b(\overline{X}_i)\big(\Delta B_i^2 - \Delta t_i\big). \qquad \text{(Milstein)}$$

Theorem 6.27. *Suppose a and b satisfy (L). Moreover let $a \in C^1$ and $b \in C^2$. Then the Milstein scheme has strong order* 1.

No proof. □

Remark 6.28. There are systematic ways to generate higher order schemes, e.g. stochastic Taylor schemes. In these schemes iterated Itô-integrals I_n appear. Iterated Itô-integrals are defined as

$$\begin{aligned} I_t^0 &= 1, \\ I_t^1 &= \int_0^t 1\,\mathrm{d}B_s = B_t, \\ I_t^2 &= \int_0^t B_s\,\mathrm{d}B_s = \frac{1}{2}(B_t^2 - 1), \\ &\vdots \\ I_t^k &= \int_0^t I_s^{(k-1)}\,\mathrm{d}B_s. \end{aligned}$$

⇝ Hermite Polynomials

In reality we have to do Monte-Carlo simulation and the schemes.

6.3.3 The Optimal Ratio of Computations

Consider the stochastic differential equation

$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + b(X_t)\,\mathrm{d}B_t.$$

We try to compute $\mathbb{E}[f(X_t)]$ by a numerical approximation of the SDE and Monte-Carlo simulation. There are some choices to be made, namely

- for the numerical approximation of SDE:
  - Choosing the scheme γ and
  - choosing the number of time steps N.
- for the Monte-Carlo simulation:
  - choosing the number of iterations M.

Figure 14: Illustration of the simulation

Now let $\overline{X}^{\mathcal{D}}$ be an approximation of X and $I_M(f, \overline{X}^{\mathcal{D}})$ be the Monte-Carlo estimate. Then the combined error is

$$\left|I_M(f, \overline{X}_T^{\mathcal{D}}) - \mathbb{E}[f(X_T)]\right| \le \underbrace{\left|I_M(f, \overline{X}_T^{\mathcal{D}}) - \mathbb{E}[f(\overline{X}_T^{\mathcal{D}})]\right|}_{\text{Monte-Carlo error}} + \underbrace{\left|\mathbb{E}[f(\overline{X}_T^{\mathcal{D}})] - \mathbb{E}[f(X_T)]\right|}_{\text{weak error of num. scheme}}$$
$$\le C_1 \cdot M^{-\frac{1}{2}} + C_2 \cdot N^{-\gamma}.$$

Thus the total computational workload is $M \cdot N$. Given an error tolerance $\varepsilon > 0$ we try to minimize the workload, i.e.,

$$\min\{M \cdot N : \; C_1 M^{-\frac{1}{2}} + C_2 \cdot N^{-\gamma} \le \varepsilon\},$$

which is a constrained minimization problem. The Lagrangian is given by

$$\mathcal{L}(M, N, \lambda) = M \cdot N + \lambda(C_1 M^{-\frac{1}{2}} + C_2 N^{-\gamma} - \varepsilon), \qquad \lambda > 0.$$

We solve the problem:

$$\frac{\partial \mathcal{L}}{\partial M} = N + \lambda C_1 \left(-\frac{1}{2} M^{-\frac{3}{2}}\right) \implies N \sim \lambda M^{-\frac{3}{2}} \qquad (*)$$
$$\frac{\partial \mathcal{L}}{\partial N} = M + \lambda C_2 (-\gamma) N^{-(\gamma+1)} \implies M \sim \lambda N^{-(\gamma+1)} = \lambda^{-\gamma} M^{\frac{3(\gamma+1)}{2}}$$
$$\implies \lambda^{\gamma} \sim M^{\frac{3\gamma}{2}+\frac{1}{2}}.$$
$$\implies \lambda \sim M^{\frac{3}{2}+\frac{1}{2\gamma}}.$$

Inserting this into $(*)$ yields

$$N \sim M^{\frac{1}{2\gamma}} \quad \Leftrightarrow \quad M \sim N^{2\gamma}.$$

Therefore it is optimal to choose the number of Monte-Carlo iterations M proportional to $(\#\text{timesteps})^{2\gamma}$. The computational cost in this case is

$$N^{(2\gamma+1)} \sim \varepsilon^{-\left(2+\frac{1}{\gamma}\right)}.$$

This means in terms of the error bound that

$$\varepsilon \sim N^{-\gamma} \quad \Longrightarrow \quad N \sim \varepsilon^{-\frac{1}{\gamma}}.$$

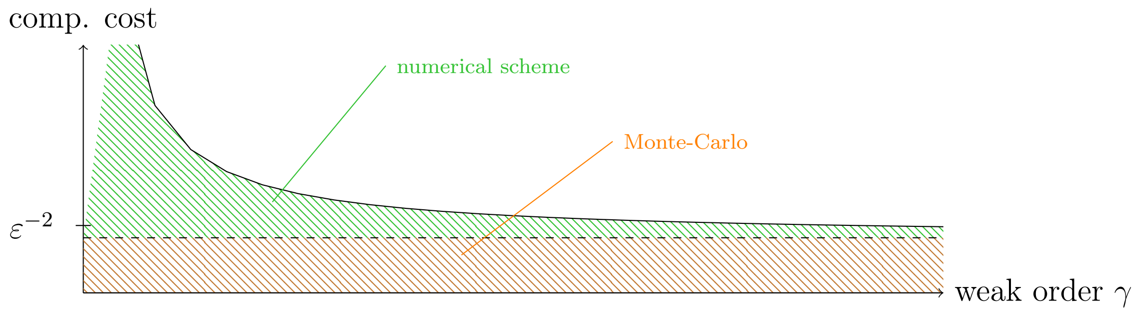


Figure 15: Illustration of the computational cost with respect to the weak order

Concluding we can say that there is no point of using schemes of order > 1 and therefore the Euler-Maruyama scheme and Milstein scheme are good enough.

6.3.4 Multilevel Monte-Carlo (Giles [2008])

Assume that we have a scheme of weak order $\gamma = 1$. Then the combined Euler-Monte-Carlo cost is $o(\varepsilon^{-3})$. The goal is now to reduce the cost.
The idea is to partition $[0, T]$ into several grids $\mathcal{D}$ with uniform time-steps h and

- choosing $h_1 > h_2 > ... > h_L$,
- estimating $f(\overline{X}^{h_1})$, $\mathbb{E}[f(\overline{X}^{h_1})]$ by a Euler-Monte-Carlo crude estimate,
- using $f(\overline{X}^{h_1})$ as a control variate for the next Euler-Monte-Carlo simulation of $f(\overline{X}^{h_2})$, $\mathbb{E}[f(\overline{X}^{h_2})]$,
- repeating this, up to the final grid $\mathcal{D}_{h_L}$.

Note that $f(\overline{X}^{h_1})$ and $f(\overline{X}^{h_2})$ are highly correlated if the same Brownian motion is used. Therefore we have an efficient variance reduction.

Theorem 6.29 (Giles 2008)**.** *Consider a stochastic differential equation with a numerical scheme of weak order* 1 *and strong order* $\frac{1}{2}$*. Fix the error tolerance* $\varepsilon > 0$*. Let*

$$L = \frac{\log\left(\frac{1}{\varepsilon}\right)}{\log(N)}.$$

Then there exists a multilevel Monte-Carlo scheme that respects the error tolerance ε *at a computational cost of*

$$o\big(\varepsilon^{-2} \cdot \log(\varepsilon)^2\big).$$

No proof. □

References

[1] Keller-Ressel, M., *Script for the lecture Stochastische Analysis*, Technische Universität Dresden, 03/2015.

[2] Keller-Ressel, M., *Script for the lecture Finanzmathematik*, Technische Universität Dresden, 07/2016.