Stochastic Differential Equations

Stefan Geiss

September 28, 2023

Contents

Chapter 1

Introduction

One goal of the lecture is to study stochastic differential equations (SDE's). So let us start with a (hopefully) motivating example: Assume that X_t is the share price of a company at time $t \geq 0$ where we assume without loss of generality that $X_0 := 1$. To get an idea of the dynamics of X let us consider the relative increments (these are the increments which are relevant in financial markets)

$$\frac{X_{t+\Delta} - X_t}{X_t} \sim b\Delta + \sigma Y_{t,\Delta}$$

with $b \in \mathbb{R}$, $\sigma > 0$, and $\Delta > 0$ being small. Here $b\Delta$ describes a general trend and $\sigma Y_{t,\Delta}$ some random events (perturbations). Asking several people about this approach we probably get answers like that:

- Statisticians: the random variables $Y_{t,\Delta}$ should be centered Gaussian random variables.

- Mathematicians: the perturbations should not have a memory, otherwise the problem gets too difficult. Hence $Y_{t,\Delta}$ is independent from X_t.

- Then, in addition, probably both of them agree to assume that the perturbations behave additively, that means

  $$Y_{t,\Delta} = Y_{t,\frac{\Delta}{2}} + Y_{t+\frac{\Delta}{2},\frac{\Delta}{2}}$$

  so that $var(Y_{t,\Delta}) = \Delta$ is a good choice.

An approach like this yields to the famous BLACK-SCHOLES option pricing model. Is it possible to make out of this a correct mathematical theory? Yes it is, if we proceed for example in the following way:

Step 1: The random variables $Y_{t,\Delta}$ will be replaced by a continuous time stochastic process $W = (W_t)_{t\geq 0}$, called Brownian motion, such that

$$Y_{t,\Delta} = W_{t+\Delta} - W_t.$$

Consequently, we have to introduce and study the Brownian motion.

Step 2: Our formal equation reads now as

$$X_{t+\Delta} - X_t \sim (bX_t)\Delta + (\sigma X_t)(W_{t+\Delta} - W_t)$$

with $X_0 = 1$ which looks nicer. Letting $\Delta \downarrow 0$ we hope to get a stochastic differential equation

$$dX_t = bX_t dt + \sigma X_t dW_t$$

we are able to explain. To this end we introduce stochastic integrals to be in a position to write the differential equation as an integral equation

$$X_t = X_0 + \int_0^t bX_u du + \int_0^t \sigma X_u dW_u.$$

Step 3: We have to solve this equation. In particular we have to study whether the solutions are unique and to find possible ways to obtain them.

Before we proceed, we give some historic data about groundbreaking work of Kiyoshi Itô and Wolfgang Döblin.

1.0.1 Wolfgang Döblin

Wolfgang Döblin was born in Berlin in 1915 as the son of the famous German writer Alfred Döblin. Since Döblins were Jewish, in 1933 they had to flee from Berlin and take refugee in Paris. There Wolfgang made a strong impression with his skills in maths, and during his short career he published 13 papers and 13 contributions. Fréchet was his adviser, and Döblin also got in touch with Paul Lévy, with whom he wrote his first note. He received his PhD in 1938, but in 1939 he was called up for front service in WWII. In February 1940 Döblin sent his work on diffusions to the Academie des Sciences in

Paris. In June the same year, he decided to take his own life as Nazi troops were just minutes away from the place where he was stationed. In May 2000 his sealed letter was finally opened, and it contained a manuscript called "On Kolmogorov's equation". The manuscript caused a sensation among mathematicians; Wolfgang Döblin had developed a result comparable with the famous Itô's formula, which was published by Itô in 1951.

1.0.2 Kiyoshi Itô

Kiyoshi Itô was born in 1915 in Japan. He studied in the Faculty of Science of the Imperial University of Tokyo, from where he graduated in 1938. The next year he was appointed to the Cabinet Statistics Bureau. He worked there until 1943 and it was during this period that he made his most outstanding contributions, for example Itô's famous paper *On stochastic processes* was published in 1942. This paper is seen today as fundamental, but Itô would have to wait several years before the importance of his ideas would be fully appreciated. In 1945 Itô was awarded his doctorate. He continued to develope his ideas on stochastic analysis, and in 1952 he was appointed to a Professorship at Kyoto University. In the following year he published his famous text *Probability theory.* In this book, Ito develops the theory of a probability space using terms and tools from measure theory. Another important publication by Itô was *Stochastic processes* in 1957, where he studied among others Markov processes and diffusion processes. Itô remained as a professor at Kyoto University until he retired in 1979, and continued to write research papers even after his retirement. He was awarded the Wolf prize in 1987, and the Gauss prize in 2006. Kiyoshi Itô died in Kyoto in 10.11.2008.

Chapter 2

Stochastic processes in continuous time

2.1 Some definitions

In this section we introduce some basic concepts concerning continuous time stochastic processes used freely later on. Let us fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and recall that a map $Z : \Omega \to \mathbb{R}$ is called a random variable if Z is measurable as a map from $(\Omega, \mathcal{F})$ into $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra on $\mathbb{R}$.

What is a continuous-time stochastic process? For us it is simply a family of random variables:

Definition 2.1.1. *Let $I = [0, T]$ for some $T \in (0, \infty)$ or $I = [0, \infty)$. A family of random variables $X = (X_t)_{t \in I}$ with $X_t : \Omega \to \mathbb{R}$ is called stochastic process with index set I.*

The definition of a stochastic process can be given more generally by allowing more general I and other state spaces than $\mathbb{R}$. In our case there are two different views on the stochastic process X:

(1) The family $X = (X_t)_{t \in I}$ describes random functions by

$$\omega \mapsto f(\omega) = (X_t(\omega))_{t \in I}.$$

The function $f(\omega) = (t \mapsto X_t(\omega))$ is called *path* or *trajectory* of X.

(2) The family $X = (X_t)_{t\in I}$ describes a *process*, which is, with respect to the time variable t, an ordered family of random variables $t \mapsto X_t$.

The two approaches differ by the roles of ω and t. Next we ask, when do two stochastic processes X and Y coincide? It turns out to be useful to have several 'degrees of coincidence' as described now.

Definition 2.1.2. Let $X = (X_t)_{t\in I}$ and $Y = (Y_t)_{t\in I}$ be stochastic processes on $(\Omega, \mathcal{F}, \mathbb{P})$. The processes X and Y are *indistinguishable* if and only if

$$\mathbb{P}(X_t = Y_t,\ t \in I) = 1.$$

It should be noted that the definition automatically requires that the set $\{\omega \in \Omega : X_t(\omega) = Y_t(\omega),\ t \in I\}$ is measurable which is not the case in general as shown by the

Example 2.1.3. Let $I = [0, \infty)$, $\Omega = [0, 2)$, $Y_t = 0$,

$$X_t(\omega) := \begin{cases} 0 & : \quad \omega \in [0,1] \\ 0 & : \quad \omega \in (1,2), t \neq \omega \\ 1 & : \quad \omega \in (1,2), t = \omega \end{cases},$$

and $\mathcal{F} := \sigma(X_t : t \geq 0) = \sigma(\{t\} : t \in (1,2))$. But

$$\{\omega \in \Omega : X_t(\omega) = Y_t(\omega),\ t \geq 0\} = [0,1] \notin \mathcal{F}.$$

Another form of coincidence is the following:

Definition 2.1.4. Let $X = (X_t)_{t\in I}$ and $Y = (Y_t)_{t\in I}$ be stochastic processes on $(\Omega, \mathcal{F}, \mathbb{P})$. The processes X and Y are *modifications* of each other provided that

$$\mathbb{P}(X_t = Y_t) = 1 \quad \text{for all} \quad t \in I.$$

Up to now we have to have that the processes are defined on the same probability space. This can be relaxed as follows:

Definition 2.1.5. Let $X = (X_t)_{t\in I}$ and $Y = (Y_t)_{t\in I}$ be stochastic processes on $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\Omega', \mathcal{F}', \mathbb{P}')$, respectively. Then X and Y have the same *finite-dimensional distributions* if

$$\mathbb{P}((X_{t_1}, \ldots, X_{t_n}) \in B) = \mathbb{P}'((Y_{t_1}, \ldots, Y_{t_n}) \in B)$$

for all $0 \leq t_1 < \ldots < t_n \in I$, where $n = 1, 2, \ldots$ and $B \in \mathcal{B}(\mathbb{R}^n)$.

Proposition 2.1.6. (i) *If X and Y are indistinguishable, then they are modifications of each other. The converse implication is not true in general.*

(ii) *If X and Y are modifications from each other, then they have the same finite-dimensional distributions. There are examples of stochastic processes defined on the same probability space having the same finite-dimensional distributions but which are not modifications of each other.*

Proof. (i) Fixing $t \in I$ we immediately get that

$$\mathbb{P}(X_t = Y_t) \geq \mathbb{P}(X_s = Y_s, s \in I) = 1.$$

To construct a counterexample for the converse implication let $I = [0, \infty)$ and $S : \Omega \to \mathbb{R}$ be a random variable such that $S \geq 0$ and $\mathbb{P}(S = t) = 0$ for all $t \geq 0$. Define $X_t := 0$ and

$$Y_t(\omega) := \begin{cases} 0 & : \quad t \neq S(\omega) \\ 1 & : \quad t = S(\omega) \end{cases} = \chi_{\{S(\omega)=t\}}.$$

The map Y_t is a random variable and

$$\mathbb{P}(X_t = Y_t) = \mathbb{P}(0 = Y_t) = \mathbb{P}(S \neq t) = 1,$$

but

$$\{\omega \in \Omega : X_t(\omega) = Y_t(\omega), t \geq 0\} = \{\omega \in \Omega : 0 = Y_t(\omega), t \geq 0\} = \emptyset.$$

(ii) Let $N_t := \{X_t \neq Y_t\}$ so that $\mathbb{P}(N_t) = 0$. Then, for $B \in \mathcal{B}(\mathbb{R}^n)$,

$$\begin{aligned} \mathbb{P}((X_{t_1}, ..., X_{t_n}) \in B) &= \mathbb{P}(\{(X_{t_1}, ..., X_{t_n}) \in B\} \setminus (N_{t_1} \cup \cdots \cup N_{t_n})) \\ &= \mathbb{P}(\{(Y_{t_1}, ..., Y_{t_n}) \in B\} \setminus (N_{t_1} \cup \cdots \cup N_{t_n})) \\ &= \mathbb{P}((Y_{t_1}, ..., Y_{t_n}) \in B). \end{aligned}$$

For the counterexample we let $\Omega = [0, 1]$ and $\mathbb{P}$ be the Lebesgue measure on $[0, 1]$. We choose $X_t = X_0$, $Y_t = Y_0$, $X_0(\omega) := \omega$, and $Y_0(\omega) := 1 - \omega$. $\square$

There are situations in which two processes are indistinguishable when they are modifications of each other.

Proposition 2.1.7. *Assume that X and Y are modifications of each other and that all trajectories of X and Y are left-continuous (or right-continuous). Then the processes X and Y are indistinguishable.*

Proof. In both cases we have that

$$A := \{X_t = Y_t, t \in I\} = \{X_t = Y_t, t \in Q \cap I\}$$

so that $A \in \mathcal{F}$ and

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(X_t = Y_t, t \in Q \cap I) \\ &= 1 - \mathbb{P}(X_t \neq Y_t \text{ for some } t \in Q \cap I) \\ &\geq 1 - \sum_{t \in Q \cap I} \mathbb{P}(X_t \neq Y_t) \\ &= 1.\end{aligned}$$

□

We also need different types of measurability for our stochastic processes. First let us recall the notion of a filtration and a stochastic basis.

Definition 2.1.8. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A family of σ-algebras $(\mathcal{F}_t)_{t \in I}$ is called filtration if $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for all $0 \leq s \leq t \in I$. The quadruple $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in I})$ is called stochastic basis.

The different types of measurability are given by

Definition 2.1.9. Let $X = (X_t)_{t \in I}$, $X_t : \Omega \to \mathbb{R}$ be a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $(\mathcal{F}_t)_{t \in I}$ be a filtration.

(i) The process X is called *measurable* provided that the function $(\omega, t) \to X_t(\omega)$ considered as map between $\Omega \times I$ and $\mathbb{R}$ is measurable with respect to $\mathcal{F} \otimes \mathcal{B}(I)$ and $\mathcal{B}(\mathbb{R})$.

(ii) The process X is called *progressively measurable* with respect to a filtration $(\mathcal{F}_t)_{t \in I}$ provided that for all $s \in I$ the function $(\omega, t) \mapsto X_t(\omega)$ considered as map between $\Omega \times [0, s]$ and $\mathbb{R}$ is measurable with respect to $\mathcal{F}_s \otimes \mathcal{B}([0, s])$ and $\mathcal{B}(\mathbb{R})$.

(iii) The process X is called *adapted* with respect to a filtration $(\mathcal{F}_t)_{t\in I}$ provided that for all $t \in I$ one has that X_t is $\mathcal{F}_t$-measurable.

Proposition 2.1.10. *A process which is progressively measurable is measurable and adapted.*
All other implications between progressively measurable, measurable, *and* adapted *do not hold true in general.*

Proof. (a) progressively measurable $\Longrightarrow$ measurable: Here only the case $I = [0,\infty)$ is of interest. Assume that X is progressively measurable. We show that X is measurable as well. Given $n = 1, 2, ...$ we have that $X^n : \Omega \times [0,n] \to \mathbb{R}$ given by $X^n(\omega,t) := X_t(\omega)$ is measurable with respect to $\mathcal{F}_n \otimes \mathcal{B}([0,n])$ by assumption. Hence the extension $\widetilde{X}^n : \Omega \times [0,\infty) \to \mathbb{R}$ given by

$$\widetilde{X}^n(\omega,t) := X_{t\wedge n}(\omega)$$

is measurable with respect to $\mathcal{F}_n \otimes \mathcal{B}([0,\infty))$ and henceforth with respect to $\mathcal{F} \otimes \mathcal{B}([0,\infty))$. This can be checked considering $\widetilde{X}^n = X^n \circ J_n$ with $J_n : \Omega \times [0,\infty) \to \Omega \times [0,n]$ given by

$$J_n(\omega,t) := (\omega, t \wedge n).$$

Finally we observe that

$$X_t(\omega) = \lim_{n\to\infty} X_{t\wedge n}(\omega)$$

and get that X is $\mathcal{F} \otimes \mathcal{B}([0,\infty))$-measurable as a limit of measurable mappings.

(b) progressively measurable $\Longrightarrow$ adapted: Fix $t \in I$. Then $X^t : \Omega \times [0,t] \to \mathbb{R}$ is measurable with respect to $\mathcal{F}_t \otimes \mathcal{B}([0,t])$ by assumption. Then FUBINI's theorem gives that X_t is $\mathcal{F}_t$-measurable.

(c) Let $I = [0,\infty)$, $B \subseteq [0,\infty)$ be a non-measurable set, and define $X_t := 1$ if $t \in B$ and $X_t := 0$ if $t \notin B$. Then X_t is constant, but

$$\{(\omega,t) \in \Omega \times [0,\infty) : X_t(\omega) = 1\} = \Omega \times B$$

is not measurable. Hence the process is adapted but not measurable or even progressively measurable.

(d) Examples of measurable, but not progressively measurable processes are trivial. □

Proposition 2.1.11. *An adapted process such that all trajectories are left-continuous (or right-continuous) is progressively measurable.*

Proof. We only consider the case of right-continuous paths. For $S = 0$ we easily have that the map $(\omega, 0) \mapsto X_0(\omega)$ is measurable when considered between $(\Omega \times \{0\}, \mathcal{F}_0 \otimes \mathcal{B}(\{0\}))$ and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Consider now $S > 0$. We have to show that $(\omega, t) \mapsto X_t(\omega)$ is measurable when considered between $(\Omega \times [0, S], \mathcal{F}_S \otimes \mathcal{B}([0, S]))$ and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. For $n \in \{1, 2, \ldots\}$ we define $F_n : \Omega \times [0, S] \to \mathbb{R}$ by $(\omega, t) \mapsto X_t^{(n)}(\omega)$ with

$$X_t^{(n)}(\omega) := X_{\frac{k}{2^n} S}(\omega)$$

for $\frac{k-1}{2^n} S < t \leq \frac{k}{2^n} S$ and $k = 1, \ldots, 2^n$, and $X_0^{(n)}(\omega) := X_0(\omega)$. It is clear that F_n is measurable when considered between $(\Omega \times [0, S], \mathcal{F}_S \otimes \mathcal{B}([0, S]))$ and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Hence our claim follows since

$$X_t(\omega) = \lim_{n \to \infty} F_n(\omega, t)$$

by the path-wise right-hand side continuity of X. $\square$

Finally let us recall the notion of a martingale.

Definition 2.1.12. Let $(X_t)_{t \in I}$ be $(\mathcal{F}_t)_{t \in I}$-adapted and such that $\mathbb{E}|X_t| < \infty$ for all $t \geq 0$.

(i) X is called *martingale* provided that for all $0 \leq s \leq t \in I$ one has

$$\mathbb{E}(X_t \mid \mathcal{F}_s) = X_s \text{ a.s.}$$

(ii) X is called *sub-martingale* provided that for all $0 \leq s \leq t \in I$ one has

$$\mathbb{E}(X_t \mid \mathcal{F}_s) \geq X_s \text{ a.s.}$$

(iii) X is called *super-martingale* provided that for all $0 \leq s \leq t \in I$ one has that

$$\mathbb{E}(X_t \mid \mathcal{F}_s) \leq X_s \text{ a.s.}$$

Finally we need some properties of the trajectories:

Definition 2.1.13. Let $X = (X_t)_{t \in I}$ be a stochastic process.

(i) The process X is *continuous* provided that $t \mapsto X_t(\omega)$ is continuous for all $\omega \in \Omega$.

(ii) The process X is *càdlàg* (*continue à droite, limites à gauche*) provided that $t \mapsto X_t(\omega)$ is right-continuous and has left limits for all $\omega \in \Omega$.

(iii) The process X is *càglàd* (*continue à gauche, limites à droite*) provided that $t \mapsto X_t(\omega)$ is left-continuous and has right limits for all $\omega \in \Omega$.

2.2 Basic examples of stochastic processes

Brownian motion. The two-dimensional Brownian motion was observed in 1828 by Robert Brown as diffusion of pollen in water. Later the one-dimensional Brownian motion was used by Louis Bachelier around 1900 in modeling of financial markets and in 1905 by Albert Einstein. A first rigorous proof of its (mathematical) existence was given by Norbert Wiener in 1921. Later on, various different proofs of its existence were given.

Proposition 2.2.1. *There exists a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and a process* $B = (B_t)_{t \geq 0}$ *with* $B_0 \equiv 0$ *such that*

(i) $(B_t)_{t \geq 0}$ *is continuous,*

(ii) *for all* $0 \leq s_1 \leq s_2 \leq \cdots \leq s_n \leq s < t < \infty$ *the random variable* $B_t - B_s$ *is independent of* $(B_{s_1}, ..., B_{s_n})$ *(independent increments),*

(iii) *for all* $0 \leq s < t < \infty$ *one has* $B_t - B_s \sim \mathcal{N}(0, t - s)$ *(stationary increments).*

Definition 2.2.2. A process satisfying the properties of Proposition 2.2.1 is called *standard Brownian motion*.

Poisson process. Let us assume that we have a lamp. Statistics says that the probability that a bulb breaks down is at any time the same. Hence it does not make sense to change a bulb before it is broken down.

How to model this? We need to have a distribution without memory, and this is the exponential distribution. This yields to the following construction: We take independent random variables $\Delta_1, \Delta_2, ... : \Omega \to \mathbb{R}$ with $\Delta_i \geq 0$ and

$$\mathbb{P}(\Delta_i \in B) = \int_B \lambda e^{-\lambda t} dt$$

for $B \in \mathcal{B}(\mathbb{R})$, where $\lambda > 0$ is a parameter to model the time for breaking down. The random variables Δ_i have an exponential distribution with parameter $\lambda > 0$. Letting

$$S_n := \Delta_1 + \cdots + \Delta_n$$

with $S_0 = 0$ gives the time that the nth bulb broke down. The inverse function

$$N_t := \max\{n \geq 0 \, : \, S_n \leqq t\}$$

describes the number of bulbs which were broken down until time t.

Definition 2.2.3. $(N_t)_{t\geq 0}$ is called *Poisson process with intensity* $\lambda > 0$.

Proposition 2.2.4. (i) *$(N_t)_{t\geq 0}$ is a càdlàg process.*

(ii) *$N_t - N_s$ is independent from $(N_{s_1}, ..., N_{s_n})$ for $0 \leq s_1 \leq \cdots s_n \leq s < t < \infty$.*

(iii) *$N_t - N_s$ has a Poisson distribution with parameter $\lambda(t-s)$, that means*

$$\mathbb{P}(N_t - N_s = k) = \tfrac{\mu^k}{k!} e^{-\mu} \textit{ for } \mu = \lambda(t-s).$$

Both processes are Lévy processes which are defined now:

Definition 2.2.5. A process $(X_t)_{t\geq 0}$, $X_0 \equiv 0$ is called *Lévy process* if

(i) X is càdlàg,

(ii) for all $0 \leq s_1 \leq \cdots \leq s_n \leq s < t < \infty$ the random variable $X_t - X_s$ is independent from $(X_{s_1}, ..., X_{s_n})$ (independent increments),

(iii) for all $0 \leq s < t < \infty$ one has that $X_t - X_s$ has the same distribution as X_{t-s} (stationary increments).

2.3 Gaussian processes

Gaussian processes form a class of stochastic processes used in several branches in pure and applied mathematics. Some typical examples are the following:

- In the theory of stochastic processes many processes can be represented and investigated as transformations of the Brownian motion.
- In real analysis the Laplace operator is directly connected to the Brownian motion.
- The modeling of telecommunication traffic, where the fractional Brownian motion is used.

We introduce Gaussian processes in two steps. First we recall Gaussian random variables with values in $\mathbb{R}^n$, then we turn to the processes.

Definition 2.3.1. (i) A random variable $f : \Omega \to \mathbb{R}$ is called *Gaussian* provided that $\mathbb{P}(f = m) = 1$ for some $m \in \mathbb{R}$ or there are $m \in \mathbb{R}$ and $\sigma > 0$ such that

$$\mathbb{P}(f \in B) = \int_B e^{-\frac{(x-m)^2}{2\sigma^2}} \frac{dx}{\sqrt{2\pi}\sigma}$$

for all $B \in \mathcal{B}(\mathbb{R})$. The parameters m and σ^2 are called *expected value* and *variance*, respectively.

(ii) A random vector $f = (f_1, ..., f_n) : \Omega \to \mathbb{R}^n$ is called *Gaussian* provided that for all $a = (a_1, ..., a_n) \in \mathbb{R}^n$ one has that

$$\langle f(\omega), a \rangle := \sum_{i=1}^{n} a_i f_i(\omega)$$

is Gaussian. The parameters $m = (m_1, ..., m_n)$ with $m_i := \mathbb{E} f_i$ and $\sigma = (\sigma_{ij})_{i,j=1}^n$ with

$$\sigma_{ij} := \mathbb{E}(f_i - m_i)(f_j - m_j)$$

are called *mean (vector)* and *covariance (matrix)*, respectively.

For a Gaussian random variable we can compute the expected value and the variance by

$$m = \mathbb{E} f \quad \text{and} \quad \sigma^2 = \mathbb{E}(f - m)^2.$$

Proposition 2.3.2. *Assume Gaussian random vectors $f, g : \Omega \to \mathbb{R}^n$ with the same parameters (m, σ). Then f and g have the same laws.*

Proof. If $\widehat{f}(a) := \mathbb{E}e^{i\langle f,a\rangle}$ and $\widehat{g}(a) := \mathbb{E}e^{i\langle g,a\rangle}$ are the characteristic functions, then by the uniqueness theorem we need to show that $\widehat{f}(a) = \widehat{g}(a)$ for all $a \in \mathbb{R}^n$. For this it is sufficient to have that the distributions of $\langle f, a\rangle$ and $\langle g, a\rangle$ are the same. By assumption both random variables are Gaussian random variables. Hence we only have to check the expected values and the variances. We get

$$\mathbb{E}\langle f, a\rangle = \sum_{i=1}^n a_i \mathbb{E} f_i = \sum_{i=1}^n a_i m_i = \sum_{i=1}^n a_i \mathbb{E} g_i = \mathbb{E}\langle g, a\rangle$$

and

$$\begin{aligned} \mathbb{E}(\langle f, a\rangle - \mathbb{E}\langle f, a\rangle)^2 &= \sum_{i,j=1}^n a_i a_j \mathbb{E}(f_i - m_i)(f_j - m_j) \\ &= \langle \sigma a, a\rangle \\ &= \sum_{i,j=1}^n a_i a_j \mathbb{E}(g_i - m_i)(g_j - m_j) \\ &= \mathbb{E}(\langle g, a\rangle - \mathbb{E}\langle g, a\rangle)^2 \end{aligned}$$

and are done. □

Now we introduce Gaussian processes.

Definition 2.3.3. A stochastic process $X = (X_t)_{t\in I}$, $X_t : \Omega \to \mathbb{R}$, is called *Gaussian* provided that for all $n = 1, 2, ...$ and all $0 \le t_1 < t_2 < \cdots < t_n \in I$ one has that

$$(X_{t_1}, ..., X_{t_n}) : \Omega \to \mathbb{R}^n$$

is a Gaussian random vector. Moreover, we let

$$m_t := \mathbb{E}X_t \quad \text{and} \quad \Gamma(s, t) := \mathbb{E}(X_s - m_s)(X_t - m_t).$$

The process $m = (m_t)_{t\in I}$ is called *mean function* and the process $(\Gamma(s, t))_{s,t\in I}$ *covariance function.*

Up to now we only defined Gaussian processes, however we do not know yet whether they exist. We will prove the existence by analyzing the finite dimensional distributions of a stochastic process $X = (X_t)_{t\in I}$. What are the properties we can expect? From now on we use the index set

$$\Delta := \{(t_1, ..., t_n) : n \geq 1, t_1, ..., t_n \in I \text{ are distinct}\}.$$

Then the family $(\mu_{t_1,...,t_n})_{(t_1,...,t_n)\in\Delta}$ with

$$\mu_{t_1,...,t_n}(B) := \mathbb{P}((X_{t_1}, ..., X_{t_n}) \in B)$$

defines a family of measures such that

$$\begin{aligned} \mu_{t_1,...,t_n}(B_1 \times \cdots \times B_n) &= \mu_{(t_{\pi(1)},...,t_{\pi(n)})}(B_{\pi(1)} \times \cdots \times B_{\pi(n)}), \\ \mu_{t_1,...,t_n}(B_1 \times \cdots \times B_{n-1} \times \mathbb{R}) &= \mu_{t_1,...,t_{n-1}}(B_1 \times \cdots \times B_{n-1}) \end{aligned}$$

for all $B_1, ..., B_n \in \mathcal{B}(\mathbb{R})$ and all permutations $\pi : \{1, ..., n\} \to \{1, ..., n\}$. This is our starting point:

Definition 2.3.4. A family of probability measures $(\mu_{t_1,...,t_n})_{(t_1,...,t_n)\in\Delta}$, where $\mu_{t_1,...,t_n}$ is a measure on $\mathcal{B}(\mathbb{R}^n)$ is called *consistent* provided that

(i) $\mu_{t_1,...,t_n}(B_1 \times \cdots \times B_n) = \mu_{(t_{\pi(1)},...,t_{\pi(n)})}(B_{\pi(1)} \times \cdots \times B_{\pi(n)})$ for all $n = 1, 2, ...,$ $B_1, ..., B_n \in \mathcal{B}(\mathbb{R})$, and all permutations $\pi : \{1, ..., n\} \to \{1, ..., n\}$,

(ii) $\mu_{t_1,...,t_n}(B_1 \times \cdots \times B_{n-1} \times \mathbb{R}) = \mu_{t_1,...,t_{n-1}}(B_1 \times \cdots \times B_{n-1})$ for all $n \geq 2$ and $B_1, ..., B_{n-1} \in \mathcal{B}(\mathbb{R})$.

We show that a consistent family of measures can be derived from one measure. The measure will be defined on the following σ-algebra:

Definition 2.3.5. We let $\mathcal{B}\left(\mathbb{R}^I\right)$ be the smallest σ-algebra which contains all cylinder sets

$$A := \{(\xi_t)_{t\in I} : (\xi_{t_1}, ..., \xi_{t_n}) \in B\}$$

for $(t_1, ..., t_n) \in \Delta$ and $B \in \mathcal{B}(\mathbb{R}^n)$.

Theorem 2.3.6 (Daniell 1918, Kolmogorov 1933)**.** *Assume a consistent family $(\mu_{t_1,...,t_n})_{(t_1,...,t_n)\in\Delta}$ of probability measures. Then there exists a unique probability measure μ on $\mathcal{B}(\mathbb{R}^I)$ such that*

$$\mu((\xi_t)_{t\in I} : (\xi_{t_1}, ..., \xi_{t_n}) \in B) = \mu_{t_1,...,t_n}(B)$$

for all $(t_1, ..., t_n) \in \Delta$ and $B \in \mathcal{B}(\mathbb{R}^n)$.

Proof. Let $\mathcal{A}$ be the algebra of cylinder sets

$$A := \{(\xi_t)_{t\in I} : (\xi_{t_1}, ..., \xi_{t_n}) \in B\}$$

with $(t_1, ..., t_n) \in \Delta$ and $B \in \mathcal{B}(\mathbb{R}^n)$, that means we have that

- $\mathbb{R}^I \in \mathcal{A}$,
- $A_1, ..., A_n \in \mathcal{A}$ implies that $A_1 \cup \cdots \cup A_n \in \mathcal{A}$,
- $A \in \mathcal{A}$ implies that $A^c \in \mathcal{A}$.

Now we define $\nu : \mathcal{A} \to [0, 1]$ by

$$\nu((\xi_t)_{t\in I} : (\xi_{t_1}, ..., \xi_{t_n}) \in B) := \mu_{t_1,...,t_n}(B).$$

In fact, ν is well-defined. Assume that

$$\{(\xi_t)_{t\in I} : (\xi_{s_1}, ..., \xi_{s_m}) \in B\} = \{(\xi_t)_{t\in I} : (\xi_{t_1}, ..., \xi_{t_n}) \in C\}.$$

Let $(r_1, ..., r_N) \in \Delta$ such that $\{r_1, ..., r_N\} = \{s_1, ..., s_m, t_1, ..., t_n\}$. By adding coordinates we find an $D \in \mathcal{B}(\mathbb{R}^N)$ such that the sets above are equal to

$$\{(\xi_t)_{t\geq 0} : (\xi_{r_1}, ..., \xi_{r_N}) \in D\}.$$

By the consistency we have that

$$\mu_{s_1,...,s_m}(B) = \mu_{r_1,...,r_N}(D) = \mu_{t_1,...,t_n}(C).$$

Now we indicate how to check that ν is σ-additive on $\mathcal{A}$, which means that

$$\nu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \nu(A_n)$$

for $A_1, A_2, ... \in \mathcal{A}$, $A_i \cap A_j = \emptyset$ for $i \neq j$, **and** $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. By subtracting $\sum_{n=1}^{N} \nu(A_n)$ and using the finite additivity of ν we see that it is sufficient to prove that

$$\lim_{N\to\infty} \nu\left(\bigcup_{n=N+1}^{\infty} A_n\right) = 0.$$

Letting $C_N := \bigcup_{n=N+1}^{\infty} A_n$ this writes as

$$\lim_n \nu(C_n) = 0$$

for $C_1 \supseteq C_2 \supseteq \cdots \in \mathcal{A}$ with $\bigcap_{n=1}^{\infty} C_n = \emptyset$. Assume that

$$\lim_n \nu(C_n) = 3\varepsilon > 0.$$

Because we only use finitely many coordinates for each C_n we can transform the problem from $\mathbb{R}^I$ to $\mathbb{R}^{\mathbb{N}}$ with $\mathbb{N} = \{1, 2, ...\}$ and consider C_n as cylinder sets in $\mathbb{R}^{\mathbb{N}}$. By the inner regularity of probability measures on $\mathcal{B}(\mathbb{R}^L)$ one can construct compact sets $K_n \subseteq \mathbb{R}^{L_n}$ such that

$$\nu(D_n^0) \geq \nu(C_n) - \varepsilon_n$$

with $\varepsilon_n := \varepsilon/2^{n+1}$, $D_n^0 := K_n \times \mathbb{R}^\infty \subseteq C_n$. We construct the new sequence

$$D_n := D_1^0 \cap \cdots \cap D_n^0$$

and get

- $D_n \subseteq D_n^0 = K_n \times \mathbb{R}^\infty \subseteq C_n$,
- $D_1 \supseteq D_2 \supseteq \cdots$,
- $\nu(D_n) \geq \nu(D_n^0) - (\varepsilon_1 + \cdots + \varepsilon_n) \geq \nu(D_n^0) - \varepsilon \geq \nu(C_n) - 2\varepsilon$.

This estimate can be seen as follows: We have

$$\begin{aligned} \nu(D_n) &= \nu(C_n) - \nu(C_n \setminus D_n) \\ &= \nu(C_n) - \nu(C_n \cap D_n^c) \end{aligned}$$

and $C_n \cap D_n^c = C_n \cap ((D_1^0)^c \cup \cdots \cup (D_n^0)^c) = \bigcup_{k=1}^n (C_n \setminus D_k^0) \subseteq \bigcup_{k=1}^n (C_k \setminus D_k^0)$ Since we have $\nu(C_k \setminus D_k^0) = \nu(C_k) - \nu(D_k^0) \leq \varepsilon_k$ this implies

$$\nu(D_n) \;\geq\; \nu(C_n) - (\varepsilon_1 + \cdots + \varepsilon_n) \geq \nu(D_n^0) - (\varepsilon_1 + \cdots + \varepsilon_n).$$

Hence $\lim_n \nu(D_n) \geq \varepsilon > 0$. Hence $\lim_n \nu(D_n) \geq \varepsilon > 0$. By construction we have $\bigcap_{n=1}^{\infty} D_n \subseteq \bigcap_{n=1}^{\infty} C_n = \emptyset$. However all coordinate sections of the D_n are compact in $\mathbb{R}$ and decreasing so that there is non-empty intersection which disproves that $\bigcap_{n=1}^{\infty} D_n = \emptyset$.

Having proved this we are allowed to apply CARATHÉODORY's extension theorem and obtain the desired probability measure. □

As an application we get the following

Proposition 2.3.7. *Let $(\Gamma(s,t))_{s,t\in I}$ be positive semi-definite and symmetric, that means*

$$\sum_{i,j=1}^{n} \Gamma(t_i,t_j)a_i a_j \geq 0 \quad \textit{and} \quad \Gamma(s,t) = \Gamma(t,s)$$

for all $s,t,t_1,...,t_n \in I$ and $a_1,...,a_n \in \mathbb{R}$. Then there exists a probability space $(\Omega,\mathcal{F},\mathbb{P})$ and a Gaussian process $X = (X_t)_{t\in I}$ defined on $(\Omega,\mathcal{F},\mathbb{P})$ with

(i) $\mathbb{E}X_t = 0$,

(ii) $\mathbb{E}X_s X_t = \Gamma(s,t)$.

Remark 2.3.8. Given any stochastic process $X = (X_t)_{t\in I} \subseteq L_2$ with $\mathbb{E}X_t = 0$ and $\Gamma(s,t) := \mathbb{E}X_s X_t$ we always have that Γ is positive semi-definite and symmetric.

Proof of Proposition 2.3.7. We will construct a consistent family of probability measures. Given $(t_1,...,t_n) \in \Delta$, we let $\mu_{t_1,...,t_n}$ be the Gaussian measure on $\mathbb{R}^n$ with mean zero and covariance

$$\int_{\mathbb{R}^n} \xi_i \xi_j d\mu_{t_1,...,t_n}(\xi_1,...,\xi_n) = \Gamma(t_i,t_j).$$

If the measure exists, then it is unique. To obtain the measure we let $C := (\Gamma(t_i,t_j))_{i,j=1}^n$, so that C is symmetric and positive semi-definite. We know from algebra that there is a matrix A such that $C = AA^T$. Let γ_n be the standard Gaussian measure on $\mathbb{R}^n$ and μ be the image with respect to $A : \mathbb{R}^n \to \mathbb{R}^n$. Then

$$\int_{\mathbb{R}^n} \langle x, e_i \rangle d\mu(x) = \int_{\mathbb{R}^n} \langle Ax, e_i \rangle d\gamma_n(x) = 0$$

and

$$\begin{aligned}
\int_{\mathbb{R}^n} \langle x, e_i \rangle \langle x, e_j \rangle d\mu(x) &= \int_{\mathbb{R}^n} \langle Ax, e_i \rangle \langle Ax, e_j \rangle d\gamma_n(x) \\
&= \int_{\mathbb{R}^n} \langle x, A^T e_i \rangle \langle x, A^T e_j \rangle d\gamma_n(x)
\end{aligned}$$

$$
\begin{aligned}
&= \langle A^T e_i, A^T e_j \rangle \\
&= \langle e_i, AA^T e_j \rangle \\
&= \langle e_i, Ce_j \rangle.
\end{aligned}
$$

The defined family of measures is easily seen to be consistent: given a permutation $\pi : \{1, ..., n\} \to \{1, ..., n\}$ we have that the covariance of $\mu_{t_{\pi(1)},...,t_{\pi(n)}}$ is $\Gamma(t_{\pi(i)}, t_{\pi(j)})$ which proves property (i). Hence $\mu_{t_{\pi(1)},...,t_{\pi(n)}}$ can be obtained from μ by permutation of the coordinates. To prove that

$$\mu_{t_1,...,t_{n-1},t_n}(B_1 \times \cdots \times B_{n-1} \times \mathbb{R}) = \mu_{t_1,...,t_{n-1}}(B_1 \times \cdots \times B_{n-1})$$

we consider the linear map $A : \mathbb{R}^n \to \mathbb{R}^{n-1}$ with $A(\xi_1, ..., \xi_n) := (\xi_1, ..., \xi_{n-1})$ so that

$$A^{-1}(B_1 \times \cdots \times B_{n-1}) = B_1 \times \cdots \times B_{n-1} \times \mathbb{R}$$

and

$$\mu_{t_1,...,t_{n-1},t_n}(B_1 \times \cdots \times B_{n-1} \times \mathbb{R}) = \mu_{t_1,...,t_{n-1},t_n}(A^{-1}(B_1 \times \cdots \times B_{n-1}))$$

and we need to show that

$$\nu := \mu_{t_1,...,t_{n-1},t_n}(A^{-1}(\cdot)) = \mu_{t_1,...,t_{n-1}}.$$

The measure ν is the image measure of $\mu_{t_1,...,t_{n-1},t_n}$ with respect to A so that it is a Gaussian measure. Moreover,

$$
\begin{aligned}
&\int_{\mathbb{R}^{n-1}} \eta_i \eta_j d\nu(\eta_1, ..., \eta_{n-1}) \\
&= \int_{\mathbb{R}^n} \langle A\xi, e_i \rangle \langle A\xi, e_j \rangle d\mu_{t_1,...,t_n}(\xi) \\
&= \int_{\mathbb{R}^n} \langle \xi, A^T e_i \rangle \langle \xi, A^T e_j \rangle d\mu_{t_1,...,t_n}(\xi) \\
&= \sum_{k,l=1}^{n} \langle e_k, A^T e_i \rangle \langle e_l, A^T e_j \rangle \int_{\mathbb{R}^n} \langle \xi, e_k \rangle \langle \xi, e_l \rangle d\mu_{t_1,...,t_n}(\xi) \\
&= \sum_{k,l=1}^{n} \langle e_k, A^T e_i \rangle \langle e_l, A^T e_j \rangle \sigma_{kl} \\
&= \sum_{k,l=1}^{n} \langle Ae_k, e_i \rangle \langle Ae_l, e_j \rangle \sigma_{kl}
\end{aligned}
$$

$$\begin{aligned} &= \sum_{k,l=1}^{n-1} \langle e_k, e_i\rangle \langle e_l, e_j\rangle \sigma_{kl} \\ &= \sigma_{ij}. \end{aligned}$$

Now the process $X = (X_t)_{t\in I}$ is obtained by $X_t : \mathbb{R}^I \to \mathbb{R}$ with $X_t((\xi_s)_I) := \xi_t$. □

Let us consider some examples where we use Proposition 2.3.7 for the existence of Gaussian processes.

Example 2.3.9 (Brownian motion)**.** For $I = [0, \infty)$ we let

$$\Gamma(s,t) := \min\{s,t\} = \int_0^\infty \chi_{[0,s]}(\xi)\chi_{[0,t]}(\xi)d\xi$$

so that

$$\begin{aligned} \sum_{i,j=1}^{n} \Gamma(t_i,t_j)a_i a_j &= \int_0^\infty \sum_{i,j=1}^{n} a_i\chi_{[0,t_i]}(\xi)a_j\chi_{[0,t_j]}(\xi)d\xi \\ &= \int_0^\infty \left(\sum_{i=1}^{n} a_i\chi_{[0,t_i]}(\xi)\right)^2 d\xi \\ &\geq 0. \end{aligned}$$

Example 2.3.10 (Brownian bridge)**.** For $I = [0,1]$ we let

$$\Gamma(s,t) := \begin{cases} s(1-t) & : \quad 0 \le s \le t \le 1 \\ t(1-s) & : \quad 0 \le t \le s \le 1 \end{cases}$$

and want to get a *Gaussian* process returning to zero at time $T = 1$. The easiest way to show that Γ is positive semi-definite is to find one realization of this process: we take the Brownian motion $W = (W_t)_{t\geq 0}$ like in Example 2.3.9, let

$$X_t := W_t - tW_1$$

and get that

$$\begin{aligned} \mathbb{E}X_sX_t &= \mathbb{E}(W_s - sW_1)(W_t - tW_1) \\ &= \mathbb{E}W_sW_t - t\mathbb{E}W_sW_1 - s\mathbb{E}W_1W_t + st\mathbb{E}W_1^2 \\ &= s - ts - st + st \\ &= s(1-t) \end{aligned}$$

for $0 \le s \le t \le 1$.

Example 2.3.11 (Fractional Brownian motion)**.** As for the Brownian motion we assume that $I = [0, \infty)$. The Fractional Brownian motion was considered in 1941 by Kolmogorov in connection with turbulence and in 1968 by Mandelbrot and Van Ness as fractional Gaussian noise. Let $H \in (0,1)$. As fractional Brownian motion with Hurst index H (Hurst was an English hydrologist) we denote a Gaussian process $(W_t^H)_{t\geq 0}$ with $\mathbb{E}W_t^H = 0$ and with the covariance function

$$\mathbb{E}W_s^H W_t^H = \Gamma(s,t) := \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right).$$

For $H = 1/2$ we get $\Gamma(s,t) = \min\{s,t\}$, that means the Brownian motion from Example 2.3.9. The main problem consists in showing that Γ is positive semi-definite. To give the idea for this proof let $t_0 := 0$ and $a_0 := -\sum_{i=1}^n a_i$ so that $\sum_{i=0}^n a_i = 0$ and

$$\sum_{i,j=1}^n \Gamma(t_i,t_j) a_i a_j = -\frac{1}{2}\sum_{i,j=0}^n |t_i - t_j|^{2H} a_i a_j.$$

Take $\varepsilon > 0$. Since $\sum_{i,j=0}^n a_i a_j = 0$ we have using Taylor's theorem that

$$\begin{aligned}
\sum_{i,j=0}^n e^{-\varepsilon|t_i-t_j|^{2H}} a_i a_j &= \sum_{i,j=0}^n \left(e^{-\varepsilon|t_i-t_j|^{2H}} - 1\right) a_i a_j \\
&= -\varepsilon \sum_{i,j=0}^n |t_i - t_j|^{2H} a_i a_j + o(\varepsilon) \\
&= 2\varepsilon \sum_{i,j=1}^n \Gamma(t_i,t_j) a_i a_j + o(\varepsilon).
\end{aligned}$$

Hence it is sufficient to show that

$$\sum_{i,j=0}^n e^{-\varepsilon|t_i-t_j|^{2H}} a_i a_j \geq 0.$$

Case $H \in (0, \frac{1}{2}]$: In this case we show that

$$\varphi(t) = e^{-\varepsilon|t|^{2H}}$$

is a characteristic function and therefore by the theorem of Bochner and Khinchin positive semidefinite using the following result:

Case $H \in (0, \frac{1}{2}]$: In this case we first construct a random variable Z such that

$$\mathbb{E}e^{itZ} = e^{-\varepsilon|t|^{2H}}$$

by the following result:

Fact 2.3.12 (Polya). *Let $\varphi : \mathbb{R} \to [0, \infty)$ be*

(i) *continuous,*

(ii) *even (i.e. $\varphi(x) = \varphi(-x)$),*

(iii) *convex on $[0, \infty)$,*

(iv) *and assume that $\varphi(0) = 1$ and $\lim_{x\to\infty} \varphi(x) = 0$.*

Then there exists a probability measure μ on $\mathbb{R}$ such that $\widehat{\mu}(x) = \varphi(x)$.

General case $H \in (0, 1)$: For $0 < p = 2H < 2$ one gets the desired random variable Z with $\mathbb{E}e^{itZ} = e^{-\varepsilon|t|^{2H}}$ by

$$Z = \sum_{j=1}^{\infty} \Gamma_j^{-\frac{1}{p}} \eta_j$$

where $(\Gamma_j)_{j=1}^{\infty}$ are the jump times of a standard Poisson process and $\eta_1, \eta_2, ...$ are i.i.d. copies of a symmetric p-integrable random variable.

Since characteristic functions are positive semi-definite, we are done in both cases.

The random variables we obtained in both cases are p-stable with $p = 2H$, i.e. $\alpha Z + \beta Z'$ and $(|\alpha|^p + |\beta|^p)^{1/p} Z$ have the same distribution if $\alpha, \beta \in \mathbb{R}$ and Z' is an independent copy of Z.

Up to now we have constructed stochastic processes with certain finite-dimensional distributions. In the case of Gaussian processes this can be done through the covariance structure. Now we go the next step and provide the path-properties we would like to have. Here we use the fundamental

Theorem 2.3.13 (Kolmogorov's continuity theorem)**.** *Assume that* $X = (X_t)_{t\in[0,1]}$, $X_t : \Omega \to \mathbb{R}$, *is a family of random variables such that there are constants* $c, \varepsilon > 0$ *and* $p \in [1,\infty)$ *with*

$$\mathbb{E}|X_t - X_s|^p \le c|t-s|^{1+\varepsilon}.$$

Then there is a modification Y *of the process* X *such that*

$$\mathbb{E}\sup_{s\neq t}\left(\frac{|Y_t - Y_s|}{|t-s|^{\alpha}}\right)^p < \infty$$

for all $0 < \alpha < \frac{\varepsilon}{p}$ *and that all trajectories are continuous.*

Remark 2.3.14. In particular we get from the proof of the theorem two things:

(i) The function $f : \Omega \to [0,\infty]$ given by

$$f(\omega) := \sup_{s\neq t}\frac{|Y_t(\omega) - Y_s(\omega)|}{|t-s|^{\alpha}}$$

is a measurable function.

(ii) The function f is almost surely finite (otherwise $\mathbb{E}|f|^p$ would be infinite), so that there is a set Ω_0 of measure one such that for all $\omega \in \Omega_0$ there is a $c(\omega) > 0$ such that

$$|Y_t(\omega) - Y_s(\omega)| \le c(\omega)|t-s|^{\alpha}$$

for all $s, t \in [0,1]$ and $\omega \in \Omega_0$. In particular, the trajectories $t \to Y_t(\omega)$ are continuous for $\omega \in \Omega_0$.

Remark 2.3.15. *Proposition 2.3.13 can be formulated as an embedding theorem: The assumption reads as*

$$\|X_t - X_s\|_p \le c^{\frac{1}{p}}|t-s|^{\frac{1}{p}+\frac{\varepsilon}{p}} =: d|t-s|^{\frac{1}{p}+\eta}.$$

Letting

$$\|X\|_{C^{\frac{1}{p}+\eta}(L_p)} := \sup_{s\neq t}\frac{\|X_t - X_s\|_p}{|t-s|^{\frac{1}{p}+\eta}}$$

and

$$\|X\|_{L_p(C^\alpha)} := \left\| \sup_{s\neq t} \frac{|Y_t - Y_s|}{|t-s|^\alpha} \right\|_p ,$$

then we get the embedding

$$C^{\frac{1}{p}+\eta}(L_p) \hookrightarrow L_p(C^\alpha) \quad \textit{for} \quad 0 < \alpha < \eta.$$

Before we prove KOLMOGOROV's theorem we consider our fundamental example, the Brownian motion.

Proposition 2.3.16. *Let $W = (W_t)_{t\geq 0}$ be a Gaussian process with mean $m(t) = 0$ and covariance $\Gamma(s,t) = \mathbb{E}W_sW_t = \min\{s,t\}$. Then there is a modification $B = (B_t)_{t\geq 0}$ of $W = (W_t)_{t\geq 0}$ such that all trajectories are continuous and*

$$\mathbb{E}\left(\sup_{0\leq s<t\leq T} \frac{|B_t - B_s|}{|t-s|^\alpha} \right)^p < \infty$$

for all $0 < \alpha < \frac{1}{2}$, $0 < p < \infty$, and $T > 0$.

Proof. First we fix $T > 0$ and define

$$X_t := W_{tT}$$

for $t \in [0,1]$. Then, for $p \in (0,\infty)$,

$$\begin{aligned}
\mathbb{E}|X_t - X_s|^p &= \mathbb{E}|W_{tT} - W_{sT}|^p \\
&= \mathbb{E}|W_{(t-s)T}|^p \\
&= \frac{1}{\sqrt{2\pi(t-s)T}} \int_{\mathbb{R}} |\xi|^p e^{-\frac{\xi^2}{2(t-s)T}} d\xi \\
&= ((t-s)T)^{\frac{p}{2}} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |\xi|^p e^{-\frac{\xi^2}{2}} d\xi \\
&= \gamma_p (t-s)^{\frac{p}{2}} T^{\frac{p}{2}}
\end{aligned}$$

where we used that the covariance structure implies $W_b - W_a \sim N(0, b-a)$ for $0 \leq a < b < \infty$. Now fix $\alpha \in (0, 1/2)$ and $p \in (2,\infty)$ such that

$$\frac{1}{\frac{1}{2} - \alpha} < p < \infty \quad \text{and let } \varepsilon > 0 \text{ satisfy} \quad 0 < \alpha < \frac{\varepsilon}{p} = \frac{1}{2} - \frac{1}{p}.$$

Then
$$\mathbb{E}|X_t - X_s|^p \le \gamma_p T^{\frac{p}{2}}(t-s)^{1+\varepsilon}.$$
Proposition 2.3.13 implies the existence of a path-wise continuous modification $Y = Y(\alpha, p)$ of X such that

$$\left(\mathbb{E} \sup_{0\le s<t\le 1} \left| \frac{|Y_t(\alpha,p) - Y_s(\alpha,p)|}{|t-s|^\alpha}\right|^q\right)^{\frac{1}{q}} \le \left(\mathbb{E} \sup_{0\le s<t\le 1} \left| \frac{|Y_t(\alpha,p) - Y_s(\alpha,p)|}{|t-s|^\alpha}\right|^p\right)^{\frac{1}{p}} < \infty. \quad (2.1)$$

for all $q \in (0, p]$. Hence for each $0 < \alpha < 1/2$ and $0 < p < \infty$ we find a modification $Y(\alpha, p)$ such that (2.1) is satisfied. However, since $Y(\alpha_1, p_1)$ and $Y(\alpha_2, p_2)$ are continuous and modifications of each other, they are indistinguishable. Hence we can pick one process $Y = Y(p_0, \alpha_0)$ which satisfies (2.1) for all $0 < \alpha < 1/2$ and all $0 < p < \infty$. Coming back to our original time-scale we have found a continuous modification $(B_t^T)_{t\in[0,T]}$ of $(W_t)_{t\in[0,T]}$ such that
$$\mathbb{E} \sup_{0\le s<t\le T} \left(\frac{|B_s^T - B_t^T|}{|t-s|^\alpha}\right)^p < \infty$$
for all $0 < \alpha < 1/2$ and $0 < p < \infty$. We are close to the end, we only have to remove the remaining parameter T. For this purpose we let

$$\Omega_T := \left\{\omega \in \Omega : B_t^T(\omega) = W_t(\omega), t \in \mathbb{Q} \cap [0,T]\right\}$$

and $\widetilde{\Omega} := \bigcap_{N=1}^\infty \Omega_N$ so that $\mathbb{P}(\widetilde{\Omega}) = 1$ and

$$B_t^{N_1}(\omega) = B_t^{N_2}(\omega) \quad \text{for} \quad t \in \mathbb{Q} \cap [0, \min\{N_1, N_2\}]$$

and $\omega \in \widetilde{\Omega}$. Since $(B_t^{N_i})_{t\in[0,N_i]}$ are continuous processes we derive that

$$B_t^{N_1}(\omega) = B_t^{N_2}(\omega) \quad \text{for} \quad t \in [0, \min\{N_1, N_2\}]$$

whenever $\omega \in \widetilde{\Omega}$. Hence we have found one process $(B_t)_{t\ge 0}$ on $\widetilde{\Omega}$ and may set the process B zero on $\Omega \setminus \widetilde{\Omega}$. □

Proof of Theorem 2.3.13. (a) For $m = 1, 2, ...$ we let

$$D_m := \left\{0, \frac{1}{2^m}, ..., \frac{2^m}{2^m}\right\} \quad \text{and} \quad D := \bigcup_{m=1}^\infty D_m.$$

Moreover, we set

$$\Delta_m := \left\{(s,t) \in D_m \times D_m : |s-t| = 2^{-m}\right\} \quad \text{and} \quad K_m := \sup_{(s,t)\in\Delta_m} |X_t - X_s|.$$

Then $\mathrm{card}(\Delta_m) \leq 2\,2^m$ and

$$\begin{aligned}
\mathbb{E}K_m^p &= \mathbb{E} \sup_{(s,t)\in\Delta_m} |X_t - X_s|^p \\
&\leq \sum_{(s,t)\in\Delta_m} \mathbb{E}|X_t - X_s|^p \\
&\leq \mathrm{card}(\Delta_m) c \left(\frac{1}{2^m}\right)^{1+\varepsilon} \\
&\leq 22^m c 2^{-m} 2^{-m\varepsilon} \\
&= 2c2^{-m\varepsilon}.
\end{aligned}$$

(b) Let $s,t \in D$ and

$$\begin{aligned}
S_k &:= \max\{s_k \in D_k : s_k \leq s\} \in D_k \\
T_k &:= \max\{t_k \in D_k : t_k \leq t\} \in D_k
\end{aligned}$$

so that $S_k \uparrow s$, $T_k \uparrow t$, and $S_k = s$ and $T_k = t$ for $k \geq k_0$. For $|t-s| \leq 2^{-m}$ we get that

$$X_s - X_t = \sum_{i=m}^{\infty} \left(X_{S_{i+1}} - X_{S_i}\right) + X_{S_m} + \sum_{i=m}^{\infty} \left(X_{T_i} - X_{T_{i+1}}\right) - X_{T_m}$$

where we note that the sums are finite sums, that $|T_m - S_m| \in \{0, 2^{-m}\}$, $S_{i+1} - S_i \in \left\{0, 2^{-(i+1)}\right\}$, and that $T_{i+1} - T_i \in \left\{0, 2^{-(i+1)}\right\}$. Hence

$$|X_t - X_s| \leq K_m + 2\sum_{i=m}^{\infty} K_{i+1} \leq 2\sum_{i=m}^{\infty} K_i.$$

(c) Let

$$M_\alpha := \sup\left\{\frac{|X_t - X_s|}{|t-s|^\alpha} : s,t \in D, s \neq t\right\}.$$

Now we estimate M_α from above by

$$M_\alpha = \sup_{m=0,1,\ldots} \sup\left\{\frac{|X_t - X_s|}{|t-s|^\alpha} : s,t \in D, s \neq t, 2^{-m-1} \leq |t-s| \leq 2^{-m}\right\}$$

$$
\begin{aligned}
&\leq \sup_{m=0,1,\dots} 2^{(m+1)\alpha} \sup\left\{|X_t - X_s| : s,t \in D, s \neq t, |t-s| \leq 2^{-m}\right\} \\
&\leq 2 \sup_{m=0,1,\dots} 2^{(m+1)\alpha} \sum_{i=m}^{\infty} K_i \\
&\leq 2^{1+\alpha} \sum_{i=0}^{\infty} 2^{\alpha i} K_i,
\end{aligned}
$$

where we used step (b), and

$$
\begin{aligned}
\|M_\alpha\|_{L_p} &\leq 2^{1+\alpha} \sum_{i=0}^{\infty} 2^{\alpha i} \|K_i\|_{L_p} \\
&\leq 2^{1+\alpha} \sum_{i=0}^{\infty} 2^{\alpha i} (2c)^{\frac{1}{p}} 2^{-\frac{i\varepsilon}{p}} \\
&= 2^{1+\alpha} (2c)^{\frac{1}{p}} \sum_{i=0}^{\infty} 2^{i\left(\alpha - \frac{\varepsilon}{p}\right)} \\
&< \infty
\end{aligned}
$$

where we used step (a).

(d) Hence there is a set $\Omega_0 \subseteq \Omega$ with $\mathbb{P}(\Omega_0) = 1$ such that $t \to X_t(\omega)$ is uniformly continuous on D for $\omega \in \Omega_0$. We define

$$
Y_t(\omega) := \left\{ \begin{array}{rcl} X_t(\omega) & : & \omega \in \Omega_0, t \in D \\ \lim_{s \uparrow t, s \in D} X_s(\omega) & : & \omega \in \Omega_0, t \notin D \\ 0 & : & \omega \notin \Omega_0 \end{array} \right. .
$$

It remains to show that $\mathbb{P}(X_t = Y_t) = 1$. Because of our assumption we have that

$$
\|X_{t_n} - X_t\|_{L_p} \to 0 \quad \text{as} \quad t_n \uparrow t.
$$

Take $t_n \in D$ and find a subsequence $(n_k)_{k=1}^{\infty}$ such that

$$
\mathbb{P}(\lim_k X_{t_{n_k}} = X_t) = 1.
$$

Since $\mathbb{P}(\lim_k X_{t_{n_k}} = Y_t) = 1$ by construction, we are done. $\square$

2.4 Brownian motion

In this section we prove the existence of the Brownian motion introduced already earlier. To this end we start with a setting which does not use a filtration. This leads, in our first step, to a Brownian motion formally a bit different than presented in Definition 2.2.2.

Definition 2.4.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $B = (B_t)_{t \geq 0}$ be a stochastic process. The process B is called *standard Brownian motion* provided that

(i) $B_0 \equiv 0$,

(ii) for all $0 \leq s < t < \infty$ the random variable $B_t - B_s$ is independent from $(B_u)_{u \in [0,s]}$, which means that for all $0 \leq s_1 \leq \cdots \leq s_n \leq s$ and $A, A_1, ..., A_n \in \mathcal{B}(\mathbb{R})$ one has

$$\mathbb{P}(B_{s_1} \in A_1, ..., B_{s_n} \in A_n, B_t - B_s \in A) = \\ \mathbb{P}(B_{s_1} \in A_1, ..., B_{s_n} \in A_n)\mathbb{P}(B_t - B_s \in A),$$

(iii) for all $0 \leq s < t < \infty$ and for all $A \in \mathcal{B}(\mathbb{R})$ one has

$$\mathbb{P}(B_t - B_s \in A) = \frac{1}{\sqrt{2\pi(t-s)}} \int_A e^{-\frac{x^2}{2(t-s)}} d\lambda(x),$$

where λ denotes the Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,

(iv) all paths $t \to B_t(\omega)$ are continuous.

Proposition 2.4.2. *The standard Brownian motion exists.*

Proof. We take the process $B = (B_t)_{t \geq 0}$ from Proposition 2.3.16.

(i) Since $\mathbb{E} B_0 B_0 = 0$ so that $B_0 = 0$ a.s. we can set the whole process B on the null-set $\{B_0 \neq 0\}$ to zero and the conclusion of Proposition 2.3.16 is still satisfied.

(iv) follows directly from Proposition 2.3.16.

(iii) follows from $\mathbb{E}(B_t - B_s) = 0$,

$$\mathbb{E}(B_t - B_s)^2 = t - 2\min\{t, s\} + s = t - s,$$

and the fact that $B_t - B_s$ is a Gaussian random variable.

(ii) The random variables $B_t - B_s, B_{s_n} - B_{s_{n-1}}, ..., B_{s_2} - B_{s_1}, B_{s_1}$ are independent since they form a Gaussian random vector and any two of them are uncorrelated. Consequently

$$\begin{aligned}
&\mathbb{P}(B_{s_1} \in A_1, ..., B_{s_n} \in A_n, B_t - B_s \in A) \\
&= \mathbb{P}((B_{s_1}, ..., B_{s_n}) \in A_1 \times \cdots \times A_n, B_t - B_s \in A) \\
&= \mathbb{P}((B_{s_1}, B_{s_2} - B_{s_1}, ..., B_{s_n} - B_{s_{n-1}}) \in C, B_t - B_s \in A) \\
&= \mathbb{P}((B_{s_1}, B_{s_2} - B_{s_1}, ..., B_{s_n} - B_{s_{n-1}}) \in C)\mathbb{P}(B_t - B_s \in A) \\
&= \mathbb{P}(B_{s_1} \in A_1, ..., B_{s_n} \in A_n)\mathbb{P}(B_t - B_s \in A)
\end{aligned}$$

where

$$C := \{(y_1, ..., y_n) \in \mathbb{R}^n : y_1 \in A_1, y_1 + y_2 \in A_2, ..., y_1 + \cdots + y_n \in A_n\}.$$

□

Proposition 2.4.3. *Let $B = (B_t)_{t \geq 0}$ be a stochastic process such that all trajectories are continuous and such that $B_0 \equiv 0$. Then the following assertions are equivalent:*

(i) *The process B is a standard Brownian motion.*

(ii) *The process B is a Gaussian process with mean $m(t) \equiv 0$ and covariance $\Gamma(s,t) = \min\{s,t\}$.*

Proof. $(ii) \Rightarrow (i)$ From Proposition 2.3.16 (see Proposition 2.4.2) we get that there is a modification $\widetilde{B}$ of B which is a standard Brownian motion. Since both processes are continuous they are indistinguishable. Hence B is a standard Brownian motion as well.

$(i) \Rightarrow (ii)$ The mean is zero by definition and the covariance is obtained by, for $0 \leq s < t < \infty$,

$$\mathbb{E}B_sB_t = \mathbb{E}B_s(B_t - B_s) + \mathbb{E}B_s^2 = \mathbb{E}B_s\mathbb{E}(B_t - B_s) + s = s.$$

It remains to show that B is a Gaussian process. To this end we have to check whether for all $n = 1, 2, ...$ and $0 \leq t_1 < t_2 < \cdots < t_n < \infty$

$$(B_{t_1}, ..., B_{t_n}) : \Omega \to \mathbb{R}^n$$

is a Gaussian random vector. We know that $B_{t_n} - B_{t_{n-1}}, ..., B_{t_2} - B_{t_1}, B_{t_1}$ are independent Gaussian random variables so that $(B_{t_n} - B_{t_{n-1}}, ..., B_{t_2} - B_{t_1}, B_{t_1})$ is a Gaussian random vector. Using the linear transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ defined by $A(\xi_1, ..., \xi_n) := (\xi_1 + \cdots + \xi_n, \xi_1 + \cdots + \xi_{n-1}, ..., \xi_1)$ we obtain

$$A((B_{t_n} - B_{t_{n-1}}, ..., B_{t_2} - B_{t_1}, B_{t_1})) = (B_{t_n}, B_{t_{n-1}}, ..., , B_{t_1})$$

and are done since the linear image of a Gaussian random vector is a Gaussian random vector. □

Proposition 2.4.4. *The trajectories of the standard Brownian motion are Hölder continuous with exponent* $\alpha \in (0, 1/2)$, *i.e. the set*

$$A_{\alpha,T} := \left\{ \omega \in \Omega : \sup_{0 \le s < t \le T} \frac{|B_t(\omega) - B_s(\omega)|}{|t-s|^\alpha} < \infty \right\}$$

is measurable and of measure one for all $\alpha \in (0, 1/2)$ *and* $T > 0$.

Proof. Since the Brownian motion is continuous we get that $A_{\alpha,T} \in \mathcal{F}$. Moreover, $\mathbb{E}B_sB_t = \min\{s,t\}$ implies by Proposition 2.3.16 that there is a continuous modification $\widetilde{B} = (\widetilde{B}_t)_{t\ge 0}$ of B such that the corresponding set $\widetilde{A}_{\alpha,T}$ has measure one. However, B and $\widetilde{B}$ are indistinguishable, so that $\mathbb{P}(A_{\alpha,T}) = 1$ as well. □

For our later purpose we need a slight modification of the definition of the Brownian motion. We include a filtration as follows:

Definition 2.4.5 ($(\mathcal{F}_t)_{t\in I}$-Brownian motion)**.** Let $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\in I})$ be a stochastic basis. An *adapted* stochastic process $B = (B_t)_{t\in I}$, $B_t : \Omega \to \mathbb{R}$, is called (standard) $(\mathcal{F}_t)_{t\in I}$-Brownian motion provided that

(i) $B_0 \equiv 0$,

(ii) for all $0 \le s < t \in I$ the random variable $B_t - B_s$ is independent from $\mathcal{F}_s$ that means that

$$\mathbb{P}(C \cap \{B_t - B_s \in A\}) = \mathbb{P}(C)\mathbb{P}(B_t - B_s \in A)$$

for $C \in \mathcal{F}_s$ and $A \in \mathcal{B}(\mathbb{R})$,

(iii) for all $0 \leq s < t \in I$ one has $B_t - B_s \sim N(0, t-s)$,

(iv) for all $\omega \in \Omega$ the trajectories $t \mapsto B_t(\omega)$ are continuous.

Now we link this form of definition to the previous one.

Proposition 2.4.6. *Let $B = (B_t)_{t \geq 0}$ be a standard Brownian motion in the sense of Definition 2.4.1 and let $(\mathcal{F}_t^B)_{t \geq 0}$ be its natural filtration, i.e. $\mathcal{F}_t^B := \sigma(B_s : s \in [0,t])$. Then $(B_t)_{t \in I}$ is an $(\mathcal{F}_t^B)_{t \in I}$-Brownian motion.*

Proof. Comparing Definitions 2.4.1 and 2.4.5 we only need to check that Definition 2.4.1(ii) implies Definition 2.4.5(ii) which is left as an exercise. □

For technical reason we have to go one step further: we have to augment the natural filtration. First we recall the completion of a probability space.

Lemma 2.4.7. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and*

$$\mathcal{N} := \{A \subseteq \Omega : \text{ there exists a } B \in \mathcal{F} \text{ with } A \subseteq B \text{ and } \mathbb{P}(B) = 0\} \cup \{\emptyset\}.$$

(i) *Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Then $B \in \mathcal{G} \vee \mathcal{N}$ if and only if there is a $A \in \mathcal{G}$ such that $A \Delta B \in \mathcal{N}$.*

(ii) *The measure $\mathbb{P}$ can be extended to a measure $\widetilde{\mathbb{P}}$ on $\widetilde{\mathcal{F}} := \mathcal{F} \vee \mathcal{N}$ by $\widetilde{\mathbb{P}}(B) := \mathbb{P}(A)$ for $A \in \mathcal{F}$ such that $A \Delta B \in \mathcal{N}$.*

Definition 2.4.8. The probability space $(\Omega, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ is called *completion* of $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 2.4.9. Let $X = (X_t)_{t \in I}$, $X_t : \Omega \to \mathbb{R}$, be a stochastic process,

$$\mathcal{F}_\infty^X := \sigma(X_s : s \in I), \quad \text{and} \quad \mathcal{F}_t^X := \sigma(X_s : s \in [0,t])$$

for $t \in I$. Define

$$\mathcal{N} := \left\{A \subseteq \Omega : \text{ there exists a } B \in \mathcal{F}_\infty^X \text{ with } A \subseteq B \text{ and } \mathbb{P}(B) = 0\right\}$$

Then $(\mathcal{F}_t)_{t \in I}$ with $\mathcal{F}_t := \mathcal{F}_t^X \vee \mathcal{N}$ is called *augmentation* of $(\mathcal{F}_t^X)_{t \in I}$.

Proposition 2.4.10. *Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion, $(\mathcal{F}_t^B)_{t\geq 0}$ be its natural filtration and $(\mathcal{F}_t)_{t\in I}$ be the augmentation of $(\mathcal{F}_t^B)_{t\in I}$. Then*

(i) *the process $(B_t)_{t\in I}$ is an $(\mathcal{F}_t)_{t\in I}$-Brownian motion,*

(ii) *the filtration $(\mathcal{F}_t)_{t\in I}$ is right-continuous that means that*

$$\mathcal{F}_t = \bigcap_{s\in(t,S)} \mathcal{F}_s$$

with $0 \leq t < S := \infty$ if $I = [0,\infty)$ and $0 \leq t < S := T$ if $I = [0,T]$.

Proof. (i) We only have to check that $B_t - B_s$ is independent of $\mathcal{F}_s$ for $0 \leq s < t < \infty$. Assume $C \in \mathcal{F}_s$ and find an $\widetilde{C} \in \mathcal{F}_s^B$ such that $\mathbb{P}(C\Delta\widetilde{C}) = 0$ where we denote the extension $\widetilde{\mathbb{P}}$ of $\mathbb{P}$ again by $\mathbb{P}$. Taking $A \in \mathcal{B}(\mathbb{R})$ we get that

$$\begin{aligned}\mathbb{P}(\{B_t - B_s \in A\} \cap C) &= \mathbb{P}(\{B_t - B_s \in A\} \cap \widetilde{C}) \\ &= \mathbb{P}(B_t - B_s \in A)\mathbb{P}(\widetilde{C}) = \mathbb{P}(B_t - B_s \in A)\mathbb{P}(C).\end{aligned}$$

(ii) The right-hand side continuity of the filtration $(\mathcal{F}_t)_{t\in I}$ is not proved here. □

Definition 2.4.11. The stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\in I})$ satisfies the *usual conditions* provided that

(i) $(\Omega, \mathcal{F}, \mathbb{P})$ is complete,

(ii) $A \in \mathcal{F}_t$ for all $A \in \mathcal{F}$ with $\mathbb{P}(A) = 0$ and $t \in I$,

(iii) the filtration $(\mathcal{F}_t)_{t\in I}$ is right-continuous that means that

$$\mathcal{F}_t = \bigcap_{s\in(t,S)} \mathcal{F}_s$$

with $0 \leq t < S := \infty$ if $I = [0,\infty)$ and $0 \leq t < S := T$ if $I = [0,T]$.

2.5 Stopping and optional times

The heuristic Reflection Principle as motivation: Stopping times and optional times are random times that form an important tool in the theory of stochastic processes. To demonstrate their usefulness let us briefly consider the *Reflection Principle* for a standard Brownian motion $B = (B_t)_{t \geq 0}$. Given $b > 0$, we are interested in the distribution of

$$\tau_b := \inf\{t \geq 0 \,:\, B_t = b\}.$$

First we write

$$\mathbb{P}(\tau_b < t) = \mathbb{P}(\tau_b < t,\, B_t > b) + \mathbb{P}(\tau_b < t,\, B_t < b).$$

Then our heuristic *Reflection Principle* says that among the paths which reach b before time t, there will be as many above b as below b at time t:

$$\mathbb{P}(\tau_b < t,\, B_t < b) = \mathbb{P}(\tau_b < t,\, B_t > b). \tag{2.2}$$

On the other hand

$$\mathbb{P}(\tau_b < t, B_t > b) = \mathbb{P}(B_t > b)$$

which implies

$$\begin{aligned} \mathbb{P}(\tau_b < t) &= \mathbb{P}(\tau_b < t,\, B_t > b) + \mathbb{P}(\tau_b < t,\, B_t < b) \\ &= 2\mathbb{P}(\tau_b < t,\, B_t > b) \\ &= 2\mathbb{P}(B_t > b). \end{aligned}$$

From this one can deduce at least two things:

- The Brownian motion reaches with probability one any level because

$$\begin{aligned} \mathbb{P}(\tau_b < \infty) = \lim_{t\to\infty} \mathbb{P}(\tau_b < t) = 2\lim_{t\to\infty} \mathbb{P}(B_t > b) \\ = 2\lim_{t\to\infty} \mathbb{P}\left(B_1 > \frac{b}{\sqrt{t}}\right) = 1. \end{aligned}$$

- One can deduce the distribution of the running maximum of the Brownian motion $M_t(\omega) := \sup_{s\in[0,t]} B_s(\omega)$ because

$$\{M_t \geq b\} = \{\tau_b \leq t\} \quad \text{so that} \quad \mathbb{P}(M_t \geq b) = 2\mathbb{P}(B_t > b).$$

To justify (2.2) would require a considerable amount of work. Here we only introduce the concepts around the random time $\tau_b : \Omega \to [0, \infty]$ like stopping times and optional times, and their relations to each other.

The precise definitions: In the following we assume in the case $I = [0,T]$ **that $\mathcal{F} = \mathcal{F}_T$.** Moreover, to treat both cases, $I = [0,\infty)$ and $I = [0,T]$, at the same time, we let

$$\bar{I} := \begin{cases} [0,\infty] & : I = [0,\infty) \\ [0,T] & : I = [0,T] \end{cases}.$$

Definition 2.5.1. Assume a measurable space $(\Omega, \mathcal{F})$ equipped with a filtration $(\mathcal{F}_t)_{t\in I}$.

(i) The map $\tau : \Omega \to \bar{I}$ is called *stopping time* with respect to the filtration $(\mathcal{F}_t)_{t\in I}$ provided that
$$\{\tau \leqq t\} \in \mathcal{F}_t$$
for all $t \in I$. Moreover,
$$\mathcal{F}_\tau := \{A \in \mathcal{F} \,:\, A \cap \{\tau \leqq t\} \in \mathcal{F}_t \text{ for all } t \in I\}.$$

(ii) The map $\tau : \Omega \to \bar{I}$ is called *optional time* with respect to the filtration $(\mathcal{F}_t)_{t\in I}$ provided that
$$\{\tau < t\} \in \mathcal{F}_t$$
for all $t \in I$. Moreover,
$$\mathcal{F}_{\tau^+} := \{A \in \mathcal{F} \,:\, A \cap \{\tau < t\} \in \mathcal{F}_t \text{ for all } t \in I\}.$$

It should be noted that the constant time $\tau \equiv T$ is an optional time, so that in the case $I = [0,T]$ we get (with our convention that $\mathcal{F} = \mathcal{F}_T$) $\mathcal{F}_{T^+} = \mathcal{F}_T$.

One might say that $\mathcal{F}_\tau$ contains those events that can be decided until time τ and that $\mathcal{F}_{\tau^+}$ contains those events that can be decided right after the random time τ occurs. Let us give a simple example of an optional time that is not a stopping time:

Example 2.5.2. Let $I = [0,\infty)$, $\Omega := \{1,2,3\}$, $\mathcal{F} = 2^\Omega$, $\mathbb{P}(\{k\}) = 1/3$, and $X_t(\omega) := (t-\omega)^+$ for $t \geq 0$. Then the natural filtration computes as

$$\mathcal{F}_t^X = \begin{cases} \{\emptyset, \Omega\} & : \ t \in [0,1] \\ \{\emptyset, \Omega, \{1\}, \{2,3\}\} & : \ t \in (1,2] \\ 2^\Omega & : \ t > 2 \end{cases}.$$

Let $\tau(\omega) := \inf\{t > 0 : X_t(\omega) > 0\}$ so that $\tau(\omega) = \omega$. The time τ is not a stopping time since

$$\{\tau \leq 1\} = \{1\} \notin \mathcal{F}_1.$$

But it is an optional time since $\{\tau < t\} \in \mathcal{F}_t$ for all $t \geq 0$.

Proposition 2.5.3. *Let $\tau : \Omega \to \bar{I}$ be a stopping time. Then*

(i) *the system of sets $\mathcal{F}_\tau$ is a σ-algebra,*

(ii) *one has $\{\tau \leqq s\} \in \mathcal{F}_\tau$ for $s \in I$ so that $\tau : \Omega \to \bar{I}$ is an extended $\mathcal{F}_\tau$-measurable random variable.*

Proof. (i) Since $\emptyset \cap \{\tau \leq t\} = \emptyset \in \mathcal{F}_t$ we have $\emptyset \in \mathcal{F}_\tau$. Assume that $B_1, B_2, ... \in \mathcal{F}_\tau$. Then

$$\left(\bigcup_{n=1}^{\infty} B_n\right) \cap \{\tau \leqq t\} = \bigcup_{n=1}^{\infty} (B_n \cap \{\tau \leqq t\}) \in \mathcal{F}_t.$$

Finally, for $B \in \mathcal{F}_\tau$ we get that

$$B^c \cap \{\tau \leqq t\} = \{\tau \leqq t\} \setminus (B \cap \{\tau \leqq t\}) \in \mathcal{F}_t.$$

(ii) For $s, t \in I$ we get that

$$\{\tau \leqq s\} \cap \{\tau \leqq t\} = \{\tau \leqq \min\{s, t\}\} \in \mathcal{F}_{\min\{s,t\}} \subseteq \mathcal{F}_t.$$

We conclude the proof by remarking that the system $(-\infty, t]$, $t \in \mathbb{R}$, generates the Borel σ-algebra and that (for example) $\{\tau \leq t\} = \emptyset \in \mathcal{F}_\tau$ for $t < 0$. □

The pendant for the optional times reads as

Proposition 2.5.4. *Let $\tau : \Omega \to \bar{I}$ be an optional time. Then*

(i) *the system of sets $\mathcal{F}_{\tau^+}$ is a σ-algebra,*

(ii) *one has $\{\tau < s\} \in \mathcal{F}_{\tau^+}$ for all $s \in I$ so that τ is an extended $\mathcal{F}_{\tau^+}$-measurable random variable.*

The proof is an exercise. Optional times are stopping times with respect to the filtration $(\mathcal{F}_{t^+})_{t \in I}$:

Proposition 2.5.5. *Let $(\Omega, \mathcal{F})$ be a measurable space equipped with a filtration $(\mathcal{F}_t)_{t\in I}$ and $\tau : \Omega \to \bar{I}$.*

(i) *Then τ is an optional time if and only if, for all $t \in I$,*

$$\{\tau \le t\} \in \mathcal{F}_{t^+} := \bigcap_{s\in(t,S)} \mathcal{F}_s$$

with $0 \le t < S := \infty$ if $I = [0, \infty)$ and $0 \le t < S := T$ if $I = [0, T]$, and with the convention that $\mathcal{F}_{T^+} = \mathcal{F}_T$ if $I = [0, T]$.

(ii) *If τ is an optional time, then $A \in \mathcal{F}_{\tau^+}$ if and only if $A \cap \{\tau \le t\} \in \mathcal{F}_{t^+}$ for $t \in I$.*

The proof is an exercise. The general relation between stopping and optional times is as follows:

Proposition 2.5.6. (i) *Every stopping time τ is an optional time and $\mathcal{F}_\tau \subseteq \mathcal{F}_{\tau^+}$.*

(ii) *If the filtration is right-continuous, then any optional time τ is a stopping time and $\mathcal{F}_\tau = \mathcal{F}_{\tau^+}$.*

Proof. (i) This follows from

$$\{\tau < t\} = \bigcup_{n=1}^{\infty} \left\{\tau \le t - \frac{1}{n}\right\}$$

and

$$\left\{\tau \le t - \frac{1}{n}\right\} \in \mathcal{F}_{\max\left\{0, t-\frac{1}{n}\right\}} \subseteq \mathcal{F}_t.$$

Hence τ is an optional time. Let $A \in \mathcal{F}_\tau$ so that by definition $A \cap \{\tau \le t\} \in \mathcal{F}_t$ for all $t \in I$. Hence

$$A \cap \{\tau < t\} = \bigcup_{n=1}^{\infty} A \cap \left\{\tau \le t - \frac{1}{n}\right\} \in \mathcal{F}_t$$

because

$$A \cap \left\{\tau \le t - \frac{1}{n}\right\} \in \mathcal{F}_{\max\left\{0, t-\frac{1}{n}\right\}} \subseteq \mathcal{F}_t.$$

(ii) For the second assertion we have to check that $\mathcal{F}_{\tau+} \subseteq \mathcal{F}_\tau$. Given $A \in \mathcal{F}$, from Proposition 2.5.5 we know that $A \in \mathcal{F}_{\tau+}$ if and only if $A \cap \{\tau \le t\} \in \mathcal{F}_{t+}$. But, by assumption $\mathcal{F}_t = \mathcal{F}_{t^+}$, so that the statement follows. $\square$

Now we consider two basic examples.

Example 2.5.7. Let $I = [0, \infty)$, $X = (X_t)_{t \ge 0}$ be continuous and adapted, $D \subseteq \mathbb{R}$ be non-empty, and define the first hitting time

$$\tau_D := \inf\{t \geqq 0 \,:\, X_t \in D\}$$

with the convention that $\inf \emptyset := \infty$.

(i) If D is open, then τ_D is an optional time.

(ii) If D is closed, then τ_D is a stopping time.

Proof. We only show (ii), part (i) is an exercise. Given $t \ge 0$ we have to show that $\{\tau_D \leqq t\} \in \mathcal{F}_t$. The condition $\tau_D(\omega) \le t$ implies the existence of a sequence $t_1 \ge t_2 \ge t_3 \ge \cdots \ge 0$ such that $X_{t_n}(\omega) \in D$ and $s := \lim_n t_n \le t$. By the continuity of X we have $X_s(\omega) = \lim_n X_{t_n}(\omega)$ and by the closedness of D that $X_s(\omega) \in D$. Hence

$$\begin{aligned} \{\tau_D \leqq t\} &= \{\omega \in \Omega : \text{ there exists an } s \in [0,t] \text{ such that } X_s(\omega) \in D\} \\ &= \{\omega \in \Omega : \inf_{s \in \mathbb{Q},\, s \le t} d(X_s(\omega), D) = 0\} \in \mathcal{F}_t \end{aligned}$$

where $d(x, D) := \inf\{|x - y| : y \in D\}$ is a continuous function, note that

$$|d(x, D) - d(y, D)| \le |x - y|,$$

which implies that $\omega \mapsto d(X_s(\omega), D)$ is $\mathcal{F}_s$-measurable. $\square$

2.6 A short excursion to Markov processes

To give a rigorous justification of the reflection principle one would need to introduce the strong Markov property. In this course we restrict ourselves to the introduction of the basic concept. Because of the continuity properties of the filtration (right-hand side continuous or not) we first work with the optional times and not with the, probably more intuitive, stopping times.

Before we begin, we make a general observation: Given a measurable process $X = (X_t)_{t\geq 0}$ and an optional time τ and a fixed $S \in [0,\infty)$, the map $\omega \mapsto X_{\tau(\omega)\wedge S}(\omega)$ is measurable because it can be written as $X_{\tau(\omega)\wedge S}(\omega) = X \circ \Phi^S(\omega)$ where $\Phi^S : \Omega \to [0,S] \times \Omega$ is given by $\Phi^S(\omega) := (\tau(\omega) \wedge S, \omega)$ and $X^S : [0,T] \times \Omega \to \mathbb{R}$ with $X(t,\omega) = X_t(\omega)$. By $(X_\tau)(\omega) = \lim_{n\to\infty}(X_{\tau\wedge n})(\omega)$ it follows that that X_τ is measurable whenever $\tau(\omega) < \infty$ for all $\omega \in \Omega$.

Definition 2.6.1. Let $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\geq 0})$ be a stochastic basis and $X = (X_t)_{t\geq 0}$, $X_t : \Omega \to \mathbb{R}$, be a stochastic process.

(i) The process X is a *Markov process* provided that X is adapted and for all $s, t \geq 0$ and $B \in \mathcal{B}(\mathbb{R})$ one has that

$$\mathbb{P}(X_{s+t} \in B|\mathcal{F}_s) = \mathbb{P}(X_{s+t} \in B|\sigma(X_s)) \ a.s.$$

(ii) The process X is a *strong Markov process* provided that X is progressively measurable and for all $t \geq 0$, optional times $\tau : \Omega \to [0,\infty]$, and $B \in \mathcal{B}(\mathbb{R})$ one has that, a.s.,

$$\mathbb{P}(\{X_{\tau+t} \in B\} \cap \{\tau < \infty\} |\mathcal{F}_{\tau+}) = \mathbb{P}(\{X_{\tau+t} \in B\} \cap \{\tau < \infty\} |\sigma(X_\tau)),$$

where $\sigma(X_\tau) := \sigma(\tau^{-1}(\infty), \{X_\tau^{-1}(B) \cap \{\tau < \infty\} : B \in \mathcal{B}(\mathbb{R})\})$.

To make the definition rigorous we need

Lemma 2.6.2. *Let X be progressively measurable and $\tau : \Omega \to [0,\infty]$ be an optional time, then*

$$X_\tau^{-1}(B) \cap \{\tau < \infty\} \in \mathcal{F}_{\tau+}$$

for all $B \in \mathcal{B}(\mathbb{R})$.

Proof. Note that $X_\tau^{-1}(B)\cap\{\tau < \infty\}\cap\{\tau < t\} = X_{\tau\wedge t}^{-1}(B)\cap\{\tau < t\} \in \mathcal{F}_t$. $\square$

We give an easy example for a Markov process which is not a strong Markov process.

Example 2.6.3. We consider $\Omega := \{-1, 1\}$, $\mathcal{F} := 2^\Omega$, and $\mathbb{P}(\{-1\}) = \mathbb{P}(\{1\}) = 1/2$. As stochastic process we take $X_t(1) := \max\{t-1, 0\}$ and $X_t(-1) := \min\{-t+1, 0\}$. The filtration is the natural filtration

$\mathcal{F}_t := \sigma(X_s : s \in [0,t])$. Taking the optional time $\tau \equiv 1$, one checks that $\sigma(X_\tau) = \{\emptyset, \Omega\}$ and $\mathcal{F}_{\tau+} = 2^\Omega$. But this gives that

$$\mathbb{P}(X_2 \in B|\mathcal{F}_{\tau+}) = \mathbb{P}(X_2 \in B|\sigma(X_\tau)) \; a.s.$$

cannot be true for all $B \in \mathcal{B}(\mathbb{R})$.

For us, the main (positive) example is

Proposition 2.6.4. *Assume a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\geq 0})$ and an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion $B = (B_t)_{t\geq 0}$ like in Definition 2.4.5. Then B is a strong Markov process.*

Finally, without proof we state

Proposition 2.6.5. *Assume a process $X = (X_t)_{t\geq 0}$ which is a strong Markov process with respect to its natural filtration $(\mathcal{F}_t^X)_{t\geq 0}$ with $\mathcal{F}_t^X := \sigma(X_u : u \in [0,t])$. Then the augmented filtration is right continuous.*

Chapter 3

Stochastic integration

Given a Brownian motion $B = (B_t)_{t\geq 0}$, we would like to define

$$\int_0^T L_t dB_t$$

for a large class of stochastic processes $L = (L_t)_{t\geq 0}$. A first approach would be to write

$$\int_0^T L_t dB_t = \int_0^T L_t \frac{dB_t}{dt} dt.$$

However this is not possible (at least in this naive form) because of the following

Proposition 3.0.1 (Paley, Wiener-Zygmund). *Given a standard Brownian motion in the sense of Definition 2.4.1, then the set*

$$\{\omega \in \Omega : t \to B_t(\omega) \text{ is nowhere differentiable } \}$$

contains a set of measure one.

So we have to proceed differently. We will first define the stochastic integral for simple processes and extend then the definition to an appropriate class of processes. Throughout the whole section we assume that the *usual conditions* hold:

- the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is complete,

- the filtration $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous that means $\bigcap_{\varepsilon>0}\mathcal{F}_{t+\varepsilon}=\mathcal{F}_t$ for all $t\geq 0$,
- all null-sets of $\mathcal{F}$ are contained in $\mathcal{F}_0$,
- the process $B=(B_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion.

3.1 The Itô integral

3.1.1 Martingales, Doob's maximal inequality

As a preparation for the definition of Itô's integral and its investigation we start with martingales: Recall that $M=(M_t)_{t\geq 0}$ is a **martingale** w.r.t. the filtration $(\mathcal{F}_t)_{t\geq 0}$ if

(i) M_t is $\mathcal{F}_t$-measurable for all $t\geq 0$ (M is adapted)

(ii) $\mathbb{E}|M_t|<\infty$ for all $t\geq 0$ (M is integrable)

(iii) $\mathbb{E}[M_t|\mathcal{F}_s]=M_s$ a.s. for all $0\leq s\leq t<\infty$ (martingale property).

Definition 3.1.1. *A martingale $M=(M_t)_{t\geq 0}$ belongs to $\mathcal{M}_2^c$ provided that*

(i) $\mathbb{E}|M_t|^2<\infty$ *for all* $t\geq 0$,

(ii) *the trajectories $t\mapsto M_t(\omega)$ are continuous for all $\omega\in\Omega$.*

In case that $M_0\equiv 0$ we write $M\in\mathcal{M}_2^{c,0}$.

Proposition 3.1.2 (Doob's maximal inequalities)**.** *Let $M=(M_t)_{t\geq 0}$ be a right-continuous martingale or a right-continuous positive sub-martingale and let $M_t^*:=\sup_{s\in[0,t]}|M_s|$. Then one has, for $\lambda,t\geq 0$ and $p\in(1,\infty)$, that*

$$\lambda\mathbb{P}(M_t^*\geqq\lambda)\leq\mathbb{E}|M_t|,$$

and

$$\mathbb{E}(M_t^*)^p\leqq\left(\frac{p}{p-1}\right)^p\mathbb{E}|M_t|^p.$$

For the proof we use the discrete time version proved in [7]:

Lemma 3.1.3. *Let $M = (M_n)_{n=0}^N$ be martingale or positive sub-martingale and let $M_n^* := \sup_{k=0,...,n} |M_k|$. Then one has, for $\lambda \geq 0$, $n = 0, ..., N$, and $p \in (1, \infty)$, that*

$$\lambda \mathbb{P}(M_n^* \geqq \lambda) \leq \mathbb{E}|M_n|,$$

and

$$\mathbb{E}(M_n^*)^p \leqq \left(\frac{p}{p-1}\right)^p \mathbb{E}|M_n|^p.$$

Proof of Proposition 3.1.2. First we remark that $M_t^* : \Omega \to \mathbb{R}$ is measurable, since

$$M_t^* = \sup_{s \in [0,t] \cap \mathbb{Q}} (|M_s| \vee |M_t|)$$

which is a countable supremum of measurable functions. From Lemma 3.1.3 it follows that, for

$$M_t^{n,*} := \sup_{\substack{s = \frac{k}{2^n} t \\ k = 0,...,2^n}} |M_s|,$$

one has that

$$\lambda \mathbb{P}(M_t^{n,*} \geq \lambda) \leq \mathbb{E}|M_t| \quad \text{and} \quad \mathbb{E}(M_t^{n,*})^p \leq \left(\tfrac{p}{p-1}\right)^p \mathbb{E}|M_t|^p.$$

Since $M_t^{n,*} \uparrow M_t^*$ a.s. as $n \to \infty$ we are done. □

The next lemma will turn out to be useful below to show the so called Itô isometry.

Lemma 3.1.4. *Let $M = (M_k)_{k=0}^N$ be a martingale with respect to $(\mathcal{G}_k)_{k=0}^N$ such that $\mathbb{E}M_k^2 < \infty$ for $k = 0, ..., N$. Then, a.s.,*

$$\mathbb{E}\left((M_N - M_n)^2 | \mathcal{G}_n\right) = \mathbb{E}\left(\sum_{l=n+1}^N (dM_l)^2 | \mathcal{G}_n\right)$$

for $0 \leq n < N$ where $dM_l := M_l - M_{l-1}$.

Proof. We have that

$$\mathbb{E}(M_N - M_n)^2 | \mathcal{G}_n) = \sum_{i,j=n+1}^{N} \mathbb{E}(dM_i \, dM_j | \mathcal{G}_n).$$

We are done if we can show that

$$\mathbb{E}(dM_i \, dM_j | \mathcal{F}_n) = 0 \;\; a.s.$$

for $n < i < j \leq N$. But this follows from, a.s.,

$$\mathbb{E}(dM_i \, dM_j | \mathcal{G}_n) = \mathbb{E}\left(\mathbb{E}(dM_i \, dM_j | \mathcal{G}_i) | \mathcal{G}_n\right) = \mathbb{E}\left(dM_i \, \mathbb{E}(dM_j | \mathcal{G}_i) | \mathcal{G}_n\right) = 0.$$

□

3.1.2 Step 1: The Itô integral for simple processes

Let $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\geq 0})$ be a stochastic basis satisfying the *usual conditions* (see Definition 2.4.11) and $B = (B_t)_{t\geq 0}$ an $(\mathcal{F}_t)_{t\geq 0}$ -Brownian motion.

First we define simple processes we are able to integrate:

Definition 3.1.5 (the space $\mathcal{L}_0$)**.** A stochastic process $H = (H_t)_{t\geq 0}$ is called **simple** provided that there exists

(i) a sequence $0 = t_0 < t_1 < t_2 < \cdots$ with $\lim_n t_n = \infty$,

(ii) $\mathcal{F}_{t_i}$-measurable random variables $v_i : \Omega \to \mathbb{R}$, $i = 0, 1, ...$, with $c := \sup_{i,\omega} |v_i(\omega)| < \infty$,

such that

$$H_t(\omega) = \sum_{i=1}^{\infty} \chi_{(t_{i-1}, t_i]}(t) v_{i-1}(\omega).$$

The class of these processes is denoted by $\mathcal{L}_0$.

For simple processes we define the stochastic integral as follows

Definition 3.1.6 (Stochastic integral for $\mathcal{L}_0$-integrands)**.** *For $H \in \mathcal{L}_0$ and $t \geq 0$ we define the* **stochastic integral**

$$I_t(H)(\omega) := \sum_{i=1}^{\infty} v_{i-1}(\omega)\left(B_{t_i \wedge t}(\omega) - B_{t_{i-1} \wedge t}(\omega)\right),$$

where $a \wedge b := min\{a, b\}$.

Proposition 3.1.7. *For $H \in \mathcal{L}_0$ one has that $I(H) := (I_t(H))_{t \geq 0} \in \mathcal{M}_2^{c,0}$.*

Proof. By definition we have that $I_0(H) \equiv 0$ and that the process

$$t \mapsto I_t(H)(\omega)$$

is continuous for all $\omega \in \Omega$. We show next that $I(H)$ is a square integrable martingale: Since

$$v_{i-1}(B_{t_i \wedge t} - B_{t_{i-1} \wedge t}) = \begin{cases} 0 & : \quad t \leq t_{i-1} \\ v_{i-1}(B_{t_i \wedge t} - B_{t_{i-1}}) & : \quad t > t_{i-1} \end{cases}$$

we get that $I_t(H)$ is $\mathcal{F}_t$-measurable. Now we observe that

$$\mathbb{E}|v_{i-1}(B_b - B_{t_{i-1}})|^2 \leq c^2 \mathbb{E}|B_b - B_{t_{i-1}}|^2 = c^2(b - t_{i-1})$$

for $t_{i-1} \leq b < \infty$ so that

$$\begin{aligned} \left(\mathbb{E}|I_t(H)|^2\right)^{\frac{1}{2}} &\leq \left(\mathbb{E}\left|\sum_{i=1}^{n_0} v_{i-1}(B_{t_i \wedge t} - B_{t_{i-1} \wedge t})\right|^2\right)^{\frac{1}{2}} \\ &\leq \sum_{i=1}^{n_0} \left(\mathbb{E}|v_{i-1}(B_{t_i \wedge t} - B_{t_{i-1} \wedge t})|^2\right)^{\frac{1}{2}} < \infty \end{aligned}$$

whenever $t_{n_0 - 1} < t \leq t_{n_0}$. It remains to show the martingale property

$$\mathbb{E}(I_t(H)|\mathcal{F}_s) = I_s(H) \quad a.s.$$

for $0 \leq s \leq t < \infty$. By introducing new time knots we can assume without loss of generality that $s = t_n$ and $t = t_N$. Then

$$I_t(L) = \sum_{i=1}^{N} v_{i-1}(B_{t_i} - B_{t_{i-1}})$$

and

$$\begin{aligned}
& \mathbb{E}(I_t(H)|\mathcal{F}_s) \\
= \;& \mathbb{E}\Big(\sum_{i=1}^N v_{i-1}(B_{t_i}-B_{t_{i-1}})|\mathcal{F}_{t_n}\Big) \\
= \;& \mathbb{E}\Big(\sum_{i=1}^n v_{i-1}(B_{t_i}-B_{t_{i-1}})|\mathcal{F}_{t_n}\Big) + \mathbb{E}\Big(\sum_{i=n+1}^N v_{i-1}(B_{t_i}-B_{t_{i-1}})|\mathcal{F}_{t_n}\Big) \\
= \;& \sum_{i=1}^n v_{i-1}(B_{t_i}-B_{t_{i-1}}) \quad a.s.
\end{aligned}$$

Indeed, for $1 \le i \le n$ the expression $v_{i-1}(B_{t_i}-B_{t_{i-1}})$ is $\mathcal{F}_{t_n}$ measurable, while for $n+1 \le i \le N$ the tower property implies that

$$\begin{aligned}
\mathbb{E}\Big(v_{i-1}(B_{t_i}-B_{t_{i-1}})|\mathcal{F}_{t_n}\Big) &= \mathbb{E}\Big(\mathbb{E}\Big(v_{i-1}(B_{t_i}-B_{t_{i-1}})|\mathcal{F}_{t_{i-1}}\Big)|\mathcal{F}_{t_n}\Big) \\
&= \mathbb{E}\Big(v_{i-1}\mathbb{E}\Big((B_{t_i}-B_{t_{i-1}})|\mathcal{F}_{t_{i-1}}\Big)|\mathcal{F}_{t_n}\Big) \\
&= 0 \quad a.s.
\end{aligned}$$

Clearly, $\sum_{i=1}^n v_{i-1}(B_{t_i}-B_{t_{i-1}}) = I_s(H)$.

□

Now we can prove our first Itô-isometry:

Proposition 3.1.8. *For $H \in \mathcal{L}_0$ and $0 \le s \le t < \infty$ one has that*

$$\mathbb{E}\left([I_t(H)-I_s(H)]^2|\mathcal{F}_s\right) = \mathbb{E}\left(\int_s^t H_u^2 du|\mathcal{F}_s\right) a.s.$$

Proof. As above, by introducing new time knots we can assume without loss of generality that $s = t_n$ and $t = t_N$. Let

$$M_k := I_{t_k}(H) \quad \text{and} \quad \mathcal{G}_k := \mathcal{F}_{t_k}$$

so that $(M_k)_{k=0}^N$ is a martingale with respect to $(\mathcal{G}_k)_{k=0}^N$. Hence, a.s.,

$$\mathbb{E}\left([I_t(H)-I_s(H)]^2|\mathcal{F}_s\right) = \mathbb{E}\left([M_N-M_n]^2|\mathcal{G}_n\right)$$

$$
\begin{aligned}
&= \mathbb{E}\left(\sum_{l=n+1}^{N}(dM_l)^2|\mathcal{G}_n\right)\\
&= \mathbb{E}\left(\sum_{l=n+1}^{N}v_{l-1}^2(B_{t_l}-B_{t_{l-1}})^2|\mathcal{F}_{t_n}\right)\\
&= \sum_{l=n+1}^{N}\mathbb{E}\left(\mathbb{E}(v_{l-1}^2(B_{t_l}-B_{t_{l-1}})^2|\mathcal{F}_{t_{l-1}})|\mathcal{F}_{t_n}\right)\\
&= \sum_{l=n+1}^{N}\mathbb{E}\left(v_{l-1}^2\mathbb{E}((B_{t_l}-B_{t_{l-1}})^2|\mathcal{F}_{t_{l-1}})|\mathcal{F}_{t_n}\right)\\
&= \sum_{l=n+1}^{N}\mathbb{E}\left(v_{l-1}^2\mathbb{E}(B_{t_l}-B_{t_{l-1}})^2|\mathcal{F}_{t_n}\right)\\
&= \sum_{l=n+1}^{N}\mathbb{E}\left(v_{l-1}^2(t_l-t_{l-1})|\mathcal{F}_{t_n}\right)\\
&= \mathbb{E}\left(\int_{t_n}^{t_N}H_u^2du|\mathcal{F}_{t_n}\right)\\
&= \mathbb{E}\left(\int_{s}^{t}H_u^2du|\mathcal{F}_{s}\right).
\end{aligned}
$$

□

Proposition 3.1.9. *For $H, K \in \mathcal{L}_0$ and $\alpha, \beta \in \mathbb{R}$ one has that*

$$I_t(\alpha H + \beta K) = \alpha I_t(H) + \beta I_t(K).$$

3.1.3 Step 2: The Itô integral extended from $\mathcal{L}_0$ to $\mathcal{L}_2$

We want to extend the map

$$I : \mathcal{L}_0 \longrightarrow \mathcal{M}_2^{c,0}$$

(which was defined ω-wise) to a larger class of integrands.

Definition 3.1.10. Let $\mathcal{L}_2$[1] be the set of all progressively measurable processes $H = (H_t)_{t\geq 0}$, $H_t : \Omega \to \mathbb{R}$, such that

$$\|H\|_{\mathcal{L}_2,T} := \left(\mathbb{E}\int_0^T H_t^2 dt\right)^{\frac{1}{2}} < \infty \quad \text{for all} \quad T > 0.$$

Moreover, we let

$$\mathtt{d}(K,H) := \sum_{n=1}^{\infty} 2^{-n} \min\{1, \|K-H\|_{\mathcal{L}_2,n}\} \quad \text{for} \quad K,H \in \mathcal{L}_2.$$

Proposition 3.1.11. *One has that $\mathcal{L}_0 \subseteq \mathcal{L}_2$. Moreover the inclusion is dense, that means for all $H \in \mathcal{L}_2$ there is a sequence $(H^{(n)})_{n=1}^{\infty} \subseteq \mathcal{L}_0$ such that*

$$\lim_n \mathtt{d}(H^{(n)}, H) = 0.$$

Proof. **step a:** Assume that $\sup_{t,\omega} |H_t(\omega)| < \infty$ and that $t \to H_t(\omega)$ is continuous for all $\omega \in \Omega$. For $T > 0$ we define

$$H_t^{(n)}(\omega) := \sum_{k=0}^{2^n-1} H_{\frac{kT}{2^n}}(\omega)\chi_{\left(\frac{kT}{2^n},\frac{(k+1)T}{2^n}\right]}(t).$$

Then $\lim_{n\to\infty} H_t^{(n)}(\omega) = H_t(\omega)$ for all $t \in (0,T]$ and, by dominated convergence,

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

step b: Let H be progressively measurable with $\sup_{t,\omega} |H_t(\omega)| < \infty$ and let $T > 0$. For $t \geq 0$ and $m = 1, 2, ...$ define

$$\widetilde{H}_t^{(m)}(\omega) := m\int_{\left(t-\frac{1}{m}\right)^+}^t H_s(\omega) ds.$$

Hence $(\widetilde{H}^{(m)})_{t\geq 0}$ is continuous and adapted (we can write $\widetilde{H}_t^{(m)}(\omega) = m[K_t(\omega) - K_{(t-\frac{1}{m})\vee 0}(\omega)]$ with $K_t(\omega) := \int_0^{t\wedge T} H_s(\omega) ds$). Moreover,

$$|\widetilde{H}_t^{(m)}(\omega)| \leq \sup_{s\geq 0} |H_s(\omega)|$$

[1] Working with $\mathcal{L}_2$ we use equivalence classes: $K \sim H$ if $(\lambda \times \mathbb{P})((t,\omega) \in [0,\infty) \times \Omega : H_t(\omega) \neq K_t(\omega)) = 0$.

and we can apply step (a) and find

$$\widetilde{H}^{(m,n)} = \left(\widetilde{H}_t^{(m,n)}\right)_{t\geq 0} \in \mathcal{H}_0$$

such that

$$\lim_{n\to\infty} \mathbb{E}\int_0^T |\widetilde{H}_t^{(m,n)} - \widetilde{H}_t^{(m)}|^2 dt = 0. \tag{3.1}$$

Let

$$\mathcal{A} := \left\{(t,\omega) \in [0,T]\times\Omega : \lim_{m\to\infty} \widetilde{H}_t^{(m)}(\omega) = H_t(\omega)\right\}^c \in \mathcal{B}([0,T])\otimes\mathcal{F}_T.$$

Let

$$\mathcal{A}_\omega := \{t \in [0,T] : (t,\omega) \in \mathcal{A}\}.$$

Fubini's theorem implies that $\mathcal{A}_\omega \in \mathcal{B}([0,T])$. Moreover, $t \in \mathcal{A}_\omega^c$ whenever

$$\widetilde{H}_t^{(m)}(\omega) = m\int_{(t-\frac{1}{m})\vee 0}^t H_s(\omega)ds \to_m H_t(\omega).$$

By the fundamental theorem of calculus we get that

$$\lambda(\mathcal{A}_\omega^c) = T \quad \text{or} \quad \lambda(\mathcal{A}_\omega) = 0$$

where λ is the Lebesgue measure. Hence by Fubini's theorem,

$$(\lambda\otimes\mathbb{P})(\mathcal{A}) = 0$$

so that by majorized convergence,

$$\lim_m \mathbb{E}\int_0^T |\widetilde{H}_t^{(m)} - H_t|^2 dt = 0. \tag{3.2}$$

Combining (3.1) and (3.2) gives the existence of $H^{(n)} = (H_t^{(n)})_{t\geq 0} \in \mathcal{L}_0$ such that

$$\lim_n \mathbb{E}\int_0^T |H_t^{(n)} - H_t|^2 dt = 0.$$

step c: From **step b** we get that there are $H^{(N)} = (H_t^{(N)})_{t\geq 0} \in \mathcal{L}_0$ such that

$$\mathbb{E}\int_0^N \left|H_t^{(N)} - H_t\right|^2 dt \leq \frac{1}{N^2}$$

for $N = 1, 2, ...$ Hence

$$\|H^{(N)} - H\|_{\mathcal{L}_2,T} \leq \frac{1}{N} \quad \text{for} \quad N \geq T$$

and $\mathtt{d}(H^{(N)}, H) \to_N 0$ as $N \to \infty$.

step d: Now we remove the condition that H is bounded. For $k = 1, 2, ...$ we let

$$H_t^{(k)} := \begin{cases} k & : \quad H_t \geq k \\ H_t & : \quad H_t \in [-k, k] \\ -k & : \quad H_t \leq -k \end{cases} .$$

By dominated convergence we have that $\|H - H^{(k)}\|_{\mathcal{L}_2,T} \to 0$ as $k \to \infty$ for all $T > 0$. But the $H^{(k)}$ can be approximated by elements of $\mathcal{L}_0$ so that we are done. □

Proposition 3.1.12 (extension of the Itô integral to $\mathcal{L}_2$)**.** *The map*

$$I : \mathcal{L}_0 \to \mathcal{M}_2^{c,0}$$

can be extended to a map

$$J : \mathcal{L}_2 \to \mathcal{M}_2^{c,0}$$

such that the following holds:

(i) *Extension property: if $H \in \mathcal{L}_0$, then*

$$I_t(H) = J_t(H), \quad t \geq 0, a.s.$$

(ii) *Expectation and Itô isometry: if $H \in \mathcal{L}_2$, then, for $t \geq 0$,*

$$\mathbb{E} J_t(H) = 0 \quad \text{and} \quad \|J_t(H)\|_{L_2} = \left(\mathbb{E} \int_0^t L_u^2 du\right)^{\frac{1}{2}}.$$

(iii) *Linearity: for $\alpha, \beta \in \mathbb{R}$ and $K, H \in \mathcal{L}_2$ one has*

$$J_t(\alpha K + \beta H) = \alpha J_t(K) + \beta J_t(H) \quad \text{for all } t \geq 0 \text{ a.s.}$$

(iv) *Uniqueness property: if $J' : \mathcal{L}_2 \to \mathcal{M}_2^{c,0}$ is another mapping satisfying (i),...,(iii), then one has that*

$$\mathbb{P}(J_t(L) = J'_t(L) : t \geq 0) = 1 \quad \text{for all} \quad L \in \mathcal{L}_2.$$

Proof. By Proposition 3.1.11 there exist $H^{(n)} \in \mathcal{L}_0$ with $\mathtt{d}(H, H^{(n)}) \to_n 0$. By DOOB's maximal inequality we have

$$\begin{aligned} \mathbb{E} \sup_{t\in[0,N]} |I_t(H^{(n)}) - I_t(H^{(m)})|^2 &\leq 4\mathbb{E}|I_N(H^{(n)}) - I_N(H^{(m)})|^2 \\ &= 4\mathbb{E}\int_0^N |H_t^{(n)} - H_t^{(m)}|^2 dt \qquad (3.3) \\ &< 4\varepsilon^2 \end{aligned}$$

for $m, n \geq n(\varepsilon, N)$. Hence we find $1 = n_0 < n_1 < n_2 < \cdots$ (depending on N) such that

$$\mathbb{E} \sup_{t\in[0,N]} |I_t\left(H^{(n_{k+1})}\right) - I_t\left(H^{(n_k)}\right)|^2 \leq \frac{1}{2^{k+1}}$$

for $k = 1, 2, ...$ Letting

$$a_k^N := \sup_{t\in[0,N]} |I_t\left(H^{(n_{k+1})}\right) - I_t\left(H^{(n_k)}\right)|$$

implies

$$\sum_{k=0}^{\infty} \left(\mathbb{E}|a_k^N|^2\right)^{\frac{1}{2}} < \infty \quad \text{and} \quad \mathbb{P}\left(\sum_{k=0}^{\infty} a_k^N < \infty\right) = 1,$$

where the latter is a consequence of $\sum_{k=0}^{\infty} \left(\mathbb{E}|a_k^N|^2\right)^{\frac{1}{2}} \geq \sum_{k=0}^{\infty} \mathbb{E}a_k^N$. Setting

$$\Omega_N := \left\{\omega \in \Omega : \sum_{k=0}^{\infty} a_k^N(\omega) < \infty\right\},$$

we define for $t \geq 0$

$$J_t^{(N)}(H)(\omega) := \begin{cases} \sum_{k=0}^{\infty} \left[I_t(H^{(n_{k+1})}) - I_t(L^{(n_k)})\right](\omega) & : \quad \omega \in \Omega_N, \\ 0 & : \quad \omega \notin \Omega_N. \end{cases}$$

The $\boxed{\text{process } (J_t^{(N)}(H))_{t\in[0,N]} \text{ is path-wise continuous}}$ as a uniform in time limit on $[0, N]$ of path-wise continuous processes $I(H^{(n_{k+1})})$. Using triangular inequality and (3.3) leads to

$$(\mathbb{E} \sup_{t\in[0,N]} |J_t^{(N)}(H) - I_t(H^{(m)})|^2)^{\frac{1}{2}} - (\mathbb{E} \sup_{t\in[0,N]} |J_t^{(N)}(H) - I_t(H^{(n_k)})|^2)^{\frac{1}{2}}$$

$$
\begin{aligned}
&\le (\mathbb{E} \sup_{t\in[0,N]} |I_t(H^{(n_k)}) - I_t(H^{(m)})|^2)^{\frac{1}{2}} \\
\le\;\; & 2\left(\mathbb{E}\int_0^N |H_t^{(n_k)} - H_t^{(m)}|^2 dt\right)^{\frac{1}{2}} \\
\le\;\; & 2\left(\mathbb{E}\int_0^N |H_t^{(n_k)} - H_t|^2 dt\right)^{\frac{1}{2}} + 2\left(\mathbb{E}\int_0^N |H_t - H_t^{(m)}|^2 dt\right)^{\frac{1}{2}},
\end{aligned}
$$

so that for $k \to \infty$ we get

$$
\boxed{\mathbb{E} \sup_{t\in[0,N]} |J_t^{(N)}(H) - I_t(H^{(m)})|^2 \le 4\mathbb{E}\int_0^N |H_t - H_t^{(m)}|^2 dt.}
$$

This implies that $J^{(N)}(H)$ does not depend on the approximating sequence.

Especially, we have that $\mathbb{E}\int_0^N |H_s - H_s^{(m)}|^2 ds \to 0$ implies

$$
\mathbb{E}|J_t^{(N)}(H) - I_t(H^{(m)})|^2 \to 0 \quad \text{for any } t \in [0,N]. \tag{3.4}
$$

Since $I(H^{(m)})$ is a martingale, and the L_2-limit of martingales is a martingale, also $\boxed{J^{(N)}(H) \text{ is a martingale.}}$

With a standard trick we find the process we are looking for: define, for $0 < N \le M < \infty$,

$$
\Omega_{N,M} := \left\{\omega \in \Omega : J_t^{(N)}(H)(\omega) = J_t^{(M)}(H)(\omega) \text{ for all } t \in [0,N]\right\}
$$

so that $\mathbb{P}(\Omega_{N,M}) = 1$ since $(J_t^{(N)}(H))_{t\in[0,N]}$ and $(J_t^{(M)}(H))_{t\in[0,N]}$ are continuous modifications. Letting

$$
\widetilde{\Omega} := \bigcap_{0<N\le M<\infty} \Omega_{N,M}
$$

we obtain that $\mathbb{P}(\widetilde{\Omega}) = 1$ and set

$$
J_t(H)(\omega) := \begin{cases} J_t^{(N)}(H)(\omega) & : \ \omega \in \widetilde{\Omega} \text{ and } t \in [0,N] \\ 0 & : \ \omega \notin \widetilde{\Omega} \end{cases}.
$$

By changing the process on a null-set (we have the usual conditions) we may assume that $J_0(H) \equiv 0$.

$\boxed{\text{We indeed have found the desired extension } J:}$

- Property (i) follows by construction since we can take $H_n = H$ in this case.
- Property (ii) holds since

$$\mathbb{E}J_t(H) = \mathbb{E}[\mathbb{E}[J_t(H)|\mathcal{F}_0]] = \mathbb{E}J_0(H) = 0$$

by the tower property. By

$$\left| \|J_t(H)\|_{L_2} - \|I_t(H^{(n)})\|_{L_2} \right| \le \|J_t(H) - I_t(H^{(n)})\|_{L_2} \to_n 0$$

and

$$\left| \left(\mathbb{E} \int_0^t H_u^2 du \right)^{\frac{1}{2}} - \left(\mathbb{E} \int_0^t (H_u^{(n)})^2 du \right)^{\frac{1}{2}} \right| \le \left| \left(\mathbb{E} \int_0^t (H_u - H_u^{(n)})^2 du \right)^{\frac{1}{2}} \right| \to 0,$$

and by the Itô isometry for the simple processes we get the Itô isometry in (ii).

- Property (iii) Follows from the linearity of $I_t(H^{(n)})$, the L_2 convergence $\mathbb{E}|I_t(H^{(n)}) - J_t(H)|^2 \to 0$ and the fact that $J(H)$ is pathwise continuous.
- Property (iv) By Proposition 3.1.11 there exist $H^{(n)} \in \mathcal{L}_0$ with $\mathtt{d}(H, H^{(n)}) \to_n 0$. Hence by (3.4), which holds for J' as well,

$$J_t(H) = L_2\text{–}\lim_n J_t(H^{(n)}) = L_2\text{–}\lim_n I_t(H^{(n)}) = L_2\text{–}\lim_n J_t'(H^{(n)}) = J_t'(H).$$

Since $(J_t(H))_{t\ge 0}$ and $(J_t'(H))_{t\ge 0}$ are continuous, the processes are indistinguishable.

$\square$

3.1.4 Examples of Itô integrals

If the integrand is not random the Itô integral is called Wiener integral.

Example 3.1.13 (Wiener integral)**.** For a continuous function $L : [0, \infty) \to \mathbb{R}$ the process

$$X_t := J_t(L)$$

is a Gaussian process with mean zero and covariance

$$\Gamma(s,t) = \mathbb{E}X_s X_t = \int_0^{\min\{s,t\}} L_u^2 du.$$

Proof. (a) First we remark that an L_2-limit of Gaussian random variables is again a Gaussian random variable, which is left as an exercise.

(b) Defining

$$L_t^{(n)} := \sum_{k=1}^{\infty} \chi_{\left(\frac{k-1}{2^n}, \frac{k}{2^n}\right]}(t)\, L\left(\tfrac{k-1}{2^n}\right)$$

we obtain that $\mathtt{d}(L, L^{(n)}) \to_n 0$. Hence by Proposition 3.1.12 (iv)

$$J_t(L^{(n)}) \to_{L_2} J_t(L).$$

But $J_t(L^{(n)}) = I_t(L^{(n)})$ are Gaussian random variables so $J_t(L)$ is Gaussian by step (a) as well.

(c) To check that we have a Gaussian process we let $0 \le t_1 \le \cdots \le t_n \le T$ and $\alpha_1, ..., \alpha_n \in \mathbb{R}$. It is easy to see that

$$\sum_{k=1}^{n} \alpha_k J_{t_k}(L^{(n)})$$

is a Gaussian random variable which converges in L_2 to the random variable $\sum_{k=1}^n \alpha_k J_{t_k}(L)$. Hence the latter is Gaussian and we have a Gaussian process.

(d) To get the covariance structure we simply use the Itô-isometry since one can show that

$$J_r(L) = J_T(K^{(r)})\ a.s.$$

for $K_u^{(r)} := L_u$ for $u \in [0, r]$ and $K_u^{(r)} := 0$ for $u > r$, where $T \ge r$. □

Next we consider our first real example.

Example 3.1.14. *One has that*

$$J_t(B) = \frac{1}{2}(B_t^2 - t) \quad \textit{for} \quad t \geq 0 \;\; \textit{a.s.}$$

Proof. Since the left and right-hand side are continuous processes starting in zero we only have to show that $J_T(B) = \frac{1}{2}(B_T^2 - T)$ a.s. for $T > 0$. Let

$$L_T^{(n)} := \sum_{i=1}^{\infty} \chi_{\left(T\frac{i-1}{2^n}, T\frac{i}{2^n}\right]}(t) B_{T\frac{i-1}{2^n}}.$$

By dominated convergence one can show that

$$\lim_n \mathtt{d}(L^{(n)}, B) = 0.$$

In the next step we would need to show that

$$J_T(L^{(n)}) = \sum_{i=1}^{2^n} B_{T\frac{i-1}{2^n}}(B_{T\frac{i}{2^n}} - B_{T\frac{i-1}{2^n}})$$

(note that $L^{(n)}$ is not a simple integrand according to our definition). This can be easily done by a truncation of the coefficients $B_{T\frac{i-1}{2^n}}$ which brings us back to the simple integrands. So we are left to show that

$$\lim_{n\to\infty} \mathbb{E} \left| \sum_{i=1}^{2^n} B_{T\frac{i-1}{2^n}}(B_{T\frac{i}{2^n}} - B_{T\frac{i-1}{2^n}}) - \frac{1}{2}(B_T^2 - T) \right|^2 = 0.$$

Let $N = 2^n$. Then

$$\begin{aligned}
& \mathbb{E} \left| \sum_{i=1}^{N} B_{T\frac{i-1}{N}}(B_{T\frac{i}{N}} - B_{T\frac{i-1}{N}}) - \frac{1}{2}(B_T^2 - T) \right|^2 \\
= \;& \mathbb{E} \left| \sum_{i=1}^{N} \left[B_{T\frac{i-1}{N}}(B_{T\frac{i}{N}} - B_{T\frac{i-1}{N}}) - \frac{1}{2}\left(B_{T\frac{i}{N}}\right)^2 + \frac{1}{2}\left(B_{T\frac{i-1}{N}}\right)^2 \right] + \frac{T}{2} \right|^2 \\
= \;& \frac{1}{4}\mathbb{E} \left| T - \sum_{i=1}^{N} \left[B_{T\frac{i}{N}} - B_{T\frac{i-1}{N}} \right]^2 \right|^2
\end{aligned}$$

$$\begin{aligned}
&= \frac{T^2}{4}\mathbb{E}\left|1-\sum_{i=1}^N\left[B_{\frac{i}{N}}-B_{\frac{i-1}{N}}\right]^2\right|^2 \\
&= \frac{T^2}{4}\mathbb{E}\left|\sum_{i=1}^N\left[\left[B_{\frac{i}{N}}-B_{\frac{i-1}{N}}\right]^2-\frac{1}{N}\right]\right|^2 \\
&= \frac{T^2}{4}\sum_{i=1}^N\mathbb{E}\left|\left[\left[B_{\frac{i}{N}}-B_{\frac{i-1}{N}}\right]^2-\frac{1}{N}\right]\right|^2 \\
&= \frac{T^2}{4}\sum_{i=1}^N[\mathbb{E}B_1^4-1]\frac{1}{N^2}\to_N 0.
\end{aligned}$$

□

Looking at the above proof we also got

Proposition 3.1.15. *For the Brownian motion $B=(B_t)_{t\ge 0}$ and $T\ge 0$ one has that*

$$\lim_{N\to\infty}\sum_{i=1}^N\left[B_{T\frac{i}{N}}-B_{T\frac{i-1}{N}}\right]^2=T$$

where the convergence in probability is taken.

3.1.5 Step 3: The Itô integral extended from $\mathcal{L}_2$ to $\mathcal{L}_2^{\mathrm{loc}}$

We are close to the stochastic integral we need to have. In this last step we carry out a *localizing procedure* to extend the integral from $\mathcal{L}_2$ to $\mathcal{L}_2^{\mathrm{loc}}$. For this purpose we first need

Lemma 3.1.16. *For any stopping time $\tau:\Omega\to[0,\infty)\cup\{\infty\}$ and progressively measurable process $L=(L_t)_{t\ge 0}$ one has that $L^\tau := (L_t\chi_{\{t\le\tau\}})_{t\ge 0}$ is progressively measurable.*

Definition 3.1.17. (i) Let $\mathcal{L}_2^{\mathrm{loc}}$ be the set of all progressively measurable processes $L=(L_t)_{t\ge 0}$ such that

$$\mathbb{P}\left(\omega\in\Omega:\int_0^t L_u^2(\omega)du<\infty\right)=1\quad\text{for all}\quad t\ge 0.$$

(ii) A sequence $(\tau_n)_{n=0}^{\infty}$ of stopping times is called **localizing** for $L = (L_t)_{t\geq 0} \in \mathcal{L}_2^{\mathrm{loc}}$ provided that

(a) $0 \leq \tau_0(\omega) \leq \tau_1(\omega) \leq \tau_2(\omega) \leq \cdots \leq \infty$ and $\lim_n \tau_n(\omega) = \infty$ for all $\omega \in \Omega$,

(b) $L^{\tau_n} \in \mathcal{L}_2$ for all $n = 0, 1, 2, \ldots$

Remark 3.1.18. (i) One has that $\mathcal{L}_2 \subseteq \mathcal{L}_2^{\mathrm{loc}}$ since

$$\mathbb{E} \int_0^t L_u^2 du < \infty$$

implies by FUBINI's theorem that

$$\int_0^t L_u^2 du < \infty \; a.s.$$

(ii) For every $L \in \mathcal{L}_2^{\mathrm{loc}}$ there exists a localizing sequence $(\tau_n)_{n\geq 1}$: Let $N = 1, 2, \ldots$ and

$$\Omega_N := \left\{ \omega \in \Omega : \int_0^N L_u^2(\omega) du < \infty \right\}.$$

Then $\mathbb{P}\left(\bigcap_{N=1}^{\infty} \Omega_N\right) = 1$ and we may set

$$\tau_n(\omega) := \begin{cases} \inf \left\{ t \geq 0 : \int_0^t L_u^2(\omega) du \geq n \right\} & : \quad \omega \in \bigcap_{N=1}^{\infty} \Omega_N \\ \infty & : \quad \text{else} \end{cases}$$

because $t \to \int_0^t L_u^2(\omega) du$ is a continuous function defined on $[0, \infty)$ with values in $[0, \infty)$ for $\omega \in \bigcap_{N=1}^{\infty} \Omega_N$.

Next we need

Lemma 3.1.19. *Assume $L \in \mathcal{L}_2^{\mathrm{loc}}$. Then there is a unique (up to indistinguishability) adapted and continuous process $X = (X_t)_{t\geq 0}$ with $X_0 \equiv 0$ such that for all localizing sequences $(\tau_n)_{n=0}^{\infty}$ of L one has*

$$\mathbb{P}\big(J_t(L^{\tau_n}) = X_t, \; t \in [0, \tau_n]\big) = 1 \quad \textit{for} \quad n = 0, 1, 2, \ldots$$

Definition 3.1.20. Let $L \in \mathcal{L}_2^{\mathrm{loc}}$. The process X obtained in Lemma 3.1.19 is denoted by

$$\left(\int_0^t L_u dB_u \right)_{t \geq 0}$$

and is called *Itô integral* of (the integrand) L with respect to (the integrator) B. Moreover, for $0 \leq s \leq t < \infty$ we let

$$\int_s^t L_u dB_u = \int_{(s,t]} L_u dB_u := X_t - X_s.$$

Before we summarize some properties of stochastic integrals we introduce local martingales.

Definition 3.1.21. A continuous adapted process $M = (M_t)_{t \geq 0}$ with $M_0 \equiv 0$ is called **local martingale** (we shall write $M \in \mathcal{M}_{\mathrm{loc}}^{c,0}$) provided that there exists an increasing sequence $(\sigma_n)_{n=0}^{\infty}$ of stopping times $0 \leq \sigma_0(\omega) \leq \sigma_1(\omega) \leq \cdots \leq \infty$ with $\lim_n \sigma_n(\omega) = \infty$ such that $M^{\sigma_n} := (M_{t \wedge \sigma_n})_{t \geq 0}$ is a martingale for all $n = 0, 1, 2, ...$

Surprisingly it is not so easy to find local martingales which are not martingales. We indicate a construction, but do not go into any details. The example is intended as motivation for Itô's formula presented in the next section.

Example 3.1.22. Given $d = 1, 2, ...$ we let $(W_t)_{t \geq 0}$ be the d-dimensional standard Brownian motion where $W_t := (B_{t,1}, ..., B_{t,d})$, $W_0 \equiv 0$, and $(B_{t,i})_{t \geq 0}$ are independent Brownian motions. The filtration is obtained as in the one-dimensional case as the augmentation of the natural filtration. Let $d = 3$ and

$$M_t := \frac{1}{|x + W_t|}$$

with $|x| > 0$ where $|\cdot|$ is the Euclidean norm on $\mathbb{R}^d$. Then $M = (M_t)_{t \geq 0}$ is a local martingale, but not a martingale. The process $|x + W_t|$ is a 3-dimensional Bessel process starting in $|x|$.

Proof. To justify the construction one would need the following:

(a) For a d-dimensional standard Brownian motion W with $d \geq 2$ the sets $\{y\}$ with $y \neq 0$ are *polar sets*, that means

$$\mathbb{P}(\tau_y < \infty) = 0 \quad \text{with} \quad \tau_y := \inf\{t \geq 0 : W_t = y\}.$$

(b) For $d \geq 3$ one has that $\mathbb{P}(\lim_{t\to\infty} |W_t| = \infty) = 1$.

(c) Assuming that M is a martingale we would get

$$\mathbb{E}M_t = \mathbb{E}M_0 = \frac{1}{|x|}.$$

But a direct computation yields to

$$\mathbb{E}\frac{1}{|x+W_t|} = \mathbb{E}\frac{1}{|x+\sqrt{t}(g_1,g_2,g_3)|} \to_{t\to\infty} 0$$

where $g_1, g_2, g_3 \sim N(0,1)$ are independent.

(d) How to show that M is a *local* martingale? This gives us a first impression of Itô-formula which will read for $f(\xi_1,\xi_2,\xi_3) := \frac{1}{\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}}$ and $X_t := x + (B_{t,1}, B_{t,2}, B_{t,3})$ as

$$\begin{aligned} f(X_t) &= f(x) + \sum_{i=1}^3 \int_0^t \frac{\partial f}{\partial x_i}(X_u) dB_{u,i} + \frac{1}{2}\int_0^t (\Delta f)(X_u) du \quad a.s. \\ &= f(x) + \sum_{i=1}^3 \int_0^t \frac{\partial f}{\partial x_i}(X_u) dB_{u,i} \end{aligned}$$

where the latter term turns out to be a local martingale. □

Now we summarize some of the properties of our stochastic integral:

Proposition 3.1.23.

(i) *For* $L \in \mathcal{L}_2^{\mathrm{loc}}$ *one has* $\left(\int_0^t L_u dB_u\right)_{t\geq 0} \in \mathcal{M}_{\mathrm{loc}}^{c,0}$.

(ii) *For* $K, L \in \mathcal{L}_2^{\mathrm{loc}}$ *and* $\alpha, \beta \in \mathbb{R}$ *one has*

$$\int_0^t (\alpha K_u + \beta L_u) dB_u = \alpha \int_0^t K_u dB_u + \beta \int_0^t L_u dB_u, \ t \geq 0, \ a.s.$$

(iii) *Itô-Isometry: for $K, L \in \mathcal{L}_2$ and $0 \le s < t < \infty$ one has*

$$\mathbb{E}\left(\int_s^t K_u dB_u \int_s^t L_u dB_u | \mathcal{F}_s\right) = \mathbb{E}\left(\int_s^t K_u L_u du | \mathcal{F}_s\right) \text{ a.s.}$$

(iv) *Given $K \in \mathcal{L}_2$, the process*

$$\left(\left(\int_0^t K_u dB_u\right)^2 - \int_0^t K_u^2 du\right)_{t \ge 0}$$

is a continuous martingale.

(v) *For $L \in \mathcal{L}_2^{\mathrm{loc}}$ and a stopping time $\tau : \Omega \to [0, \infty]$ one has that*

$$\left(\int_0^{t \wedge \tau(\omega)} L_u dB_u\right)(\omega) = \left(\int_0^t L_u \chi_{\{u \le \tau\}} dB_u\right)(\omega)$$

for $t \ge 0$ a.s.

(vi) *For $0 \le s \le t < \infty$, $B \in \mathcal{F}_s$, and $K \in \mathcal{L}_2^{\mathrm{loc}}$ one has that*

$$\int_s^t [\chi_B K_u] dB_u = \chi_B \int_s^t K_u dB_u \text{ a.s.}$$

with the convention that

$$\int_s^t [\chi_B K_u] dB_u := \int_0^t \widetilde{K}_u dB_u$$

with

$$\widetilde{K}(u) := \begin{cases} 0 & : \quad u \in [0, s] \\ \chi_B K_u & : \quad u > s \end{cases} .$$

Proof. *(iii)* $\Rightarrow$ *(iv)* Let

$$M_t := \int_0^t K_u dB_u.$$

Then, a.s.,

$$\mathbb{E}\left(M_t^2 - \int_0^t K_u^2 du | \mathcal{F}_s\right)$$

$$
\begin{aligned}
&= \mathbb{E}\left((M_t - M_s + M_s)^2 - \int_0^t K_u^2 du | \mathcal{F}_s\right) \\
&= \mathbb{E}\left((M_t - M_s)^2 + 2(M_t - M_s)M_s + M_s^2 - \int_0^t K_u^2 du | \mathcal{F}_s\right) \\
&= \mathbb{E}\left((M_t - M_s)^2 - \int_s^t K_u^2 du | \mathcal{F}_s\right) + \mathbb{E}\left(2(M_t - M_s)M_s | \mathcal{F}_s\right) \\
&\quad + M_s^2 - \int_0^s K_u^2 du \\
&= M_s^2 - \int_0^s K_u^2 du
\end{aligned}
$$

where the first term is zero because of (iii) and the second one because of

$$\mathbb{E}\left(2(M_t - M_s)M_s | \mathcal{F}_s\right) = 2M_s \mathbb{E}\left(M_t - M_s | \mathcal{F}_s\right) = 0 \;\; a.s.$$

(iii) Using the polarization formula $ab = \frac{1}{4}\left((a+b)^2 - (a-b)^2\right)$ it is enough to show the assertion for $K = L$. We take $L^{(n)} \in \mathcal{L}_0$ such that $d(L^{(n)}, L) \to_n 0$ and get that

$$\mathbb{E}\int_0^t |L_u^{(n)} - L_u|^2 du \to_n 0.$$

By the construction of the stochastic integral we have

$$\mathbb{E}\int_0^t |L_u^{(n)} - L_u|^2 du = \mathbb{E}\left|\int_0^t L_u^{(n)} dB_u - \int_0^t L_u dB_u\right|^2 \to_n 0.$$

Fact 3.1.24. *Let (M, Σ, μ) be a probability space and $\mathcal{G} \subseteq \Sigma$ be a sub-σ-algebra. Assume random variables $f, f_n \in L_2$ such that $\mathbb{E}|f_n - f|^2 \to_n 0$. Then*

$$\mathbb{E}\left|\mathbb{E}(f_n^2 | \mathcal{G}) - \mathbb{E}(f^2 | \mathcal{G})\right| \to_n 0.$$

Let now $f_n := \int_s^t L_u^{(n)} dB_u$ and $f := \int_s^t L_u dB_u$. Then

$$\|f_n - f\|_{L_2} = \left(\mathbb{E}\int_s^t (L_u^{(n)} - L_u)^2 du\right)^{\frac{1}{2}} \le \left(\mathbb{E}\int_0^t (L_u^{(n)} - L_u)^2 du\right)^{\frac{1}{2}} \to_n 0.$$

Hence, by our fact,

$$\mathbb{E}\left(\left(\int_s^t L_u^{(n)} dB_u\right)^2 | \mathcal{F}_s\right) \to_{L_1} \mathbb{E}\left(\left(\int_s^t L_u dB_u\right)^2 | \mathcal{F}_s\right).$$

Considering the product space $\Omega \times [s,t]$ with $\mu := \mathbb{P} \times \frac{\lambda}{t-s}$ and $\mathcal{G} := \mathcal{F}_s \otimes \mathcal{B}([s,t])$ we get in the same way that

$$\mathbb{E}\left(\int_s^t \left(L_u^{(n)}\right)^2 du|\mathcal{F}_s\right) \to_{L_1} \mathbb{E}\left(\int_s^t (L_u)^2 du|\mathcal{F}_s\right).$$

Now we can finish with Proposition 3.1.8.

We do not give the details for (i), (ii), (v) and (vi). □

3.2 Itô's formula

In calculus there is the fundamental formula

$$f(y) = f(x) + \int_x^y f'(u)du$$

for, say, $f \in C^1(\mathbb{R})$ and $-\infty < x < y < \infty$. Is there a similar formula for stochastic integrals? We develop such a formula for a class of processes, called Itô-processes.

Definition 3.2.1. *Let $f : [0,\infty) \to \mathbb{R}$ be a function. Then*

$$\mathrm{V}_t^1(f) := \sup_{n\in\mathbb{N}} \sup_{0=t_0\le\cdots\le t_n=t} \sum_{k=1}^n |f(t_k) - f(t_{k-1})| \in [0,\infty].$$

Lemma 3.2.2.

(i) *The function $t \mapsto \mathrm{V}_t^1(f)$ is increasing.*

(ii) *If $f : [0,\infty) \to \mathbb{R}$ is continuous, then $t \mapsto \mathrm{V}_t^1(f)$ is left-continuous.*

Definition 3.2.3. A stochastic process $A = (A_t)_{t\ge 0}$, $A_t : \Omega \to \mathbb{R}$, is called of *bounded variation* provided that

$$\mathrm{V}_t^1(A_\cdot(\omega)) = \sup_{n\in\mathbb{N}} \sup_{0=t_0\le\cdots\le t_n=t} \sum_{k=1}^n |A_{t_k}(\omega) - A_{t_{k-1}}(\omega)| < \infty \ a.s. \quad \text{for all} \quad t \ge 0.$$

Lemma 3.2.4. *If $M = (M_t)_{t\geq 0} \in \mathcal{M}_2^{c,0}$ is of bounded variation, then*

$$\mathbb{P}\left(\{\omega \in \Omega : M_t(\omega) = 0, t \geq 0\}\right) = 1.$$

Proof. Since M has continuous paths it is sufficient to show that

$$\mathbb{P}(M_t = 0) = 1 \quad \text{for all} \quad t \geq 0.$$

(a) Assume that

$$\mathrm{V}_t^1(M.(\omega)) \leq c < \infty \ a.s.$$

and let $t_i^n := \frac{it}{n}$. Then

$$\begin{aligned}
\mathbb{E}M_t^2 &= \mathbb{E}\left[\sum_{i=1}^n \left[M_{t_i^n} - M_{t_{i-1}}^n\right]\right]^2 \\
&= \sum_{i=1}^n \mathbb{E}\left[M_{t_i^n} - M_{t_{i-1}}^n\right]^2 \\
&\leq \mathbb{E}\mathrm{V}_t^1(M.) \sup_{i=1,\dots,n} \left|M_{t_i^n} - M_{t_{i-1}^n}\right| \\
&\leq c\,\mathbb{E} \sup_{i=1,\dots,n} \left|M_{t_i^n} - M_{t_{i-1}^n}\right|.
\end{aligned}$$

Since

$$\sup_{i=1,\dots,n} \left|M_{t_i^n}(\omega) - M_{t_{i-1}^n}(\omega)\right| \to_n 0$$

for all $\omega \in \Omega$ by the uniform continuity of the paths of M on compact intervals and

$$\sup_{i=1,\dots,n} |M_{t_i^n} - M_{t_{i-1}^n}| \leq 2 \sup_{u\in[0,t]} |M_u| \in L_2$$

by DOOB's maximal inequality, dominated convergence implies that

$$\lim_n \mathbb{E} \sup_{i=1,\dots,n} \left|M_{t_i^n} - M_{t_{i-1}^n}\right| = 0 \quad \text{so that} \quad \mathbb{E}M_t^2 = 0.$$

(b) Now let $N \in \{1, 2, ...\}$, $T > 0$, and

$$\tau_N(\omega) := \inf\left\{t \geq 0 : \mathrm{V}_t^1(M.(\omega)) > N\right\} \wedge T.$$

Because of Lemma 3.2.2 the random time τ_N is a stopping time. To check this it is sufficient to show that

$$\sigma_N(\omega) := \inf\left\{t \geq 0 : \mathrm{V}_t^1(M.(\omega)) > N\right\}$$

is a stopping time. Indeed

$$\{t \le \sigma_N(\omega)\} = \left\{ \mathrm{V}_t^1(M_\cdot(\omega)) \le N \right\} \in \mathcal{F}_t$$

yields that σ_N is an optional time, so that we conclude that σ_N is a stopping time by the usual conditions. Moreover,

$$(M_{t\wedge\tau_N})_{t\ge 0} \in \mathcal{M}_2^{c,0}$$

by stopping and

$$\mathrm{V}_t^1(M_\cdot^{\tau_N}(\omega)) \le N.$$

By stopping we mean here that

$$\mathbb{E}(M_{t\wedge\rho}|\mathcal{F}_s) = M_{s\wedge\rho} \quad a.s. \tag{3.5}$$

for a stopping time $0 \le \rho \le T$. This can be proved by considering approximating stopping times $T \ge \rho_l \downarrow \rho$ where the ρ_l takes only values in $\{0, T/2^l, ..., kT/2^l, ..., T, s\}$. Stopping from discrete time martingales implies

$$\mathbb{E}(M_{t\wedge\rho_l}|\mathcal{F}_s) = M_{s\wedge\rho_l} \quad a.s.$$

Letting $l \to \infty$, using the pathwise continuity of M, and that (by Doob's maximal inequality) $\mathbb{E}\sup_{t\in[0,T]} M_t^2 < \infty$, we arrive at (3.5). After all, applying (a) gives

$$\mathbb{E}M_{\sigma_N\wedge T}^2 = 0.$$

Now our assumption implies $\sigma_N \to \infty$ a.s. Using again $\mathbb{E}|M_T^*|^2 < \infty$ we derive $\mathbb{E}|M_T| = \mathbb{E}\lim_{N\to\infty}|M_{\tau_N\wedge T}| = \lim_{N\to\infty}\mathbb{E}|M_{\tau_N\wedge T}| = 0$. □

Now we derive

Lemma 3.2.5. *If $M = (M_t)_{t\ge 0} \in \mathcal{M}_{\mathrm{loc}}^{c,0}$ is of bounded variation, then*

$$\mathbb{P}\left(\omega \in \Omega : M_t(\omega) = 0, t \ge 0\right) = 1.$$

Proof. We assume a localizing sequence $(\sigma_n)_{n=0}^{\infty}$ for M. In addition, we let

$$\rho_n := \inf\left\{t \ge 0 : |M_t| \ge n\right\}$$

so that $\tau_n := \sigma_n \wedge \rho_n$ is a localizing sequence with $|M_t^{\tau_n}| \leq n$ (again, stopping is used here). The variation of M^{τ_n} is bounded by the variation of M, so that

$$\mathbb{P}(M_{t \wedge \tau_n} = 0) = 1$$

for all $t \geq 0$ and $n = 0, 1, 2, ...$ by Lemma 3.2.4. Consequently,

$$\mathbb{P}(M_t = 0) = \mathbb{E}\left(\lim_{n\to\infty} \chi_{\{M_{t\wedge\tau_n}=0\}}\right) = \lim_{n\to\infty} \mathbb{E}\left(\chi_{\{M_{t\wedge\tau_n}=0\}}\right) = 1$$

where we have used dominated convergence and $\lim_n \tau_n(\omega) = \infty$ for all $\omega \in \Omega$. □

Definition 3.2.6 (Itô process)**.** A continuous and adapted process $X = (X_t)_{t\geq 0}$, $X_t : \Omega \to \mathbb{R}$, is called Itô-*process* provided there exist $L \in \mathcal{L}_2^{\mathrm{loc}}$ and a progressively measurable process $a = (a_t)_{t\geq 0}$ with

$$\int_0^t |a_u(\omega)| du < \infty$$

for all $t \geq 0$ and $\omega \in \Omega$, and $x_0 \in \mathbb{R}$ such that

$$X_t(\omega) = x_0 + \left(\int_0^t L_u dB_u\right)(\omega) + \int_0^t a_u(\omega) du \quad \text{for} \quad t \geq 0, \ a.s.$$

Proposition 3.2.7. *Assume that $X = (X_t)_{t\geq 0}$ is an* Itô*-process with representations (L, a) and (L', a'). Then*

$$(\lambda \times \mathbb{P})\big((t,\omega) \in [0,\infty) \times \Omega : L_t(\omega) \neq L'_t(\omega) \ \text{or} \ a_t(\omega) \neq a'_t(\omega)\big) = 0.$$

Proof. We only prove the part concerning the processes L and L'. From our assumption and the linearity of the stochastic integral it follows that

$$M_t := \int_0^t (L_u - L'_u) dB_u = \int_0^t (a'_u - a_u) du \quad \text{for} \quad t \geq 0 \ a.s.$$

Hence $M = (M_t)_{t\geq 0}$ is a local martingale of bounded variation because, a.s.,

$$\mathrm{var}(M_\cdot(\omega), t) \quad = \quad \sup_{t_k} \left\{ \sum_{k=1}^n \left| M_{t_k}(\omega) - M_{t_{k-1}}(\omega) \right| \right\}$$

$$\begin{aligned}
&= \sup_{t_k} \left\{ \sum_{k=1}^{n} \left| \int_{t_{k-1}}^{t_k} (a'_u(\omega) - a_u(\omega)) du \right| \right\} \\
&\leq \sup_{t_k} \left\{ \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} |a'_u(\omega) - a_u(\omega)| du \right\} \\
&= \int_0^t |a'_u(\omega) - a_u(\omega)| du \\
&< \infty.
\end{aligned}$$

By Lemma 3.2.5 this implies $M_t = 0$ a.s. so that by the continuity of the process M one ends up with

$$M_t = 0 \quad \text{for} \quad t \geq 0, \ \ a.s.$$

and

$$\int_0^t (L_u - L'_u) dB_u = 0 \quad \text{for} \quad t \geq 0, \ \ a.s.$$

Since $L - L' \in \mathcal{L}_2^{\text{loc}}$ we find a sequence of stopping times $0 \leq \tau_0 \leq \tau_1 \leq \cdots$ converging to infinity such that $(L - L')^{\tau_n} \in \mathcal{L}_2$ and by the definition of the stochastic integral

$$\int_0^{t \wedge \tau_n} (L_u - L'_u) \chi_{\{u \leq \tau_n\}} dB_u = \int_0^{t \wedge \tau_n} (L_u - L'_u) dB_u = 0 \ \ a.s.$$

By Proposition 3.1.23(v)

$$\int_0^t (L_u - L'_u) \chi_{\{u \leq \tau_n\}} dB_u = \int_0^{t \wedge \tau_n} (L_u - L'_u) \chi_{\{u \leq \tau_n\}} dB_u = 0 \ \ a.s.$$

and by the Itô isometry,

$$\mathbb{E} \int_0^t |L_u - L'_u|^2 \chi_{\{u \leq \tau_n\}} du = 0.$$

Monotone convergence gives

$$\mathbb{E} \int_0^\infty |L_u - L'_u|^2 du = 0$$

which implies our assertion with respect to L and L'. □

To formulate Itô's formula we need

Definition 3.2.8. A continuous function $f : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ belongs to $C^{1,2}([0,\infty) \times \mathbb{R})$ provided that all partial derivatives $\partial f/\partial t$, $\partial f/\partial x$, and $\partial^2 f/\partial x^2$ exist on $(0,\infty) \times \mathbb{R}$, are continuous, and can be continuously extended to $[0,\infty) \times \mathbb{R}$.

Theorem 3.2.9 (Itô's formula). *Let $X = (X_t)_{t\geq 0}$ be an Itô-process with representation*

$$X_t = x_0 + \int_0^t L_u dB_u + \int_0^t a_u du,\ t \geq 0, a.s.$$

and let $f \in C^{1,2}([0,\infty) \times \mathbb{R})$. Then one has that

$$\begin{aligned} f(t, X_t) = f(0, X_0) &+ \int_0^t \frac{\partial f}{\partial u}(u, X_u) du + \int_0^t \frac{\partial f}{\partial x}(u, X_u) L_u dB_u \\ &+ \int_0^t \frac{\partial f}{\partial x}(u, X_u) a_u du + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(u, X_u) L_u^2 du \end{aligned}$$

for $t \geq 0$ a.s.

Remark 3.2.10. (i) The assumptions on f and the continuity of the process X ensure that the right-hand side of Itô's formula is well-defined. In particular,

$$\left(\frac{\partial f}{\partial x}(u, X_u) L_u \right)_{u \geq 0} \in \mathcal{L}_2^{\mathrm{loc}}.$$

(ii) To shorten the notation we shall use

$$\int_0^t K_u dX_u := \int_0^t K_u L_u dB_u + \int_0^t K_u a_u du$$

where we fix the decomposition of the process X in the following.

Before we discuss the proof of Itô's formula we consider some examples for its application.

Example 3.2.11 (Compensator)**.** For $f(t,x) = f(x) := x^2$ we obtain

$$X_t^2 = x_0^2 + 2\int_0^t X_u dX_u + \int_0^t L_u^2 du, \ \ t \geq 0, \ a.s.$$

If $a_u \equiv 0$, then we get that

$$X_t^2 - \int_0^t L_u^2 du = x_0^2 + 2\int_0^t X_u L_u dB_u \ \ t \geq 0, \ a.s.$$

is a local martingale. Sometimes the term $\int_0^t L_u^2 du$ is called *compensator* (it compensates X_t^2 to get a local martingale) and denoted by

$$\langle X \rangle_t := \int_0^t L_u^2 du.$$

Example 3.2.12 (Exponential martingale)**.** Let $L \in C[0,\infty)$ and $X_t := \int_0^t L_u dB_u$. Then

$$\mathcal{E}(X)_t := e^{X_t - \frac{1}{2}\int_0^t L_u^2 du} = e^{X_t - \frac{1}{2}\langle X \rangle_t}$$

is a martingale and called *exponential martingale*. To check this, we let

$$f(t,x) := e^{x - \frac{1}{2}\int_0^t L_u^2 du}.$$

Applying Itô's formula gives, a.s.,

$$\begin{aligned}
\mathcal{E}(X)_t &= f(t, X_t) \\
&= f(0, X_0) + \int_0^t \frac{\partial f}{\partial u}(u, X_u) du + \int_0^t \frac{\partial f}{\partial x}(u, X_u) L_u dB_u \\
&\quad + \int_0^t \frac{\partial f}{\partial x}(u, X_u) a_u du + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(u, X_u) L_u^2 du \\
&= f(0, X_0) + \int_0^t f(u, X_u)\left(-\frac{1}{2}\right) L_u^2 du + \int_0^t f(u, X_u) L_u dB_u \\
&\quad + \int_0^t f(u, X_u) a_u du + \frac{1}{2}\int_0^t f(u, X_u) L_u^2 du \\
&= 1 + \int_0^t f(u, X_u) L_u dB_u
\end{aligned}$$

with

$$\begin{aligned}\int_0^t \mathbb{E}|f(u,X_u)L_u|^2du &= \int_0^t \mathbb{E}e^{2\left(X_u-\frac{1}{2}\int_0^u L_v^2dv\right)}L_u^2du \\ &\leq \left[\sup_{u\in[0,t]} L_u^2\right]\int_0^t \mathbb{E}e^{2X_u}du.\end{aligned}$$

So we have to compute $\mathbb{E}e^{2X_u}$. This is easy to verify since X_u is a centered Gaussian random variable with variance $c(u) = \int_0^u L_v^2 dv$. Hence using the change of variable formula and a simple calculation we have

$$\begin{aligned}\mathbb{E}e^{2X_u} &= \mathbb{E}e^{2\sqrt{c_u}g} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{2\sqrt{c_u}x}e^{-\frac{x^2}{2}}dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-2\sqrt{c_u})^2}{2}}e^{2c_u}dx = e^{2c_u},\end{aligned}$$

and

$$\mathbb{E}e^{2X_u} = \mathbb{E}e^{2\sqrt{c_u}g} = e^{2c_u},$$

where $g \sim N(0,1)$. This implies

$$\int_0^t \mathbb{E}|f(u,X_u)L_u|^2du < \infty$$

and that $(\mathcal{E}(X)_t - 1)_{t\geq 0} \in \mathcal{M}_2^{c,0}$. The above integral equation can also be written as a differential equation

$$df(t,X_t) = f(t,X_t)L_t dB_t \quad \text{with} \quad f(0,X_0) = 1$$

or

$$d\mathcal{E}(X)_t = \mathcal{E}(X)_t dX_t \quad \text{with} \quad \mathcal{E}(X)_0 = 1.$$

Example 3.2.13 (Integration by parts)**.** For $\psi \in C^1[0,\infty)$ one has

$$\psi(t)X_t = \psi(0)X_0 + \int_0^t \psi(u)dX_u + \int_0^t X_u\psi'(u)du \text{ a.s.}$$

which follows by using $f(t,x) := \psi(t)x$.

Example 3.2.14. Using Itô's formula we can now establish an important connection between stochastic differential equations and partial differential equations by one example. Assume the following parabolic PDE

$$\frac{\partial G}{\partial t} + \frac{y^2}{2}\frac{\partial^2 G}{\partial y^2} = 0 \tag{3.6}$$

for $(t,y) \in [0,T) \times (0,\infty)$ with the formal boundary condition $G(T,y) = g(y)$ for some fixed $T > 0$. The PDE is called *backwards* equation since the boundary condition is a condition about the *final* time point T. Assume now the *geometric Brownian motion*

$$S_t(\omega) := e^{B_t(\omega) - \frac{t}{2}} \quad \text{for} \quad t \geq 0.$$

Applying Itô's formula we see that S is an Itô-process with

$$S_t = 1 + \int_0^t S_u dB_u.$$

What is the connection between the geometric Brownian motion and the PDE (3.6). Assume that $\mathbb{E} g(S_T)^2 < \infty$ and define

$$G(t,y) := \mathbb{E} g(y S_{T-t}).$$

One can show that there is some $\varepsilon > 0$ such that $G \in C^\infty((-\varepsilon, T) \times (0,\infty))$ and that f satisfies the PDE (3.6). The principal way (without details) is as follows:

Fact 3.2.15. *Let $p(t,x,y) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}}$ for $t > 0$ and $x, y \in \mathbb{R}$. Then*

$$\frac{\partial p}{\partial t} - \frac{1}{2}\frac{\partial^2 p}{\partial x^2} = 0.$$

Moreover, letting $h : \mathbb{R} \to \mathbb{R}$ be a Borel function such that

$$\int_{\mathbb{R}} e^{-ax^2} |h(x)| dx < \infty,$$

then for

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

one has that

$$\frac{\partial u}{\partial t} - \frac{1}{2}\frac{\partial^2 u}{\partial x^2} = 0$$

for $0 < t < \frac{1}{2a}$ and $x \in \mathbb{R}$.

Let now $f(x) := g(\exp(x-(T/2))$. Applying the above fact we get an $\varepsilon > 0$ (without proof) such that for

$$F(t,x) := \mathbb{E}f(x+B_{T-t})$$

one has

$$\frac{F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial x^2} = 0$$

for $(t,x) \in (-\varepsilon, T) \times \mathbb{R}$. Using

$$G(t,y) = \mathbb{E}g(yS_{T-t}) = \mathbb{E}f\left(B_{T-t} + \frac{t}{2} + \log y\right) = F\left(t, \frac{t}{2} + \log y\right)$$

we can derive that G satisfies the PDE (3.6). Applying Itô's formula gives that

$$\begin{aligned} G(t,S_t) &= G(0,S_0) + \int_0^t \frac{\partial G}{\partial y}(u,S_u)S_u dB_u \\ &\quad + \int_0^t \frac{\partial G}{\partial u}(u,S_u)du + \frac{1}{2}\int_0^t \frac{\partial^2 G}{\partial y^2}(u,S_u)S_u^2 du \\ &= G(0,S_0) + \int_0^t \frac{\partial G}{\partial y}(u,S_u)S_u dB_u \end{aligned}$$

for $t \in (0,T)$ a.s. Without proof, we remark that

$$g(S_T) = G(0,S_0) + \int_0^T \frac{\partial G}{\partial y}(u,S_u)S_u dB_u \; a.s.$$

by $t \uparrow T$ where the integrand of the stochastic integral is defined to be zero for $u = T$ and one has that

$$\int_0^T \mathbb{E}\left|\frac{\partial G}{\partial y}(u,S_u)S_u\right|^2 du < \infty.$$

Example 3.2.16 (Computation of moments)**.** ([5], [6]) The following example is of importance in Stochastic Finance when the Brownian motion is replaced by the geometric Brownian motion. For simplicity we use the Brownian motion. Let $f : \mathbb{R} \to \mathbb{R}$ be a Borel function such that $\mathbb{E}f(B_T)^2 < \infty$ for some $T > 0$. Similarly to Example 3.2.14 we can define

$$F(t,x) := \mathbb{E}f(x+W_{T-t})$$

and get, for some $\varepsilon > 0$, a function $F \in C^\infty((-\varepsilon, T) \times \mathbb{R})$ which solves the PDE

$$\frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial x^2} = 0.$$

Again, by ITÔ's formula

$$f(S_T) = F(0, S_0) + \int_0^T \frac{\partial F}{\partial x}(u, B_u) dB_u \; a.s.$$

where

$$\int_0^T \mathbb{E}\left|\frac{\partial F}{\partial x}(u, B_u)\right|^2 du < \infty.$$

We are interested in

$$\mathbb{E}\left|\int_a^b \left[\frac{\partial F}{\partial x}(u, B_u) - \frac{\partial F}{\partial x}(a, B_a)\right] dB_u\right|^2$$

for $0 \leq a < b < T$, which can be interpreted as the quadratic one-step error if $\int_a^b \frac{\partial F}{\partial x}(u, B_u) dB_u$ is approximated by $\frac{\partial F}{\partial x}(a, B_a)(B_b - B_a)$. To compute the error we proceed formally as follows: by the ITÔ-isometry we have

$$\begin{aligned}
\mathbb{E}&\left|\int_a^b \left[\frac{\partial F}{\partial x}(u, B_u) - \frac{\partial F}{\partial x}(a, B_a)\right] dB_u\right|^2 \\
&= \int_a^b \mathbb{E}\left[\frac{\partial F}{\partial x}(u, B_u) - \frac{\partial F}{\partial x}(a, B_a)\right]^2 du.
\end{aligned}$$

Now we rewrite the expression under the integral by ITÔ's formula. We let

$$\mathcal{A} = \frac{\partial}{\partial t} + \frac{1}{2}\frac{\partial^2}{\partial x^2}$$

and get by ITO's formula that

$$\begin{aligned}
&\mathbb{E}\left[\frac{\partial F}{\partial x}(u, B_u) - \frac{\partial F}{\partial x}(a, B_a)\right]^2 \\
&\quad = \mathbb{E}\left[\frac{\partial F}{\partial x}(u, B_u) - \frac{\partial F}{\partial x}(a, B_a)\right]^2 |_{u=a} \\
&\quad\quad + \mathbb{E}\int_a^v \frac{\partial}{\partial x}\left[\frac{\partial F}{\partial x}(v, B_v) - \frac{\partial F}{\partial x}(a, B_a)\right]^2 dB_v
\end{aligned}$$

$$
\begin{aligned}
& +\mathbb{E}\int_a^u \mathcal{A}\left[\frac{\partial F}{\partial x}(v,B_v)-\frac{\partial F}{\partial x}(a,B_a)\right]^2 dv \\
= \quad & \mathbb{E}\int_a^u \mathcal{A}\left[\frac{\partial F}{\partial x}(v,B_v)-\frac{\partial F}{\partial x}(a,B_a)\right]^2 dv \\
= \quad & \mathbb{E}\int_a^u \left[\frac{\partial^2 F}{\partial x^2}(v,B_v)\right]^2 dv
\end{aligned}
$$

since

$$
\begin{aligned}
& \frac{1}{2}\mathcal{A}_{(v,x)}\left[\frac{\partial F}{\partial x}(v,x)-\frac{\partial F}{\partial x}(a,y)\right]^2 \\
= \quad & \left[\frac{\partial F}{\partial x}(v,x)-\frac{\partial F}{\partial x}(a,y)\right]\frac{\partial^2 F}{\partial x\partial t}(v,x) \\
& +\frac{1}{2}\frac{\partial}{\partial x}\left(\left[\frac{\partial F}{\partial x}(v,x)-\frac{\partial F}{\partial x}(a,y)\right]\frac{\partial^2 F}{\partial x^2}(v,B_v)\right) \\
= \quad & \left[\frac{\partial F}{\partial x}(v,x)-\frac{\partial F}{\partial x}(a,y)\right]\frac{\partial^2 F}{\partial x\partial t}(v,x) \\
& +\frac{1}{2}\frac{\partial^2 F}{\partial x^2}(v,x)\frac{\partial^2 F}{\partial x^2}(v,B_v)+\frac{1}{2}\left[\frac{\partial F}{\partial x}(v,x)-\frac{\partial F}{\partial x}(a,y)\right]\frac{\partial^3 F}{\partial x^3}(v,B_v) \\
= \quad & \left[\frac{\partial F}{\partial x}(v,x)-\frac{\partial F}{\partial x}(a,y)\right]\left(\frac{\partial^2 F}{\partial x\partial t}(v,x)+\frac{1}{2}\frac{\partial^3 F}{\partial x^3}(v,B_v)\right) \\
& \qquad\qquad +\frac{1}{2}\left(\frac{\partial^2 F}{\partial x^2}(v,x)\frac{\partial^2 F}{\partial x^2}(v,B_v)\right) \\
= \quad & \frac{1}{2}\left(\frac{\partial^2 F}{\partial x^2}(v,x)\right)^2.
\end{aligned}
$$

Summarizing the computations gives

$$
\begin{aligned}
\mathbb{E}\left[\frac{\partial F}{\partial x}(u,B_u)-\frac{\partial F}{\partial x}(a,B_a)\right]^2 &= \mathbb{E}\int_a^b\int_a^u \mathbb{E}\left(\frac{\partial^2 F}{\partial x^2}(v,x)\right)^2 dvdu \\
&= \mathbb{E}\int_a^b (b-u)\mathbb{E}\left(\frac{\partial^2 F}{\partial x^2}(u,B_u)\right)^2 du.
\end{aligned}
$$

Now take a net $0=t_0<t_1<\cdots<t_n=T$ one can show that

$$
\mathbb{E}\left|\int_0^T \frac{\partial F}{\partial x}(u,B_u)dB_u-\sum_{k=1}^n \frac{\partial F}{\partial x}(t_{k-1},B_{t_{k-1}})(B_{t_k}-B_{t_{k-1}})\right|^2
$$

$$= \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} (t_k - u)\mathbb{E}\left(\frac{\partial^2 F}{\partial x^2}(u, B_u)\right)^2 du.$$

Taking for example $\psi(x) := \chi_{[K,\infty)(x)}$ one can compute that

$$\mathbb{E}\left(\frac{\partial^2 F}{\partial x^2}(u, B_u)\right)^2 \sim \frac{1}{(T-u)^{\frac{3}{2}}}.$$

3.3 Proof of Itô's formula in a simple case

Throughout this section we assume X to be an Itô process such that $L \in \mathcal{L}_2$. Before we start to prove Itô's formula we need the following lemmata:

Lemma 3.3.1. *Let $Y_n \to 0$ a.s. and $Z_n \to Z$ in probability. Then*

$$Y_n Z_n \to_{\mathbb{P}} 0.$$

The lemma will be an exercise.

Lemma 3.3.2. *Let $Y = (Y_t)_{t \geq 0}$ be continuous and adapted, and assume that*

$$\sup_{t \geq 0, \omega \in \Omega} |Y_t(\omega)| < \infty.$$

Then, for $t_i^n := \frac{i}{n} t$, one has that

$$\sum_{i=1}^{n} Y_{t_{i-1}^n}(X_{t_i^n} - X_{t_{i-1}^n}) \quad \to_{\mathbb{P}} \quad \int_0^t Y_u dX_u, \tag{3.7}$$

$$\sum_{i=1}^{n} Y_{t_{i-1}^n}(X_{t_i^n} - X_{t_{i-1}^n})^2 \quad \to_{\mathbb{P}} \quad \int_0^t Y_u L_u^2 du. \tag{3.8}$$

The proof of this lemma is indicated at the end of the section.

Proof of Theorem 3.2.9. We shall prove Itô's formula in the case that $f^{(3)}$ exists and is continuous, and satisfies

$$\sup_{x \in \mathbb{R}, k=0,1,2,3} |f^{(k)}(x)| < \infty.$$

We fix $t > 0$ and let

$$t_i^n := \frac{i}{n} t$$

for $i = 0, ..., n$ be the equidistant time-net on $[0, t]$. Then, a.s.,

$$\begin{aligned}
& f(X_t) - f(X_0) \\
&= \sum_{i=1}^{n} [f(X_{t_i^n} - f(X_{t_{i-1}^n})] \\
&= \sum_{i=1}^{n} \Big[f'(X_{t_{i-1}^n})(X_{t_i^n} - X_{t_{i-1}^n}) + \frac{1}{2} f''(X_{t_{i-1}^n})(X_{t_i^n} - X_{t_{i-1}^n})^2 \\
&\qquad + \frac{1}{6} f^{(3)}(\widetilde{X}_{t_{i-1}^n})(X_{t_i^n} - X_{t_{i-1}^n})^3 \Big] \\
&= \sum_{i=1}^{n} f'(X_{t_{i-1}^n})(X_{t_i^n} - X_{t_{i-1}^n}) + \sum_{i=1}^{n} \frac{1}{2} f''(X_{t_{i-1}^n})(X_{t_i^n} - X_{t_{i-1}^n})^2 \\
&\qquad + \sum_{i=1}^{n} \frac{1}{6} f^{(3)}(\widetilde{X}_{t_{i-1}^n})(X_{t_i^n} - X_{t_{i-1}^n})^3 \\
&=: I_1^n + \frac{1}{2} I_2^n + \frac{1}{6} I_3^n
\end{aligned}$$

where we used TAYLOR's formula with the LAGRANGE remainder. Applying Lemma 3.3.2 we get, in probability, that

$$\begin{aligned}
\mathbb{P} - \lim_n I_1^n &= \int_0^t f'(X_u) dX_u, \\
\mathbb{P} - \lim_n I_2^n &= \int_0^t f''(X_u) L_u^2 du.
\end{aligned}$$

To get

$$\mathbb{P} - \lim_n I_3^n = 0$$

we observe that

$$\begin{aligned}
& \left| \sum_{i=1}^{n} f^{(3)}(\widetilde{X}_{t_{i-1}^n})(X_{t_i^n} - X_{t_{i-1}^n})^3 \right| \\
&\quad \leq \left[\sup_{x \in \mathbb{R}} |f^{(3)}(x)| \sup_{|u-v| \leq t/n} |X_u - X_v| \right] \left[\sum_{i=1}^{n} (X_{t_i^n} - X_{t_{i-1}^n})^2 \right]
\end{aligned}$$

$$=: \quad Y_n Z_n$$

with $Y_n \to 0$ a.s. and $Z_n \to_{\mathbb{P}} \int_0^t L_u^2 du$ according to Lemma 3.3.2. □

Proof of Lemma 3.3.2. (a) First we consider (3.7) and get

$$\sum_{i=1}^n Y_{t_{i-1}^n}(X_{t_i^n} - X_{t_{i-1}^n}) = \sum_{i=1}^n Y_{t_{i-1}^n} \int_{t_{i-1}^n}^{t_i^n} L_u dB_u + \sum_{i=1}^n Y_{t_{i-1}^n} \int_{t_{i-1}^n}^{t_i^n} a_u du.$$

By standard calculus the second term converges for all $\omega \in \Omega$ to

$$\int_0^t Y_u(\omega) a_u(\omega) du.$$

To consider the first term we let

$$K_u^{(n)} := Y_{t_{i-1}^n} L_u \quad \text{for} \quad u \in (t_{i-1}^n, t_i^n]$$

and otherwise zero. Then

$$\begin{aligned}
\mathbb{E}&\left|\sum_{i=1}^n Y_{t_{i-1}^n} \int_{t_{i-1}^n}^{t_i^n} L_u dB_u - \int_0^t Y_u L_u dB_u\right|^2 \\
&= \mathbb{E}\left|\sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n} (Y_{t_{i-1}^n} L_u) dB_u - \int_0^t Y_u L_u dB_u\right|^2 \\
&= \mathbb{E}\left|\int_0^t (K_u^{(n)} - Y_u L_u) dB_u\right|^2 \\
&= \mathbb{E}\int_0^t \left|K_u^{(n)} - Y_u L_u\right|^2 du \\
&\to_n 0
\end{aligned}$$

by dominated convergence since

$$\left|K_u^{(n)} - Y_u L_u\right| \leq [2 \sup_{t \geq 0, \omega \in \Omega} |Y_t(\omega)|] |L_u|$$

and $\lim_n |K_u^{(n)} - Y_u L_u| = 0$ for all $\omega \in \Omega$.

(b) Now we consider (3.8) and prove this statement for $a_u \equiv 0$ and $|L_u(\omega)| \le c$ for all $u \ge 0$ and $\omega \in \Omega$. We get

$$
\begin{aligned}
&\sum_{i=1}^n Y_{t_{i-1}^n} (X_{t_i^n} - X_{t_{i-1}^n})^2 - \int_0^t Y_u L_u^2 du \\
&\quad = \sum_{i=1}^n Y_{t_{i-1}^n} \left[(X_{t_i^n} - X_{t_{i-1}^n})^2 - \int_{t_{i-1}^n}^{t_i^n} L_u^2 du \right] + \sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n} (Y_{t_{i-1}^n} - Y_u) L_u^2 du
\end{aligned}
$$

where the second term converges to zero for all $\omega \in \Omega$. To treat the first sum we let

$$
d_i := (X_{t_i^n} - X_{t_{i-1}^n})^2 - \int_{t_{i-1}^n}^{t_i^n} L_u^2 du
$$

and get a martingale difference sequence since

$$
\mathbb{E}\left((X_{t_i^n} - X_{t_{i-1}^n})^2 | \mathcal{F}_{t_{i-1}^n} \right) = \mathbb{E}\left(\int_{t_{i-1}^n}^{t_i^n} L_u^2 du | \mathcal{F}_{t_{i-1}^n} \right) \quad a.s.
$$

Consequently,

$$
\begin{aligned}
&\mathbb{E}\left| \sum_{i=1}^n Y_{t_{i-1}^n} \left[(X_{t_i^n} - X_{t_{i-1}^n})^2 - \int_{t_{i-1}^n}^{t_i^n} L_u^2 du \right] \right|^2 \\
&\qquad\qquad = \sum_{i=1}^n \mathbb{E} Y_{t_{i-1}^n}^2 d_i^2 \le \sup_{\omega,u} Y_u(\omega)^2 \sum_{i=1}^n \mathbb{E} d_i^2 .
\end{aligned}
$$

Finally,

$$
\mathbb{E} d_i^2 = \mathbb{E}(X_{t_i^n} - X_{t_{i-1}^n})^4 - 2\mathbb{E}(X_{t_i^n} - X_{t_{i-1}^n})^2 \int_{t_{i-1}^n}^{t_i^n} L_u^2 du + \mathbb{E}\left(\int_{t_{i-1}^n}^{t_i^n} L_u^2 du \right)^2
$$

where

$$
\int_{t_{i-1}^n}^{t_i^n} L_u^2 du \le c^2 \frac{t}{n}.
$$

Since

$$
\sum_{i=1}^n \mathbb{E}\left(X_{t_i^n} - X_{t_{i-1}^n} \right)^2 = \mathbb{E} \int_0^t L_u^2 du < \infty
$$

it remains to show that

$$\sum_{i=1}^{n} \mathbb{E}\left(X_{t_i^n} - X_{t_{i-1}^n}\right)^4 \to_n 0.$$

— But this follows from the Burkholder-Davis-Gundy inequality (see Theorem 4.3.1 below)

$$\mathbb{E}|X_{t_i^n} - X_{t_{i-1}^n}|^4 \le c_4 \mathbb{E}\left(\int_{t_{i-1}^n}^{t_i^n} L_u^2 du\right)^{\frac{4}{2}} \le c_4 \left(\frac{t}{n}\right)^2 c^4.$$

□

3.4 For extended reading

Definition 3.4.1. A continuous adapted stochastic process $X = (X_t)_{t\ge 0}$ is called *continuous semi-martingale* provided that

$$X_t = x_0 + M_t + A_t$$

where $x_0 \in \mathbb{R}$, $M \in \mathcal{M}_{\text{loc}}^{c,0}$, and A is of bounded variation with $A_0 \equiv 0$.

Because of Lemma 3.2.5 the decomposition is unique.

Proposition 3.4.2 (Itô's formula for continuous semimartingales)**.** *Let $f \in C^2(\mathbb{R}^d)$ and $X_t = (X_t^1, ..., X_t^d)$ be a vector of continuous semi-martingales. Then one has that, a.s.,*

$$f(X_t) = f(X_0) + \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_u) dX_u^i + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_u) d\langle M^i, M^j\rangle_u$$

where $dX_u^i = dM_u^i + dA_u^i$ and

$$\langle M^i, M^j\rangle_u := \frac{1}{4}\left[\langle M^i + M^j\rangle_u - \langle M^i - M^j\rangle_u\right].$$

3.4.1 Local time

Given a Borel set $A \subseteq \mathbb{R}$ and a Brownian motion $B = (B_t)_{t\geq 0}$ we want to compute the occupation time of B in A until time t, i.e.

$$\Gamma_t(A,\omega) := \int_0^t \chi_A(B_s(\omega))ds = \lambda(s \in [0,t] : B_s(\omega) \in A).$$

It is not difficult to show that $\Gamma_t(A,\omega) = 0$ $\mathbb{P}$-a.s. if $\lambda(A) = 0$ so that one can ask for a density

$$\Gamma_t(A,\omega) = \int_A 2L_t(x,\omega)dx$$

where the factor 2 is for cosmetics reason.

Definition 3.4.3. A stochastic process $L = (L_t(x,\cdot))_{t\geq 0, x\in\mathbb{R}}$ is called *Brownian local time* provided that

(i) $L_t(x,\cdot) : \Omega \to \mathbb{R}$ is $\mathcal{F}_t$-measurable,

(ii) there exists $\Omega_0 \in \mathcal{F}$ of measure one such that for all $\omega \in \Omega_0$ one has

   (a) $(t,x) \to L_t(x,\omega)$ is continuous,

   (b) $\Gamma_t(A,\omega) = \int_A 2L_t(x,\omega)dx$ for all Borel sets $A \subseteq \mathbb{R}$.

To get a candidate for $L_t(x,\cdot)$ we use Itô's formula: Let $\varphi_\varepsilon \in C_0^\infty$ be such that $\operatorname{supp}(\varphi_\varepsilon) \subseteq [-\varepsilon,\varepsilon]$, $\varphi_\varepsilon \geq 0$, and $\int_\mathbb{R} \varphi_\varepsilon(x)dx = 1$. Let

$$f_\varepsilon(x) := \int_{-\infty}^x \int_{-\infty}^y \varphi_\varepsilon(u)dudy$$

so that

$$\begin{aligned} f'_\varepsilon(x) &= \int_{-\infty}^x \varphi_\varepsilon(u)du, \\ f''_\varepsilon(x) &= \varphi_\varepsilon(x). \end{aligned}$$

By Itô's formula, a.s.

$$f_\varepsilon(B_t - a) = f_\varepsilon(-a) + \int_0^t f'_\varepsilon(B_s - a)dB_s + \frac{1}{2}\int_0^t f''_\varepsilon(B_s - a)ds$$

$$= \quad f_\varepsilon(-a) + \int_0^t f_\varepsilon'(B_s - a)dB_s + \frac{1}{2}\int_0^t \varphi_\varepsilon(B_s - a)ds.$$

Now

$$\mathbb{E}\int_0^t |f_\varepsilon'(B_s - a) - \chi_{(0,\infty)}(B_s - a)|^2 ds \to 0$$

and

$$\sup_x |f_\varepsilon(x) - x^+| \to 0$$

as $\varepsilon \downarrow 0$, so that, a.s.,

$$\lim_{\varepsilon_n \downarrow 0} \frac{1}{2}\int_0^t \varphi_{\varepsilon_n}(B_s - a)ds = (B_t - a)^+ - (-a)^+ - \int_0^t \chi_{(a,\infty)}(B_s)dB_s$$

for some sequence $\varepsilon_n \downarrow 0$. But the left-hand side is - formally -

$$\frac{1}{2}\int_0^t \delta(B_s - a)ds = L_t(a, \cdot).$$

Proposition 3.4.4 (Trotter)**.** *The Brownian local time exists.*

Proof. (Idea) (a) Let

$$M_t(a,\omega) := (B_t(\omega) - a)^+ - (-a)^+ - \left(\int_0^t \chi_{(a,\infty)}(B_s)dB_s\right)(\omega).$$

By a version of Kolmogorov's Proposition 2.3.13 one can show that there exists a continuous (in (t,x)) version $L = (L_t(x,\cdot))_{t\geq 0, x\in\mathbb{R}}$ of M. Clearly, $L_t(x,\cdot)$ is $\mathcal{F}_t$-measurable.

(b) We still need to show that for all Borel sets $A \in \mathcal{B}(\mathbb{R})$:

$$\Gamma_t(A,\omega) = \int_A 2L_t(x,\omega)dx. \tag{3.9}$$

Let $-\infty < a_1 < a_2 < b_2 < b_1 < \infty$ and define the continuous function $h : \mathbb{R} \to \mathbb{R}$ as $h(x) = 1$ on $[a_2, b_2]$, zero outside $[a_1, b_1]$, and linear otherwise. Let

$$H(x) := \int_{-\infty}^x \int_{-\infty}^y h(u)dudy = \int_{\mathbb{R}} (x-u)^+ h(u)du$$

so that

$$H'(x) \quad = \quad \int_{-\infty}^x h(u)du = \int_{\mathbb{R}} h(u)\chi_{(u,\infty)}(x)du$$

$$H''(x) \quad = \quad h(x).$$

By Itô's formula,

$$\begin{aligned}
\frac{1}{2}\int_0^t h(B_s)ds &= H(B_t) - H(B_0) - \int_0^t H'(B_s)dB_s \\
&= \int_{\mathbb{R}} \left[h(u)(B_t - u)^+ - h(u)(-u)^+ - \int_0^t h(u)\chi_{(u,\infty)}(B_s)dB_s \right] du \\
&= \int_{\mathbb{R}} M_t(u,\cdot)h(u)du.
\end{aligned}$$

(c) In the last step we replace M by L. Then for each h it holds for $\mathbb{P}$-almost all ω

$$\frac{1}{2}\int_0^t h(B_s(\omega))ds = \int_{\mathbb{R}} L_t(u,\omega)h(u)du.$$

Since both sides are continuous in t, we can find a set $\Omega^* \in \mathcal{F}$ with $\mathbb{P}(\Omega^*) = 1$ such that the equation holds for all functions h with rational a_1, a_2, b_2, b_1 (this is obviously a countable set of functions). We can approximate indicator functions $\chi_{[a_2,b_2]}$ by the functions h and monotone convergence from above implies

$$\frac{1}{2}\int_0^t \chi_{[a_2,b_2]}(B_s(\omega))ds = \int_{\mathbb{R}} L_t(u,\omega)\chi_{[a_2,b_2]}(u)du$$

for all $\omega \in \Omega^*$. From this one can deduce by the monotone class theorem for functions that for every Borel measurable function $f : \mathbb{R} \to [0,\infty)$

$$\frac{1}{2}\int_0^t f(B_s(\omega))ds = \int_{\mathbb{R}} L_t(u,\omega)f(u)du, \quad \omega \in \Omega^*$$

which implies (3.9). □

Formally, we also get the following:

$$\Gamma((a-\varepsilon, a+\varepsilon), \omega) = \int_{a-\varepsilon}^{a+\varepsilon} 2L_t(x,\omega)dx$$

and

$$\lim_{\varepsilon \downarrow 0} \frac{1}{4\varepsilon}\Gamma((a-\varepsilon, a+\varepsilon), \omega) = L_t(a,\omega).$$

Proposition 3.4.5 (Tanaka formulas)**.** *One has that, a.s.,*

$$L_t(a) = (B_t - a)^+ - (-a)^+ - \int_0^t \chi_{(a,\infty)}(B_s)dB_s$$

and

$$2L_t(a) = |B_t - a| - |-a| - \int_0^t \operatorname{sgn}(B_s - a)dB_s,$$

where $\operatorname{sgn}(x) = -1$ *for* $x < 0$ *and* $\operatorname{sgn}(x) = 1$ *for* $x \geq 0$

Proposition 3.4.6 (Itô's formula for convex functions)**.** *For a convex function f and its second derivative μ one has, a.s.,*

$$f(B_t) = f(0) + \int_0^t D^- f(B_s)dB_s + \int_{\mathbb{R}} L_t(x)d\mu(x)$$

where

$$D^- f(x) := \lim_{h \downarrow 0} \frac{1}{h}[f(x) - f(x-h)]$$

and μ is determined by

$$\mu([a,b)) := D^- f(b) - D^- f(a).$$

3.4.2 Three-dimensional Brownian motion is transient

We would like to prove, that the three-dimensional Brownian motion is transient[2]. For this we let $B_t = (B_t^1, ..., B_t^d)$ a d-dimensional standard Brownian motion where the filtration is taken to be the augmentation of the natural filtration and the usual conditions are satisfied. The process

$$R_t := |x_0 + B_t|$$

where $|\cdot|$ is the d-dimensional euclidean norm is called d-dimensional Bessel process starting in $x_0 \in \mathbb{R}^d$. We want to prove the following

Proposition 3.4.7. *Let $d = 3$ and $0 < c < r = |x_0|$. Then one has that*

$$\mathbb{P}\left(\inf_{t \geq 0} R_t \leq c\right) = \frac{c}{r}.$$

[2]Transient: passing especially quickly into and out of existence.

Proof. Let

$$\tau := \inf\{t \geq 0 : R_t = c\} \quad \text{and} \quad \sigma_k := \inf\{t \geq 0 : R_t = k\}$$

for an integer $k > r$. Let

$$\rho_{k,n} := \tau \wedge \sigma_k \wedge n.$$

By Itô's formula we get for $f(x) := 1/|x|$

$$\begin{aligned} \frac{1}{R_{\rho_{k,n}}} &= \frac{1}{r} + \sum_{i=1}^{3} \int_0^{\rho_{k,n}} \frac{\partial}{\partial x_i} f(x_0 + B_u) dB_u^i \\ &\quad + \frac{1}{2} \sum_{i,j=1}^{3} \int_0^{\rho_{k,n}} \frac{\partial^2}{\partial x_i \partial x_j} f(x_0 + B_u) d < B^i, B^j >_u \\ &= \frac{1}{r} + \sum_{i=1}^{3} \int_0^{\rho_{k,n}} \frac{\partial}{\partial x_i} f(x_0 + B_u) dB_u^i \end{aligned}$$

since $< B^i, B^j >_u = 0$ for $i \neq j$,

$$\frac{\partial}{\partial x_i} \frac{1}{(x_1^d + ... + x_d^2)^{\frac{1}{2}}} = -\frac{x_i}{(x_1^d + ... + x_d^2)^{\frac{3}{2}}}$$

and

$$\sum_{i=1}^{3} \frac{\partial^2}{\partial x_i^2} \frac{1}{(x_1^d + ... + x_d^2)^{\frac{1}{2}}} = -\frac{3}{(x_1^d + ... + x_d^2)^{\frac{3}{2}}} + 3 \sum_{i=1}^{3} \frac{x_i^2}{(x_1^d + ... + x_d^2)^{\frac{5}{2}}} = 0.$$

Taking the expected value gives

$$\frac{1}{r} = \mathbb{E} \frac{1}{R_{\rho_{k,n}}} = \frac{1}{c} \mathbb{P}(\tau \leq \sigma_k \wedge n) + \frac{1}{k} \mathbb{P}(\sigma_k \leq \tau \wedge n) + \mathbb{E} \frac{1}{R_n} \chi_{\{n < \sigma_k \wedge \tau\}}.$$

By $n \to \infty$ we get that

$$\frac{1}{r} = \frac{1}{c} \mathbb{P}(\tau \leq \sigma_k) + \frac{1}{k} \mathbb{P}(\sigma_k \leq \tau).$$

By $k \to \infty$ we end up with

$$\frac{1}{r} = \frac{1}{c} \mathbb{P}(\tau < \infty).$$

The observation that $\{\tau < \infty\} = \{\inf_{t \geq 0} R_t \leq c\}$ finishes the proof. $\square$

Now we can prove:

Proposition 3.4.8. *The 3-dimensional Brownian motion is transiient, i.e.*

$$\mathbb{P}(\lim_{t\to\infty} |B_t| = \infty) = 1.$$

Proof. (a) We begin with $R = (R_t^r)_{t\geq 0}$, a 3-dimensional Bessel process starting at some point x_0 with $R_0 = |x_0| = r > 0$. For $c_k = r/2^k$ we get by Proposition 3.4.7 that

$$\mathbb{P}\left(\inf_{t\geq 0} R_t^r \leq \frac{r}{2^k}\right) = \frac{1}{2^k} \quad \text{for} \quad k = 1, 2, \ldots$$

(b) Given $\epsilon \in (0,1)$ and $L > 0$, we find a k with $1/2^k < \epsilon$ and r such that $r/2^k = L$. If the Brownian motion would start at level r, then

$$\mathbb{P}\left(\inf_{t\geq 0} R_t^r \leq L\right) = 1/2^k < \epsilon.$$

Let $\sigma := \inf\{t \geq 0 : |B_t| = r\}$, so that we get $\mathbb{P}(\sigma < \infty) = 1$. This means that the modulus of the Brownian motion can reach the level r with probability one in finite time. By the strong Markov property one can decompose the Brownian motion into a Brownian motion which is stopped at the first hitting of the level r and an independent Brownian motion starting from there. Hence for our process B_t starting in 0, we get

$$\mathbb{P}\left(\liminf_{t\to\infty} |B_t| \leq L\right) < \epsilon \text{ for all } \epsilon \in (0,1) \text{ and } L > 0.$$

This implies $\mathbb{P}\left(\liminf_{t\to\infty} |B_t| \leq L\right) = 0$ for all $L > 0$ and

$$\mathbb{P}\left(\liminf_{t\to\infty} |B_t| = \infty\right) = 1.$$

□

Chapter 4

Stochastic differential equations

Stochastic differential equations (SDEs) play an important role in stochastic modeling. For example, in economics solutions of the SDEs considered below are used to model share prices. In biology solutions of stochastic partial differential equations (not considered here) describe sizes of populations.

4.1 What is a stochastic differential equation?

Stochastic differential equations are (for us) a formal abbreviation of *integral* equations as described now. Throughout the whole chapter we assume a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\geq 0})$ which satisfies the usual assumptions and an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion $B = (B_t)_{t\geq 0}$.

Definition 4.1.1. Let $x_0 \in \mathbb{R}$, $D \subseteq \mathbb{R}$ be an open set, and $\sigma, a : [0,\infty)\times D \to \mathbb{R}$ be continuous. A path-wise continuous and adapted stochastic process $X = (X_t)_{t\geq 0}$ is a solution of the *stochastic differential equation* (SDE)

$$dX_t = \sigma(t, X_t)dB_t + a(t, X_t)dt \quad \text{with} \quad X_0 = x_0 \tag{4.1}$$

provided that the following conditions are satisfied:

(i) $X_t(\omega) \in D$ for all $t \geq 0$ and $\omega \in \Omega$.

(ii) $X_0 \equiv x_0$.

(iii) $X_t = x_0 + \int_0^t \sigma(u, X_u)dB_u + \int_0^t a(u, X_u)du$ for $t \geq 0$ a.s.

It follows from the continuity of X and (σ, b), and definition of the solution, that

$$\int_0^t |a(u, X_u(\omega))| du + \int_0^t |\sigma(u, X_u(\omega))|^2 du < \infty \quad \text{for all} \quad \omega \in \Omega$$

and that (in particular) $(\sigma(u, X_u))_{u \geq 0} \in \mathcal{L}_2^{\mathrm{loc}}$.

Let us give some examples of SDEs.

Example 4.1.2 (Brownian motion)**.** A solution of

$$dX_t = dB_t \quad \text{and} \quad X_0 = 0$$

is the Brownian motion itself $B = (B_t)_{t \geq 0}$ since $B_t = \int_0^t 1 dB_u$. We can take $D = \mathbb{R}$.

Example 4.1.3 (Geometric Brownian motion with drift)**.** Letting $X_t := x_0 e^{cB_t + bt}$ with $x_0, b, c \in \mathbb{R}$ we obtain by Itô's formula that, a.s.,

$$\begin{aligned} X_t &= x_0 + \int_0^t c X_u dB_u + \int_0^t b X_u du + \frac{1}{2} \int_0^t c^2 X_u du \\ &= x_0 + \int_0^t c X_u dB_u + \int_0^t \left[b + \frac{1}{2} c^2 \right] X_u du \\ &= x_0 + \int_0^t \sigma X_u dB_u + \int_0^t a X_u du \end{aligned}$$

with

$$\begin{aligned} \sigma &:= c, \\ a &:= b + \frac{1}{2} c^2. \end{aligned}$$

Going the other way round by starting with a and σ, we get that

$$\begin{aligned} c &= \sigma, \\ b &= a - \frac{1}{2} \sigma^2. \end{aligned}$$

Consequently, the SDE

$$dX_t = \sigma X_t dB_t + a X_t dt \quad \text{with} \quad X_0 = x_0$$

is solved by

$$X_t = x_0 e^{\sigma B_t + \left(a - \frac{1}{2}\sigma^2\right) t}.$$

We may use $D = \mathbb{R}$ for $\sigma(t, x) := \sigma x$ and $a(t, x) := a x$.

The following examples only provide SDEs formally. We do not discuss solvability at this point.

Example 4.1.4 (Ornstein-Uhlenbeck process)**.** Here one considers the SDE

$$dX_t = -cX_t dt + \sigma dB_t \quad \text{with} \quad X_0 = x_0.$$

We close by some examples from Stochastic Finance.

Example 4.1.5 (Vasicek interest rate model)**.** Here one considers that

$$dr_t = [a - br_t]dt + \sigma dB_t \quad \text{with} \quad r_0 \geq 0,$$

$\sigma \geq 0$, and $a, b > 0$ models an interest rate in Stochastic Finance. The problem with this model is that r_t might be negative if $\sigma > 0$. If $\sigma = 0$, then one gets as one solution

$$r_t = r_0 e^{-bt} + \frac{a}{b}(1 - e^{-bt})$$

so that the meaning of a and b become more clear: the interest rate moves from its initial value r_0 to the value $\frac{a}{b}$ as $t \to \infty$ with a speed determined by the parameter a. By assuming $\sigma > 0$ one tries to add a random perturbation to this.

Both, the Ornstein-Uhlenbeck process and the process in the Vasicek interest rate model are Gaussian processes since the diffusion coefficient is not random. Moreover, the Vasicek interest rate model process is a generalization of the Ornstein-Uhlenbeck process.

The drawback of a negative interest rate in the Vasicek model can be removed by the following model:

Example 4.1.6 (Cox-Ingersoll-Ross Model)**.** For $a, b > 0$ and $\sigma \geq 0$ one proposes the SDE

$$dr_t = [a - br_t]dt + \sigma\sqrt{r_t}dB_t \quad \text{with} \quad r_0 > 0.$$

The difference to the Vasicek interest rate model is that the factor $\sqrt{r_t}$ is added in the diffusion part. This guarantees that the fluctuation is getting smaller if r_t is close to zero. In fact, the parameters can be adjusted such that the trajectories stay positive (which should be surprising).

Instead of considering the interest rate r_0 as initial condition one can take into the account the whole interest curve as anticipated by the market at time $t = 0$ as initial condition. This yields to a considerably more complicated model, the HEATH-JARROW-MORTON model.

Example 4.1.7 (HEATH-JARROW-MORTON model)**.** We assume that $f(s,t)$ stands for the instantaneous interest rate at time t as anticipated by the market at time s with $0 \leq s \leq t < \infty$. In particular, $r_t = f(t,t)$ is the interest rate at time t. Now one considers the equation

$$f(t,u) = f(0,u) + \int_0^t \alpha(v,u)dv + \int_0^t \sigma(f(v,u))dB_v$$

with $f(0,u) = \Phi(u)$.

4.2 Strong Uniqueness of SDEs

We shall start with a beautiful lemma, the GRONWALL lemma.

Lemma 4.2.1 (GRONWALL)**.** *Let $A, B, T \geq 0$ and $f : [0,T] \to [0,\infty)$ be a continuous function such that*

$$f(t) \leq A + B \int_0^t f(s)ds$$

for all $t \in [0,T]$. Then one has that $f(T) \leq Ae^{BT}$.

Proof. Letting $g(t) := e^{-Bt} \int_0^t f(s)ds$ we deduce

$$\begin{aligned} g'(t) &= -Be^{-Bt} \int_0^t f(s)ds + e^{-Bt} f(t) \\ &= e^{-Bt} \left(f(t) - B \int_0^t f(s)ds \right) \leq Ae^{-Bt} \end{aligned}$$

and

$$g(T) = \int_0^T g'(t)dt \leq A \int_0^T e^{-Bt}dt = \frac{A}{B}\left(1 - e^{-BT}\right).$$

Consequently,

$$f(T) \leq A + B\int_0^T f(t)dt = A + Be^{BT}g(T)$$
$$\leq A + Be^{BT}\frac{A}{B}\left(1 - e^{-BT}\right) = Ae^{BT}.$$

□

Theorem 4.2.2 (Strong uniqueness)**.** *Suppose that for all $n = 1, 2, ...$ there is a constant $c_n > 0$ such that*

$$|\sigma(t,x) - \sigma(t,y)| + |a(t,x) - a(t,y)| \leq c_n|x - y|$$

for $|x| \leq n$, $|y| \leq n$, and $t \geq 0$. Assume that $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are solutions of the SDE (4.1). Then

$$\mathbb{P}\left(X_t = Y_t, t \geq 0\right) = 1.$$

Proof. We use the stopping times

$$\sigma_n := \inf\left\{t \geq 0 : |X_t| \geq n\right\} \quad \text{and} \quad \tau_n := \inf\left\{t \geq 0 : |Y_t| \geq n\right\}$$

where we assume that $n > |x_0|$. Letting $\rho_n := \min\left\{\sigma_n, \tau_n\right\}$ we obtain, a.s., that

$$X_{t\wedge\rho_n} - Y_{t\wedge\rho_n}$$
$$= \int_0^{t\wedge\rho_n} \left[a(u,X_u) - a(u,Y_u)\right] du + \int_0^{t\wedge\rho_n} \left[\sigma(u,X_u) - \sigma(u,Y_u)\right] dB_u.$$

Hence

$$\begin{aligned}
\mathbb{E}\left|X_{t\wedge\rho_n} - Y_{t\wedge\rho_n}\right|^2 &\leq 2\mathbb{E}\left|\int_0^{t\wedge\rho_n} \left[a(u,X_u) - a(u,Y_u)\right] du\right|^2 \\
&\quad +2\mathbb{E}\left|\int_0^{t\wedge\rho_n} \left[\sigma(u,X_u) - \sigma(u,Y_u)\right] dB_u\right|^2 \\
&\leq 2t\mathbb{E}\int_0^{t\wedge\rho_n} \left|a(u,X_u) - a(u,Y_u)\right|^2 du
\end{aligned}$$

$$
\begin{aligned}
& +2\mathbb{E}\int_0^{t\wedge\rho_n}[\sigma(u,X_u)-\sigma(u,Y_u)]^2\,du \\
\leq\;& (2t+2)c_n^2\mathbb{E}\int_0^{t\wedge\rho_n}|X_u-Y_u|^2du \\
\leq\;& (2t+2)c_n^2\mathbb{E}\int_0^{t}|X_{u\wedge\rho_n}-Y_{u\wedge\rho_n}|^2du.
\end{aligned}
$$

Now fix $T>0$. The above computation gives

$$
\mathbb{E}\left|X_{t\wedge\rho_n}-Y_{t\wedge\rho_n}\right|^2\leq(2T+2)c_n^2\int_0^t\mathbb{E}\left|X_{u\wedge\rho_n}-Y_{u\wedge\rho_n}\right|^2du
$$

for $t\in[0,T]$. For

$$
f(t):=\mathbb{E}\left|X_{t\wedge\rho_n}-Y_{t\wedge\rho_n}\right|^2
$$

we may apply GRONWALL's lemma. The function f is continuous since for $t_k\to t$ one gets

$$
\begin{aligned}
\lim_k f(t_k) &= \lim_k\mathbb{E}\left|X_{t_k\wedge\rho_n}-Y_{t_k\wedge\rho_n}\right|^2 \\
&= \mathbb{E}\lim_k\left|X_{t_k\wedge\rho_n}-Y_{t_k\wedge\rho_n}\right|^2 \\
&= \mathbb{E}\left|X_{t\wedge\rho_n}-Y_{t\wedge\rho_n}\right|^2 \\
&= f(t)
\end{aligned}
$$

by dominated convergence as a consequence of (for example)

$$
\mathbb{E}\sup_{t\in[0,T]}\left|X_{t\wedge\rho_n}\right|^2\leq n^2
$$

and the continuity of the processes X and Y. Exploiting GRONWALL's lemma with $A:=0$ and $B:=(2T+2)c_n^2$ yields

$$
f(T)\leq Ae^{BT}=0\quad\text{and}\quad\mathbb{E}\left|X_{t\wedge\rho_n}-Y_{t\wedge\rho_n}\right|^2=0.
$$

Since

$$
\lim_n\rho_n=\infty
$$

because X and Y are continuous processes, we get by FATOU's lemma that

$$
\mathbb{E}\left|X_t-Y_t\right|^2=\mathbb{E}\liminf_n\left|X_{t\wedge\rho_n}-Y_{t\wedge\rho_n}\right|^2\leq\liminf_n\mathbb{E}\left|X_{t\wedge\rho_n}-Y_{t\wedge\rho_n}\right|^2=0.
$$

Hence $\mathbb{P}(X_t = Y_t) = 1$ and, by the continuity of X and Y,

$$\mathbb{P}(X_t = Y_t, t \geq 0) = 1.$$

$\square$

Sometimes the assumptions of the above criteria are too strong. There is a nice extension:

Theorem 4.2.3 (Yamada-Tanaka)**.** *Suppose that*

$$\sigma, a : [0, \infty) \times \mathbb{R} \to \mathbb{R}$$

are continuous such that

$$\begin{aligned} |\sigma(t,x) - \sigma(t,y)| &\leq h(|x-y|), \\ |a(t,x) - a(t,y)| &\leq K(|x-y|) \end{aligned}$$

for $x, y \in \mathbb{R}$, where $h : [0,\infty) \to [0,\infty)$ is strictly increasing with $h(0) = 0$ and $K : [0,\infty) \to \mathbb{R}$ is strictly increasing and concave with $K(0) = 0$, such that

$$\int_0^\varepsilon \frac{du}{K(u)} = \int_0^\varepsilon \frac{du}{h(u)^2} = \infty$$

for all $\varepsilon > 0$. Then any two solutions of (4.1) are indistinguishable.

Example 4.2.4. *One can take $h(x) := x^\alpha$ for $\alpha \geq \frac{1}{2}$.*

For $\alpha = 1/2$ we have in the Cox-Ingersoll-Ross model that $\sigma(t,x) = \sigma\sqrt{|x|}$. This implies that

$$|\sigma(t,x) - \sigma(t,y)| \leq \sigma|\sqrt{x} - \sqrt{y}| \leq \sigma\sqrt{|x-y|}.$$

However, there is also the following example:

Example 4.2.5. *Let $\sigma : \mathbb{R} \to [0,\infty)$ be continuous such that*

(i) $\sigma(x_0) = 0$,

(ii) $\int_{x_0-\varepsilon}^{x_0+\varepsilon} \frac{dx}{\sigma^2(x)} < \infty$ *and $\sigma(x) \geq 1$ if $|x - x_0| > \varepsilon$ for some $\varepsilon > 0$.*

Then the SDE

$$dX_t = \sigma(X_t) dB_t \quad \text{with} \quad X_0 = x_0$$

has infinitely many solutions.

4.3 Existence of strong solutions of SDEs

First we state the Burkholder-Davis-Gundy inequalities:

Theorem 4.3.1 (Burkholder-Davis-Gundy inequalities)**.** *For any $0 < p < \infty$ there exist constants $\alpha_p, \beta_p > 0$ such that, for $L \in \mathcal{L}_2^{\mathrm{loc}}$, one has that*

$$\beta_p \left\| \sqrt{\int_0^T L_t^2 dt} \right\|_p \leq \left\| \sup_{t\in[0,T]} \left| \int_0^t L_s dB_s \right| \right\|_p \leq \alpha_p \left\| \sqrt{\int_0^T L_t^2 dt} \right\|_p . \tag{4.2}$$

Moreover, one has $\alpha_p \leq c\sqrt{p}$ for $p \in [2,\infty)$ for some absolute constant $c > 0$.

The result we want to prove in this section is

Theorem 4.3.2. *Suppose that $\sigma, a : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ are continuous such that*

$$|\sigma(t,x) - \sigma(t,y)| + |a(t,x) - a(t,y)| \leq K|x-y|$$

for all $x, y \in \mathbb{R}$ and some $K > 0$. Then there exists a solution $X = (X_t)_{t\geq 0}$ to the SDE (4.1). Moreover, for $2 \leq p < \infty$ we have

$$\left\| \sup_{t\in[0,T]} |X_t| \right\|_p \leq \sqrt{2}[|x_0| + 1] e^{K_T^2 T[\alpha_p + \sqrt{T}]^2}$$

for $T > 0$, $\alpha_p > 0$ being the constant from the Burkholder-Davis-Gundy inequalities, and $K_T := \max\{K, \sup_{t\in[0,T]}[|\sigma(t,0)| + |a(t,0)|]\}$.

Proof. (a) Let $p \geq 2$ and $T > 0$. We define the space $L_p^{C[0,T]}(\Omega, \mathcal{F}, \mathbb{P})$ to be the linear space of all $f : \Omega \to C[0,T]$ such that $f_t : \Omega \to \mathbb{R}$ is measurable for all $t \in [0,T]$ and such that

$$\|f\|_{L_p^{C[0,T]}} := \left(\int_\Omega ||f(\omega)||_{C[0,T]}^p d\mathbb{P}(\omega) \right)^{1/p} < \infty$$

with $\|g\|_{C[0,T]} := \sup_{t\in[0,T]} |g_t|$ for $g \in C[0,T]$. We obtain a complete normed space, i.e. a *Banach space*. We define the sequence of processes $X^{(k)} = (X_t^{(k)})_{t\geq 0}$ by

$$X_t^{(0)} :\equiv x_0,$$

$$X_t^{(k+1)} \quad := \quad x_0 + \int_0^t \sigma(u, X_u^{(k)}) dB_u + \int_0^t a(u, X_u^{(k)}) du, \quad k \geq 0.$$

(b) $X^{(k)} \in L_p^{C[0,T]}(\Omega, \mathcal{F}, \mathbb{P})$ for all $k \in \mathbb{N}$: We check this by induction, where for and $X_t^{(0)} = x_0$ this is evident. So we assume $X^{(0)}, \ldots, X^{(k)} \in L_p^{C[0,T]}(\Omega, \mathcal{F}, \mathbb{P})$ and decompose

$$\begin{aligned} X_t^{(k+1)} - X_t^{(k)} &= \int_0^t \left[\sigma(u, X_u^{(k)}) - \sigma(u, X_u^{(k-1)})\right] dB_u \\ &\quad + \int_0^t \left[a(u, X_u^{(k)}) - a(u, X_u^{(k-1)})\right] du \\ &=: M_t + C_t. \end{aligned}$$

Now

$$\mathbb{E} \sup_{t \in [0,T]} |C_t|^p \leq T^{p-1} K^p \int_0^T \mathbb{E} \left|X_u^{(k)} - X_u^{(k-1)}\right|^p du$$

by Hölder's inequality and

$$\begin{aligned} \mathbb{E} \sup_{t \in [0,T]} |M_t|^p &\leq \alpha_p^p \mathbb{E} \left| \int_0^T \left|\sigma(u, X_u^{(k)}) - \sigma(u, X_u^{(k-1)})\right|^2 du \right|^{p/2} \\ &\leq \alpha_p^p K^p T^{\frac{p-2}{2}} \int_0^T \mathbb{E} \left|X_u^{(k)} - X_u^{(k-1)}\right|^p du \end{aligned}$$

by the Burkholder-Davis-Gundy inequality with constant $\alpha_p > 0$ and Hölder's inequality. Consequently,

$$\begin{aligned} \left(\mathbb{E} \sup_{t \in [0,T]} \left| X_t^{(k+1)} - X_t^{(k)} \right|^p \right)^{1/p} &\leq \left(\mathbb{E} \sup_{t \in [0,T]} |M_t|^p \right)^{1/p} + \left(\mathbb{E} \sup_{t \in [0,T]} |C_t|^p \right)^{1/p} \\ &\leq D_p \left(\int_0^T \mathbb{E} \left|X_u^{(k)} - X_u^{(k-1)}\right|^p du \right)^{1/p} \end{aligned}$$

with $D_p := K[T^{\frac{p-1}{p}} + \alpha_p T^{\frac{p-2}{2p}}]$ where we used the triangle-inequality in $L_p^{C[0,T]}(\Omega, \mathcal{F}, \mathbb{P})$. Since by assumption $X^{(k)}, X^{(k-1)} \in L_p^{C[0,T]}(\Omega, \mathcal{F}, \mathbb{P})$, we conclude that the right hand side of the estimate is finite. Because of the triangle-inequality,

$$||X^{(k+1)}||_{L_p^{C[0,T]}} \leq ||X^{(k+1)} - X^{(k)}||_{L_p^{C[0,T]}} + ||X^{(k)}||_{L_p^{C[0,T]}}$$

this implies $X^{(k+1)} \in L_p^{C[0,T]}(\Omega, \mathcal{F}, \mathbb{P})$.

(c) Iterating step (b) yields to

$$\left(\mathbb{E} \sup_{t \in [0,T]} \left| X_t^{(k+1)} - X_t^{(k)} \right|^p \right) \leq \frac{(TD_p^p)^k}{k!} \sup_{t \in [0,T]} \mathbb{E} \left| X_t^{(1)} - x_0 \right|^p =: \frac{(TD_p^p)^k}{k!} c_{x_0,T,p}^p.$$

Therefore, $(X^{(k)})_{k=1}^\infty$ is a Cauchy-sequence in $L_p^{C[0,T]}(\Omega, \mathcal{F}, \mathbb{P})$ as for $0 \leq k < l < \infty$ we have

$$\begin{aligned} & \left\| X^{(l)} - X^{(k)} \right\|_{L_p^{C[0,T]}} \\ \leq & \left\| X^{(l)} - X^{(l-1)} \right\|_{L_p^{C[0,T]}} + \cdots + \left\| X^{(k+1)} - X^{(k)} \right\|_{L_p^{C[0,T]}} \\ \leq & \; c_{x_0,T,p} \left(\frac{(TD_p^p)^{l-1/p}}{((l-1)!)^{1/p}} + \cdots + \frac{(TD_p^p)^{k/p}}{(k!)^{1/p}} \right) \\ \leq & \; c_{x_0,T,p} \sum_{i=k}^\infty \frac{(TD_p^p)^{i/p}}{(i!)^{1/p}} \to_k 0. \end{aligned}$$

Therefore there is a limit X in $L_p^{C[0,T]}(\Omega, \mathcal{F}, \mathbb{P})$ and

$$\mathbb{E} \sup_{t \in [0,T]} |X_t|^p < \infty.$$

We show that this is our solution. Re-using the computation from step (b) we get

$$\begin{aligned} & \left\| X_t - x_0 - \int_0^t \sigma(u, X_u) dB_u - \int_0^t a(u, X_u) du \right\|_{L_p} \\ & = \Bigg\| \left[X_t - x_0 - \int_0^t \sigma(u, X_u) dB_u - \int_0^t a(u, X_u) du \right] \\ & \quad - \left[X_t^{(k+1)} - x_0 - \int_0^t \sigma(u, X_u^{(k)}) dB_u - \int_0^t a(u, X_u^{(k)}) du \right] \Bigg\|_{L_p} \\ & \leq \| X_t - X_t^{(k+1)} \|_{L_p} + L_p \left(\int_0^T \mathbb{E} | X_u - X_u^{(k)} |^p du \right)^{\frac{1}{p}} \\ & \to 0 \end{aligned}$$

as $k \to \infty$. Hence, a.s.,

$$X_t = x_0 - \int_0^t \sigma(u, X_u) dB_u + \int_0^t a(u, X_u) du.$$

(d) By the uniqueness argument for the strong solutions we also get that

$$\mathbb{P}(X_t^{(T_1)} = X_t^{(T_2)}) = 1$$

for $t \in [0, \min\{T_1, T_2\}]$ when $X^T = (X_t^T)_{t \in [0,T]}$ is the solution constructed on $[0, T]$. Hence we may find a continuous and adapted process $X = (X_t)_{t \geq 0}$ such that

$$\mathbb{P}(X_t = X_t^{(n)}) = 1 \quad \text{for all} \quad t \in [0, n].$$

(e) Now we consider the solved equation and $2 \leq p < \infty$. Similarly as before we deduce, for $t \in [0, T]$, that

$$\begin{aligned}
& \left\| \sup_{s \in [0,t]} |X_s| \right\|_p \\
\leq \;& |x_0| + \alpha_p \left(\mathbb{E} \left| \int_0^t |\sigma(u, X_u)|^2 \, du \right|^{p/2} \right)^{\frac{1}{p}} + \left(\mathbb{E} \left| \int_0^t |a(u, X_u)| \, du \right|^p \right)^{\frac{1}{p}} \\
\leq \;& |x_0| + \alpha_p K_T \left(\mathbb{E} \left| \int_0^t [1 + |X_u|]^2 du \right|^{p/2} \right)^{\frac{1}{p}} + \\
& K_T \left(\mathbb{E} \left| \int_0^t [1 + |X_u|] du \right|^p \right)^{\frac{1}{p}} \\
\leq \;& |x_0| + \alpha_p K_T \left| \int_0^t \| 1 + |X_u| \|_p^2 du \right|^{1/2} + K_T \int_0^t \| 1 + |X_u| \|_p du \\
\leq \;& |x_0| + K_T [\alpha_p + \sqrt{T}] \left| \int_0^t \| 1 + |X_u| \|_p^2 du \right|^{1/2}.
\end{aligned}$$

Consequently,

$$\left\| 1 + \sup_{s \in [0,t]} |X_s| \right\|_p^2 \leq 2[|x_0| + 1]^2 + 2K_T^2 [\alpha_p + \sqrt{T}]^2 \int_0^t \left\| 1 + \sup_{s \in [0,u]} |X_s| \right\|_p^2 du$$

for $t \in [0,T]$ and Gronwall's lemma gives

$$\left\| 1 + \sup_{s\in[0,T]} |X_s| \right\|_p^2 \le 2[|x_0|+1]^2 e^{2K_T^2 T[\alpha_p+\sqrt{T}]^2}.$$

□

Remark 4.3.3. (1) From the above proof it follows that we obtain a Gaussian process in case of $\sigma(t,x) = \sigma(t)$ and $a(t,x) = a_1(t)x + a_2(t)$ with σ, a_1, a_2 continuous and bounded as the approximating processes $X^{(k)}$ are Gaussian and the L_2-limit of Gaussian random variables is Gaussian as well.

(2) From the assumptions of Theorem 4.3.2 we get that

$$|\sigma(t,x)| + |a(t,x)| \le \sup_{t\in[0,T]} [|\sigma(t,0)| + |a(t,0)|] + K|x|$$

which is a standard growth condition that is satisfied in our context automatically.

4.4 Lévy's characterization of the Brownian motion and the Girsanov theorem

The first theorem, Lévy's characterization of the Brownian motion, is a characterization of the Brownian motion by the quadratic variation. To introduce the quadratic variation we need

Proposition 4.4.1. *Let $M = (M_t)_{t\ge 0} \in \mathcal{M}_{\rm loc}^{c,0}$ be a continuous local martingale starting in zero. Then there exists a continuous and adapted process $\langle M \rangle = (\langle M \rangle_t)_{t\ge 0}$, unique up to indistinguishability, such that*

(i) $0 = \langle M \rangle_0 \le \langle M \rangle_s \le \langle M \rangle_t$ *for all* $0 \le s \le t < \infty$,

(ii) $\lim_n \left[\sum_{i=1}^n \left(M_{t_i^n} - M_{t_{i-1}^n} \right)^2 \right] = \langle M \rangle_t$ *in probability for all* $0 = t_0^n \le \cdots \le t_n^n = t$ *such that* $\lim_n \sup_{i=1,\dots,n} |t_i^n - t_{i-1}^n| = 0$.

Definition 4.4.2. The process $\langle M \rangle$ is called *quadratic variation* of the local martingale M.

Proposition 4.4.3. *For $L = (L_u)_{u \ge 0} \in \mathcal{L}_2^{\mathrm{loc}}$ one has that*

$$\left\langle \int_0^{\cdot} L_u dB_u \right\rangle_t = \int_0^t L_u^2 du \quad \textit{for} \quad t \ge 0 \textit{ a.s.}$$

In particular, we have

Example 4.4.4. For the Brownian motion $B = (B_t)_{t \ge 0}$ one has that $\langle B \rangle_t = t$, $t \ge 0$, a.s.

The converse is true as well:

Proposition 4.4.5 (P. Lévy)**.** *Let $M = (M_t)_{t \ge 0}$ be a continuous adapted process such that $M_0 \equiv 0$. Then the following assertions are equivalent:*

(i) *M is an $(\mathcal{F}_t)_{t \ge 0}$-Brownian motion.*

(ii) *$M \in \mathcal{M}_{\mathrm{loc}}^{c,0}$ and $\langle M \rangle_t = t$, $t \ge 0$, a.s.*

Proof. We only have to show that (ii) implies (i). Let

$$\tau_N(\omega) := \inf\{t \ge 0 : |M_t(\omega)| = N\},$$

and let σ_k be an increasing sequence of stopping times, with $\lim_k \sigma_k(\omega) = \infty$, such that $M^{\sigma_k} := M_t \chi_{\{t < \sigma_k\}} + M_{\sigma_k} \chi_{\{\sigma_k \le t\}} = M_{t \wedge \sigma_k}$ is a martingale for all $k \in \mathbb{N}$. Now

$$\mathbb{E}|M_t^{\tau_N}|^2 \le N^2 < \infty \text{ for all } N = 0, 1, 2, ..., \text{ and } t \ge 0,$$

and M^{τ_N} is $(\mathcal{F}_t)_{t \ge 0}$-adapted (see [11, Proposition 2.18 in Chapter 1]). Moreover, for $0 \le s \le t < \infty$ (we use $\mathbb{E}[\mathbb{E}[X|\mathcal{F}_\tau|\mathcal{F}_\sigma] = \mathbb{E}[X|\mathcal{F}_{\tau \wedge \sigma}]$ which holds for any integrable X and stopping times τ, σ):

$$\begin{aligned}
\mathbb{E}\left(M_{t \wedge \tau_N} | \mathcal{F}_s\right) &= \mathbb{E}\left(\mathbb{E}\left(M_{t \wedge \tau_N} | \mathcal{F}_{t \wedge \tau_N}\right) | \mathcal{F}_s\right) \\
&= \mathbb{E}\left(M_{t \wedge \tau_N} | \mathcal{F}_{s \wedge \tau_N}\right) \\
&= \mathbb{E}\left(\lim_{k \to \infty} M_{t \wedge \sigma_k \wedge \tau_N} | \mathcal{F}_{s \wedge \tau_N}\right)
\end{aligned}$$

$$\begin{aligned}
&= \lim_{k\to\infty} \mathbb{E}\left(M_{t\wedge\sigma_k\wedge\tau_N}|\mathcal{F}_{s\wedge\tau_N}\right) \\
&= \lim_{k\to\infty} M_{s\wedge\sigma_k\wedge\tau_N} \\
&= M_{s\wedge\tau_N},
\end{aligned}$$

where we have used dominated convergence and the optional stopping theorem (sometimes called the optional sampling theorem) (note that $s\wedge\tau_N \leq t\wedge\tau_N \leq t$, for all $k \geq k_0$). Hence $M^{\tau_N} \in \mathcal{M}_2^{c,0}$. Let

$$f(x) := e^{i\lambda x} \quad \text{for some} \quad \lambda \in \mathbb{R}.$$

By Itô's formula for martingales (used in the complex setting which follows directly from the real setting), which states that for a C^2 function g and a continuous local martingale (X_t)

$$g(X_t) = g(X_0) + \int_0^t g'(X_u)dX_u + \frac{1}{2}\int_0^t g''(X_u)d\langle X\rangle_u$$

one has

$$\begin{aligned}
e^{i\lambda(M_t^{\tau_N}-M_s^{\tau_N})}\chi_A &= \chi_A + \chi_A\int_s^t i\lambda e^{i\lambda(M_u^{\tau_N}-M_s^{\tau_N})}dM_u^{\tau_N} \\
&\quad -\frac{\lambda^2}{2}\chi_A\int_{s\wedge\tau_N}^{t\wedge\tau_N} e^{i\lambda(M_u^{\tau_N}-M_s^{\tau_N})}d\langle M^{\tau_N}\rangle_u
\end{aligned}$$

for all $A \in \mathcal{F}_s$. Using $\langle M^{\tau_N}\rangle_u = \langle M\rangle_u = u$ for $s\wedge\tau_N \leq u \leq t\wedge\tau_N$ and taking the expected value implies that

$$\mathbb{E}e^{i\lambda(M_t^{\tau_N}-M_s^{\tau_N})}\chi_A = \mathbb{P}(A) - \frac{\lambda^2}{2}\int_{s\wedge\tau_N}^{t\wedge\tau_N}\mathbb{E}e^{i\lambda(M_u^{\tau_N}-M_s^{\tau_N})}\chi_A du.$$

As $N\to\infty$ we get by dominated converge

$$\mathbb{E}e^{i\lambda(M_t-M_s)}\chi_A = \mathbb{P}(A) - \frac{\lambda^2}{2}\int_s^t \mathbb{E}e^{i\lambda(M_u-M_s)}\chi_A du.$$

Letting now

$$H(u) := \mathbb{E}e^{i\lambda(M_u-M_s)}\chi_A$$

yields to a continuous and bounded function. Moreover, we have

$$H(t) = \mathbb{P}(A) - \frac{\lambda^2}{2}\int_s^t H(u)du$$

which implies

$$H(t) = \mathbb{P}(A) e^{-\frac{\lambda^2}{2}(t-s)}.$$

This can be seen by expanding the right-hand side of the integral equation successively to a series expansion. We conclude from

$$\mathbb{E} e^{i\lambda(M_u - M_s)} \chi_A = \mathbb{P}(A) e^{-\frac{\lambda^2}{2}(t-s)}$$

that $M_t - M_s$ is independent from $\mathcal{F}_s$ and that $M_t - M_s \sim N(0, t-s)$. $\square$

Now we turn to our second fundamental theorem:

Proposition 4.4.6 (Girsanov)**.** *Let $L = (L_t)_{t\geq 0} \in \mathcal{L}_2$ and assume that the process $(\mathcal{E}_t)_{t\geq 0}$ defined by*

$$\mathcal{E}_t := \exp\left(-\int_0^t L_u dB_u - \frac{1}{2}\int_0^t L_u^2 du \right)$$

is a martingale. Let $T > 0$ and

$$dQ_T := \mathcal{E}_T d\mathbb{P}.$$

Then $(W_t)_{t\geq 0}$ with

$$W_t := B_t + \int_0^t \chi_{[0,T]}(u) L_u du$$

defines a Brownian motion $(W_t)_{t\geq 0}$ with respect to $(\Omega, \mathcal{F}, Q_T, (\mathcal{F}_t)_{t\geq 0})$.

Lemma 4.4.7. *Let $0 \leq t \leq T < \infty$.*

(i) *The measures Q_t and Q_T coincide on $\mathcal{F}_t$.*

(ii) *Assume that $Z : \Omega \to \mathbb{R}$ is $\mathcal{F}_T$-measurable such that $\mathbb{E}_{Q_T}|Z| < \infty$. Then*

$$\mathbb{E}_{Q_T}(Z|\mathcal{F}_t) = \frac{\mathbb{E}(Z\mathcal{E}_T|\mathcal{F}_t)}{\mathcal{E}_t} \quad a.s.$$

Proof. (i) For $B \in \mathcal{F}_t$ one has

$$Q_T(B) = \int_B \mathcal{E}_T d\mathbb{P} = \int_B \mathbb{E}(\mathcal{E}_T|\mathcal{F}_t) d\mathbb{P} = \int_B \mathcal{E}_t d\mathbb{P} = Q_t(B)$$

where we have used that $(\mathcal{E}_t)_{t\geq 0}$ is a martingale.

(ii) We show the assertion only for $0 \leq Z \leq c$ where all terms are defined. The general case follows from the decomposition $Z = Z^+ - Z^-$ and dominated convergence. We will show that

$$\int_B \mathcal{E}_t \mathbb{E}_{Q_T}(Z|\mathcal{F}_t) d\mathbb{P} = \int_B Z \mathcal{E}_T d\mathbb{P}$$

which follows from

$$\begin{aligned}
\int_B \mathcal{E}_t \mathbb{E}_{Q_T}(Z|\mathcal{F}_t) d\mathbb{P} &= \int_B \mathbb{E}_{Q_T}(Z|\mathcal{F}_t) dQ_t \\
&= \int_B \mathbb{E}_{Q_T}(Z|\mathcal{F}_t) dQ_T \\
&= \int_B Z dQ_T \\
&= \int_B Z \mathcal{E}_T d\mathbb{P}.
\end{aligned}$$

□

Proof of Proposition 4.4.6. We only prove the statement for *bounded* $L \in \mathcal{L}_2$.

(a) We assume that $K \in \mathcal{L}_2$ is deterministic and bounded. We will show that

$$\mathbb{E}_{Q_T} e^{\int_0^T iK_u dW_u} = e^{-\frac{1}{2}\int_0^T K_u^2 du}. \tag{4.3}$$

It holds

$$\begin{aligned}
\mathbb{E}_{Q_T} e^{\int_0^T iK_u dW_u} &= \mathbb{E} e^{-\int_0^T L_u dB_u - \frac{1}{2}\int_0^T L_u^2 du} e^{\int_0^T iK_u dW_u} \\
&= \mathbb{E} e^{-\int_0^T L_u dB_u - \frac{1}{2}\int_0^T L_u^2 du} e^{\int_0^T iK_u dB_u + \int_0^T iK_u L_u du} \\
&= \mathbb{E} e^{\int_0^T (iK_u - L_u) dB_u - \frac{1}{2}\int_0^T (iK_u - L_u)^2 du} e^{-\frac{1}{2}\int_0^T K_u^2 du} \\
&= e^{-\frac{1}{2}\int_0^T K_u^2 du} \mathbb{E} e^{\int_0^T (iK_u - L_u) dB_u - \frac{1}{2}\int_0^T (iK_u - L_u)^2 du}.
\end{aligned}$$

We realise that M given by $M_t = e^{\int_0^t (iK_u - L_u)dB_u - \frac{1}{2}\int_0^t (iK_u - L_u)^2 du}$ is a complex-valued martingale with $M_0 = 1$ so that $\mathbb{E}M_T = 1$. This implies (4.3).

(b) Relation (4.3) holds especially for $K_u = \sum_{k=1}^n x_k \chi_{(t_{k-1}, t_k]}(u)$ with $x_k \in \mathbb{R}$ and $0 < t_0 < t_1 < ... < t_n < T$. Then

$$\begin{aligned}\mathbb{E}_{Q_T} e^{\int_0^T iK_u dW_u} &= \mathbb{E}_{Q_T} e^{i\sum_{k=1}^n x_k (W_{t_k} - W_{t_{k-1}})} \\ &= e^{-\frac{1}{2}\sum_{k=1}^n (x_k)^2 (t_k - t_{k-1})}.\end{aligned}$$

But this means that w.r.t. Q_T the vector $W_{t_n} - W_{t_{n-1}}, ..., W_{t_1} - W_{t_0}$ consists of independent normal distributed random variables with $\mathbb{E}(W_{t_k} - W_{t_{k-1}})^2 = t_k - t_{k-1}$. Hence W is a Brownian motion w.r.t. Q_T.

(c) In order to see that W is a Brownian motion w.r.t. $(\Omega, \mathcal{F}, Q_T, (\mathcal{F}_t)_{t\geq 0})$ we first notice that W is $(\mathcal{F}_t)_{t\geq 0}$-adapted. To show that $W_t - W_s$ is independent from $\mathcal{F}_s$ w.r.t. Q_T we write for $x, y \in \mathbb{R}$ and $A \in \mathcal{F}_s$

$$\begin{aligned}\mathbb{E}_{Q_T} e^{ix(W_t - W_s)}\chi_A &= \mathbb{E}_{Q_T}\mathbb{E}_{Q_T}[e^{ix(W_t - W_s)}\chi_A \mid \mathcal{F}_s] \\ &= \mathbb{E}_{Q_T}\chi_A \mathbb{E}_{Q_T}[e^{ix(W_t - W_s)} \mid \mathcal{F}_s] \\ &= \mathbb{E}_{Q_T}\chi_A \mathbb{E}[e^{ix(W_t - W_s)} e^{-\int_s^T L_u dB_u - \frac{1}{2}\int_s^T L_u^2 du} \mid \mathcal{F}_s].\end{aligned}$$

where we used Lemma 4.4.7 for the last line. Using the tower property we get similarly to (a) that

$$\begin{aligned}&\mathbb{E}[e^{ix(W_t - W_s)} e^{-\int_s^T L_u dB_u - \frac{1}{2}\int_s^T L_u^2 du} \mid \mathcal{F}_s] \\ &\quad= \mathbb{E}[e^{ix(W_t - W_s)} \mathbb{E}[e^{-\int_s^T L_u dB_u - \frac{1}{2}\int_s^T L_u^2 du} \mid \mathcal{F}_t] \mid \mathcal{F}_s] \\ &\quad= \mathbb{E}[e^{ix(W_t - W_s)} e^{-\int_s^t L_u dB_u - \frac{1}{2}\int_s^t L_u^2 du} \mid \mathcal{F}_s] \\ &\quad= e^{-\frac{x^2}{2}(t-s)} \mathbb{E}\left[\frac{M_T}{M_s} \mid \mathcal{F}_s\right]\end{aligned}$$

with $M_t = e^{\int_0^t (ix - L_u)dB_u - \frac{1}{2}\int_0^t (ix - L_u)^2 du}$ and $\mathbb{E}\left[\frac{M_T}{M_s} \mid \mathcal{F}_s\right] = \frac{1}{M_s}\mathbb{E}[M_T \mid \mathcal{F}_s] = 1$. Consequently,

$$\mathbb{E}_{Q_T} e^{ix(W_t - W_s)}\chi_A = \mathbb{E}_{Q_T}\chi_A e^{-\frac{x^2}{2}(t-s)}.$$

□

An important condition to decide whether $(\mathcal{E}_t)_{t\geq 0}$ is a martingale is NOVIKOV's condition.

Proposition 4.4.8. *Assume that $M = (M_t)_{t\geq 0}$ is a continuous local martingale with $M_0 \equiv 0$ and $T > 0$ such that*

$$\mathbb{E}e^{\frac{1}{2}\langle M\rangle_T} < \infty.$$

Then $\mathcal{E} = (\mathcal{E}_{t\wedge T})_{t\geq 0}$ with

$$\mathcal{E}_t := e^{M_t - \frac{1}{2}\langle M\rangle_t}$$

is a martingale.

4.5 Solving SDEs by a transformation of drift

Now we explain how to solve a SDEs by a transformation of drift.

Proposition 4.5.1 (Transformation of drift)**.** *Let $\sigma, a : [0,\infty)\times\mathbb{R}\to\mathbb{R}$ be continuous such that*

$$\begin{aligned} |\sigma(t,x)-\sigma(t,y)| + |a(t,x)-a(t,y)| &\leq K|x-y| \\ |\sigma(t,x)| + |a(t,x)| &\leq K(1+|x|) \end{aligned}$$

for all $x,y\in\mathbb{R}$ and $t\geq 0$. Let $X = (X_t)_{t\geq 0}$ be the unique strong solution of

$$dX_t = \sigma(t,X_t)dB_t + a(t,X_t)dt$$

with $X_0 \equiv x_0 \in \mathbb{R}$. Let $T > 0$ and $L = (L_t)_{t\geq 0} \in \mathcal{L}_2$ be continuous such that

$$\mathbb{E}e^{\frac{1}{2}\int_0^T L_u^2 du} < \infty$$

and let

$$W_t := B_t + \int_0^t L_u du$$

for $t\in[0,T]$. Then, under Q_T with

$$dQ_T := \mathcal{E}_T d\mathbb{P} \quad \textit{where} \quad \mathcal{E}_t := e^{-\int_0^t L_u dB_u - \frac{1}{2}\int_0^t L_u^2 du},$$

one has that X solves

$$dX_t = \sigma(t,X_t)dW_t + \left[a(t,X_t) - \sigma(t,X_t)L_t\right]dt \quad \textit{for} \quad t\in[0,T].$$

What is the philosophy in this case? We wish to solve

$$dX_t = \sigma(t, X_t)dW_t + [a(t, X_t) - \sigma(t, X_t)L_t]\, dt \quad \text{for} \quad t \in [0, T].$$

For this purpose we construct a specific Brownian motion $W = (W_t)_{t\in[0,T]}$ on an appropriate stochastic basis $(\Omega, \mathcal{F}, Q_T; (\mathcal{F}_t)_{t\in[0,T]})$ so that this problem has the solution $X = (X_t)_{t\in[0,T]}$ which is called *weak solution*.

Proof of Proposition 4.5.1. By Theorems 4.2.2 and 4.3.2 there is a unique strong solution $X = (X_t)_{t\geq 0}$. Setting

$$M_t := \int_0^t (-L_u)dB_u$$

we get that $(\mathcal{E}_t)_{t\geq 0}$ is a martingale by NOVIKOV's condition (Proposition 4.4.8). The GIRSANOV Theorem (Proposition 4.4.6) gives that $(W_t)_{t\in[0,T]}$ is a Brownian motion with respect to Q_T. And finally (and also a bit formally)

$$\begin{aligned} dX_t &= \sigma(t, X_t)dB_t + a(t, X_t)dt \\ &= \sigma(t, X_t)(dB_t + L_t dt) - \sigma(t, X_t)L_t dt + a(t, X_t)dt \\ &= \sigma(t, X_t)dW_t + (a(t, X_t) - \sigma(t, X_t)L_t)dt. \end{aligned}$$

□

Example 4.5.2. Let $\sigma(t, x) = x$, $a \equiv 0$, $x_0 = 1$, $S_t = e^{B_t - \frac{t}{2}}$, and

$$\mathbb{E}e^{\frac{1}{2}\int_0^T L_u^2 du} < \infty.$$

Then

$$dS_t = S_t dW_t - S_t L_t dt, \quad t \in [0, T], \quad \text{under} \quad Q_T.$$

4.6 Weak solutions

In this section we indicate the principle of weak solutions: we do not start with a stochastic basis, but we construct a particular basis to our problem. The formal definition is as follows:

Definition 4.6.1. Assume that $\sigma, a : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ are measurable. A *weak solution* of

$$dX_t = \sigma(t, X_t)dB_t + a(t, X_t)dt \quad \text{with} \quad X_0 \equiv x_0$$

is a pair $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$, $(X_t, W_t)_{t\geq 0}$, such that

(i) $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ satisfies the usual conditions, i.e.

- $(\Omega, \mathcal{F}, \mathbb{P})$ is complete,
- all null-sets of $\mathcal{F}$ belong to $\mathcal{F}_0$,
- the filtration is right-continuous, i.e. $\mathcal{F}_t = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}$,

(ii) X is continuous and $(\mathcal{F}_t)_{t\geq 0}$ adapted,

(iii) $(W_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion,

(iv) $X_t = x_0 + \int_0^t \sigma(u, X_u)dW_u + \int_0^t a(u, X_u)du$, $t \geq 0$, a.s.,

where $(\sigma(u, X_u))_{u\geq 0} \in \mathcal{L}_2^{\text{loc}}$ and $\int_0^t |a(u, X_u(\omega))|du < \infty$ for all $\omega \in \Omega$ and all $t \geq 0$.

Example 4.6.2 (Tanaka)**.** Assume the SDE

$$dX_t = \text{sign}(X_t)dB_t \quad \text{with} \quad X_0 = 0$$

where $\text{sign}(x) = 1$ if $x > 0$ and $\text{sign}(x) = -1$ if $x \leq 0$.

(a) Non-uniqueness of the solution: Assume that $(X_t)_{t\geq 0}$ is a solution. Because

$$\int_0^t \text{sign}(X_s)^2 ds = t$$

Lévy's theorem applies and $(X_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$- Brownian motion. Then we also get that

$$(-X_t) = \int_0^t [-\text{sign}(X_s)]dB_s.$$

Because

$$\mathbb{E}\int_0^\infty |\text{sign}(-X_s) + \text{sign}(X_s)|^2 ds = 0,$$

we also have that

$$(-X_t) = \int_0^t [\text{sign}(-X_s)]dB_s,$$

so that $(-X_t)_{t\geq 0}$ is a solution as well and uniqueness fails.

(b) Existence of a solution: Define

$$M_t := \int_0^t \text{sign}(B_s)dB_s.$$

Again, by Lévy's theorem, $(M_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion. We verify that we have that

$$B_t = \int_0^t \text{sign}(B_s)dM_s, t \geq 0, a.s.$$

We find simple $(\lambda_s^N)_{s\in[0,t]}$ such that

$$\mathbb{E}\int_0^t |\lambda_s^N - \text{sign}(B_s)|^2 ds \to 0$$

as $N \to \infty$. Then

$$\mathbb{E}\int_0^t |\lambda_s^N \text{sign}(B_s) - 1|^2 ds \to 0$$

and

$$\int_0^t [\lambda_s^N \text{sign}(B_s)]dB_s \to B_t$$

in L_2 as $N \to \infty$. On the other hand,

$$\int_0^t \lambda_s^N dM_s \to \int_0^t \text{sign}(B_s)dM_s$$

in L_2 as $N \to \infty$. Finally, by a direct computation we see that

$$\int_0^t \lambda_s^N dM_s = \int_0^t [\lambda_s^N \text{sign}(B_s)]dB_s$$

as for $\lambda_t^N := \sum_{k=1}^N v_{k-1}^N 1_{(t_{k-1}^N, t_k^N]}(t)$ we get, a.s.,

$$\int_0^t \lambda_s^N dM_s = \sum_{k=1}^N v_{k-1}^N [M_{t_k^N} - M_{t_{k-1}^N}]$$

$$
\begin{aligned}
&= \sum_{k=1}^{N} v_{k-1}^N \int_{t_{k-1}^N}^{t_k^N} \mathrm{sign}(B_s) dB_s \\
&= \sum_{k=1}^{N} \int_{t_{k-1}^N}^{t_k^N} v_{k-1}^N \mathrm{sign}(B_s) dB_s \\
&= \int_0^t [\lambda_s^N \mathrm{sign}(B_s)] dB_s.
\end{aligned}
$$

To explain the usage of weak solutions we also introduce the notion of *path-wise uniqueness*:

Definition 4.6.3. The SDE

$$dX_t = \sigma(t, X_t) dB_t + a(t, X_t) dt \quad \text{with} \quad X_0 \equiv x_0,$$

$t \geq 0$, a.s., $X_0 \equiv x_0$, satisfies the *path-wise uniqueness* if any two solutions with respect to the same stochastic basis and Brownian motion are indistinguishable.

The application of this concept consists in

Proposition 4.6.4 (Yamada and Watanabe)**.** *The existence of weak solutions together with the path-wise uniqueness implies the existence of strong solutions.*

For more information see [11, Section 5.3].

4.7 The Cox-Ingersoll-Ross SDE

Now we consider an example that does not fall into the setting we discussed so far. Instead, we have some kind of *boundary problem*. Formally, we consider the Cox-Ingersoll-Ross SDE

$$dX_t = (a - bX_t) dt + \sigma \sqrt{X_t} dB_t \quad \text{with} \quad X_0 \equiv x_0 > 0 \tag{4.4}$$

on $[0, \tau]$ where $a, \sigma > 0$, $b \in \mathbb{R}$, and

$$\tau(\omega) := \inf \{ t \geq 0 : X_t(\omega) = 0 \}.$$

First we make this formal equation precise:

Definition 4.7.1. *A stopping time $\tau : \Omega \to [0,\infty]$ and an adapted and continuous process $X = (X_t)_{t\geq 0}$ with $X_t(\omega) = X_{t\wedge\tau(\omega)}(\omega)$ for all $\omega \in \Omega$ is called solution of the* Cox-Ingersoll-Ross *SDE provided that*

$$X_{t\wedge\tau} = \int_0^{t\wedge\tau} [a - bX_s]ds + \sigma \int_0^{t\wedge\tau} \sqrt{X_s} dB_s$$

for $t \geq 0$ a.s. and $X_{\tau(\omega)}(\omega) = 0$ if $\tau(\omega) < \infty$ and $X_t(\omega) > 0$ for all $t \in [0, \tau(\omega))$.

Proposition 4.7.2. *There exists a unique solution to the SDE (4.4).*

Idea of the proof. We only give the idea of the construction of the process. Let $1/n \in (0, x_0)$ and find a Lipschitz function $\sigma_n : \mathbb{R} \to \mathbb{R}$ with $\sigma_n(x) = \sigma\sqrt{x}$ whenever $x \geq 1/n$. Then the SDE

$$dX_t^n = (a - bX_t^n)dt + \sigma_n(X_t^n)dB_t \quad \text{with} \quad X_0^n \equiv x_0 > 0 \tag{4.5}$$

has a unique strong solution. For an adapted and continues process X let

$$\tau_X^n := \inf\{t \geq 0 : X_t = 1/n\} \in [0,\infty].$$

By adapting the proof of our uniqueness theorem we can check that $(X^n_{t\wedge\tau^n_{X^n}})_{t\geq 0}$ and $(X^m_{t\wedge\tau^n_{X^m}})_{t\geq 0}$ are indistinguishable for $1 \leq n \leq m < \infty$. Let Ω_0 be a set of measure one such that on Ω_0 the trajectories of $\left(X^n_{t\wedge\tau^n_{X^n}}\right)_{t\geq 0}$ and $\left(X^m_{t\wedge\tau^n_{X^m}}\right)_{t\geq 0}$ coincide for $1 \leq n \leq m$. By construction we have that

$$\tau^n_{X^n}(\omega) \leq \tau^m_{X^m}(\omega) \quad \text{for} \quad \omega \in \Omega_0$$

and may set

$$\tau(\omega) := \begin{cases} \lim_{n\to\infty} \tau^n_{X^n}(\omega) & : \omega \in \Omega_0 \\ \infty & : \omega \notin \Omega_0 \end{cases}$$

and

$$X_t(\omega) := \begin{cases} \lim_{n\to\infty} X^n_t(\omega) & : \omega \in \Omega_0 \\ x_0 & : \omega \notin \Omega_0 \end{cases}.$$

$\square$

What we can do in more detail is to study the quantitative behavior of this equation.

Proposition 4.7.3. *One has the following:*

(i) *If* $a \geq \frac{\sigma^2}{2}$*, then* $\mathbb{P}(\tau = \infty) = 1$,

(ii) *if* $0 < a < \frac{\sigma^2}{2}$ *and* $b \geq 0$*, then* $\mathbb{P}(\tau = \infty) = 0$,

(iii) *if* $0 < a < \frac{\sigma^2}{2}$ *and* $b < 0$*, then* $\mathbb{P}(\tau = \infty) \in (0,1)$.

Proof. For $x, M > 0$ we let $(X_t^x)_{t \geq 0}$ be the solution of the Cox-Ingersoll-Ross SDE starting in $x \geq 0$ and

$$\tau_M^x(\omega) := \inf\{t \geq 0 : X_t^x(\omega) = M\}.$$

(a) Define the *scale function*

$$s(x) := \int_1^x e^{\frac{2by}{\sigma^2}} y^{-\frac{2a}{\sigma^2}} dy.$$

Then

$$\frac{\sigma^2}{2} x s''(x) + (a - bx)s'(x) = 0 \tag{4.6}$$

by a computation.

(b) Let $0 < \varepsilon < x < M$ and $\tau_{\varepsilon,M}^x := \tau_\varepsilon^x \wedge \tau_M^x$. By Itô's formula

$$\begin{aligned} s(X_{t\wedge\tau_{\varepsilon,M}^x}^x) &= s(x) + \int_0^{t\wedge\tau_{\varepsilon,M}^x} s'(X_s^x) dX_s^x + \frac{1}{2}\int_0^{t\wedge\tau_{\varepsilon,M}^x} s''(X_s^x)\sigma^2 X_s^x ds \\ &= s(x) + \int_0^{t\wedge\tau_{\varepsilon,M}^x} s'(X_s^x)\sigma\sqrt{X_s^x} dB_s \\ &\quad + \int_0^{t\wedge\tau_{\varepsilon,M}^x} \left[(a - bX_s^x)s'(X_s^x) + \frac{1}{2}s''(X_s^x)\sigma^2 X_s^x\right] ds \\ &= s(x) + \int_0^{t\wedge\tau_{\varepsilon,M}^x} s'(X_s^x)\sigma\sqrt{X_s^x} dB_s. \end{aligned}$$

(c) Since $X_{t\wedge\tau_{\varepsilon,M}^x} \in [\varepsilon, M]$ for all $t \geq 0$ we have that

$$\mathbb{E}\int_0^{t\wedge\tau_{\varepsilon,M}^x} s'(X_s^x)^2\sigma^2 X_s^x ds = \mathbb{E}\left|s\left(X_{t\wedge\tau_{\varepsilon,M}^x}^x\right) - s(x)\right|^2$$

$$\begin{aligned} &\leq 4 \sup_{y \in [\varepsilon, M]} |s(y)|^2 \\ &=: c < \infty. \end{aligned}$$

Letting $t \to \infty$ gives that

$$\mathbb{E} \int_0^{\tau_{\varepsilon,M}^x} s'(X_s^x)^2 X_s^x \sigma^2 ds < \infty.$$

Since $X_s^x \geq \varepsilon$ for $s \in [0, \tau_{\varepsilon,M}^x]$ by definition and since

$$s'(x) = e^{\frac{2bx}{\sigma^2}} x^{-\frac{2a}{\sigma^2}} \geq e^{-2\frac{|b|M}{\sigma^2}} M^{-2\frac{a}{\sigma^2}} =: d > 0$$

we get that

$$\mathbb{E} \int_0^{\tau_{\varepsilon,M}^x} ds < \infty$$

so that $\mathbb{E}\tau_{\varepsilon,M}^x < \infty$ and $\tau_{\varepsilon,M}^x < \infty$ a.s.

(d) Now

$$s(x) = \mathbb{E} \left(s(X_{\tau_{\varepsilon,M}^x \wedge t}^x) - \int_0^{\tau_{\varepsilon,M}^x \wedge t} s'(X_s^x) \sigma \sqrt{X_s^x} dB_s \right)$$

and the boundedness of the integrand of the stochastic integral on $[0, \tau_{\varepsilon,M}^x \wedge t]$ yields that

$$s(x) = \mathbb{E} s(X_{\tau_{\varepsilon,M}^x \wedge t}^x).$$

By $t \to \infty$, dominated convergence, and the fact that $\tau_{\varepsilon,M}^x$ is almost surely finite, we conclude that

$$\begin{aligned} s(x) = \mathbb{E} s(X_{\tau_{\varepsilon,M}^x}^x) &= s(M)\mathbb{P}(\tau_M^x < \tau_\varepsilon^x) + s(\varepsilon)\mathbb{P}(\tau_M^x > \tau_\varepsilon^x) \\ &= s(M)(1 - \mathbb{P}(\tau_M^x > \tau_\varepsilon^x)) + s(\varepsilon)\mathbb{P}(\tau_M^x > \tau_\varepsilon^x) \end{aligned}$$

i.e.

$$\mathbb{P}(\tau_\varepsilon^x < \tau_M^x) = \frac{s(M) - s(x)}{s(M) - s(\varepsilon)}. \tag{4.7}$$

(e) Now we can prove our assertion.

Case (i): Assume $a \geq \frac{\sigma^2}{2}$ and set $\theta := 2\frac{a}{\sigma^2} \geq 1$. Observe that

$$\lim_{\varepsilon \downarrow 0} s(\varepsilon) = \lim_{\varepsilon \downarrow 0} \int_1^\varepsilon e^{\frac{2by}{\sigma^2}} y^{-\frac{2a}{\sigma^2}} dy$$

$$
\begin{aligned}
&= -\lim_{\varepsilon \downarrow 0} \int_{\varepsilon}^{1} \left[e^{\frac{2by}{\sigma^2}} \right] y^{-\theta} dy \\
&= -\infty.
\end{aligned}
$$

Hence the from (4.7) we get

$$\lim_{\varepsilon \downarrow 0} \mathbb{P}(\tau_{\varepsilon}^{x} < \tau_{M}^{x}) = 0.$$

Since τ_{ε}^{x} is monotone in ε we conclude that

$$\mathbb{P}\left(\tau_0^x < \tau_M^x\right) \leq \mathbb{P}\left(\bigcap_{1/N < x}^{\infty} \left\{\tau_{\frac{1}{N}}^x < \tau_M^x\right\}\right) = 0.$$

Letting $M \to \infty$ gives $\tau_M^x(\omega) \uparrow \infty$ so that

$$\mathbb{P}(\tau_0^x < \infty) = \lim_{M \uparrow \infty} \mathbb{P}(\tau_0^x < \tau_M^x) = 0.$$

Hence $\mathbb{P}(\tau_0^x = \infty) = 1$.

Case (ii) and (iii): Our first aim is to show that $\mathbb{P}(\tau_{0,M}^x < \infty) = 1$. We follow [11, Proposition 5.5.32] and define the so called *speed measure* for (4.4), i.e.

$$m(dy) := \frac{2dy}{s'(y)\sigma^2 y} \quad \text{for} \quad y \in (0, M)$$

as well as the function

$$A(x) := \int_0^M G(x,y) m(dy),$$

with

$$G(x,y) := \frac{(s(x \wedge y) - s(0))(s(M) - s(x \vee y))}{s(M) - s(0)} \quad \text{for} \quad x, y \in [0, M].$$

Obviously,

$$G(0,y) = 0 = G(M,y) \quad \text{and} \quad A(0) = A(M) = 0.$$

For $x \in (0, M)$ we have

$$\int_0^x G(x,y) m(dy) = \frac{s(M) - s(x)}{s(M) - s(0)} \int_0^x (s(y) - s(0)) m(dy)$$

and
$$\int_x^M G(x,y)m(dy) = \frac{s(x)-s(0)}{s(M)-s(0)}\int_x^M (s(M)-s(y))m(dy).$$
We get
$$\begin{aligned}
A(x) &= \frac{s(M)-s(x)}{s(M)-s(0)}\int_0^x (s(y)-s(0))m(dy) \\
&\quad +\frac{s(x)-s(0)}{s(M)-s(0)}\int_x^M (s(M)-s(y))m(dy) \\
&= \frac{s(M)-s(x)}{s(M)-s(0)}\int_0^x (s(y)-s(0))m(dy) \\
&\quad -\frac{s(x)-s(0)}{s(M)-s(0)}\int_0^x (s(M)-s(y))m(dy) \\
&\quad +\frac{s(x)-s(0)}{s(M)-s(0)}\int_0^M (s(M)-s(y))m(dy) \\
&= -\int_0^x (s(x)-s(y))m(dy) \\
&\quad +\frac{s(x)-s(0)}{s(M)-s(0)}\int_0^M (s(M)-s(y))m(dy).
\end{aligned}$$
Using the notation
$$z(x) := \frac{1}{s(M)-s(0)}\int_0^x (s(M)-s(y))m(dy),$$
we get
$$A'(x) = -s'(x)\int_0^x \frac{2dy}{s'(y)\sigma^2 y} + s'(x)z(M)$$
and
$$A''(x) = -s''(x)\int_0^x \frac{2dy}{s'(y)\sigma^2 y} - \frac{2}{\sigma^2 x} + s''(x)z(M).$$
We use that s solves (4.6) to deduce
$$\begin{aligned}
&\frac{\sigma^2}{2}xA''(x) + (a-bx)A'(x) \\
= {}& \frac{\sigma^2}{2}xs''(x)z(M) - \frac{\sigma^2}{2}xs''(x)\int_0^x \frac{2dy}{s'(y)\sigma^2 y} - \frac{\sigma^2}{2}x\frac{2}{\sigma^2 x}
\end{aligned}$$

$$
\begin{aligned}
&\quad -(a-bx)s''(x)\int_0^x \frac{2dy}{s'(y)\sigma^2 y} - (a-bx)\frac{2}{\sigma^2 x} + s''(x)z(M) \\
&= z(M)\left(\frac{\sigma^2}{2}xs''(x) + (a-bx)s'(x)\right) \\
&\quad -\int_0^x \frac{2dy}{s'(y)\sigma^2 y}\left(\frac{\sigma^2}{2}xs''(x) + (a-bx)s'(x)\right) - 1 \\
&= -1.
\end{aligned}
$$

By Itô's formula we get for $0 < \varepsilon < x < M$ that

$$
\begin{aligned}
A(X^x_{t\wedge\tau^x_{\varepsilon,M}}) &= A(x) + \int_0^{t\wedge\tau^x_{\varepsilon,M}} A'(X^x_u)\sigma\sqrt{X^x_u}dB_u \\
&\quad + \int_0^{t\wedge\tau^x_{\varepsilon,M}} \left(A'(X^x_u)(a-bX_u) + \frac{1}{2}A''(X^x_u)\sigma^2 X^x_u\right) du \\
&= A(x) + \int_0^{t\wedge\tau^x_{\varepsilon,M}} A'(X^x_u)\sigma\sqrt{X^x_u}dB_u - \int_0^{t\wedge\tau^x_{\varepsilon,M}} du.
\end{aligned}
$$

For $0 < \frac{2a}{\sigma^2} < 1$ one can show that

$$\mathbb{E}\int_0^{t\wedge\tau^x_{\varepsilon,M}} (A'(X^x_u))^2\sigma^2 X^x_u du < \infty.$$

Therefore, taking the expectation yields

$$\mathbb{E}(t\wedge\tau^x_{\varepsilon,M}) = A(x) - \mathbb{E}A(X^x_{t\wedge\tau_{\varepsilon,M}}) \le A(x)$$

and limit $\varepsilon \downarrow 0$ gives

$$\mathbb{E}(t\wedge\tau^x_{0,M}) \le A(x)$$

so that, by $t\to\infty$,

$$\mathbb{E}\tau^x_{0,M} \le A(x) < \infty$$

and in particular $\mathbb{P}(\tau^x_{0,M} < \infty) = 1$. Since $\tau^x_0 \wedge \tau^x_M < \infty$ a.s. and the process can not hit 0 and M at the same time, it holds $\mathbb{P}(\tau^x_0 = \tau^x_M) = 0$. Hence we may conclude

$$\lim_{N\to\infty}\mathbb{P}(\tau^x_M < \tau^x_{\frac{1}{N}}) = \mathbb{P}(\tau^x_M < \tau^x_0)$$

and

$$\lim_{N\to\infty}\mathbb{P}(\tau^x_{\frac{1}{N}} < \tau^x_M) = \mathbb{P}(\tau^x_0 \le \tau^x_M) = \mathbb{P}(\tau^x_0 < \tau^x_M).$$

The condition $0 < a < \frac{\sigma^2}{2}$ gives that $\theta = 2\frac{a}{\sigma^2} < 1$ and

$$\lim_{\varepsilon \downarrow 0} s(\varepsilon) = -\int_0^1 e^{2b\frac{y}{\sigma^2}} y^{-\theta} dy \in \mathbb{R}$$

which is denoted by $s(0)$. Hence

$$\begin{aligned} s(x) &= s(M) \lim_{N\to\infty} \mathbb{P}(\tau_M^x < \tau_{\frac{1}{N}}^x) + s(0) \lim_{N\to\infty} \mathbb{P}(\tau_{\frac{1}{N}}^x < \tau_M^x) \\ &= s(M)\mathbb{P}(\tau_M^x < \tau_0^x) + s(0)\mathbb{P}(\tau_0^x < \tau_M^x). \end{aligned}$$

If $b \geq 0$, then we have

$$\lim_{M\to\infty} s(M) = \infty$$

so that

$$\lim_{M\to\infty} \mathbb{P}(\tau_M^x < \tau_0^x) = 0 \quad \text{and} \quad \lim_{M\to\infty} \mathbb{P}(\tau_M^x > \tau_0^x) = 1$$

(note that $\mathbb{P}(\tau_M^x = \tau_0^x) = 0$) and

$$\mathbb{P}\left(\bigcup_{M>0} \{\tau_M^x > \tau_0^x\}\right) = 1.$$

Because $\lim_M \tau_M^x(\omega) = \infty$ for all $\omega \in \Omega$, this implies

$$\mathbb{P}(\tau_0^x < \infty) = 1.$$

If $b < 0$, then we have

$$\lim_{M\to\infty} s(M) =: s(\infty) \in (0, \infty)$$

and

$$s(x) = s(\infty)\mathbb{P}(\tau_0^x = \infty) + s(0)\mathbb{P}(\tau_0^x < \infty),$$

and $\mathbb{P}(\tau_0^x = \infty) \in (0,1)$ as well as $\mathbb{P}(\tau_0^x < \infty) \in (0,1)$. $\square$

4.8 The martingale representation theorem

Let $B = (B^1, \ldots, B^d)$ be a d-dimensional (standard) Brownian motion on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. $B^i = (B^i_t)_{t \in [0.T]}$ are Brownian motions for $i = 1, \ldots, n$ that are independent from each other, meaning that all families

$$B^1_{t^1_1} - B^1_{t^1_0}, \ldots, B^1_{t^1_{N^1}} - B^1_{t^1_{N^1-1}}, \ldots, B^d_{t^d_1} - B^d_{t^d_0}, \ldots, B^d_{t^d_{N^d}} - B^d_{t^d_{N^d-1}}$$

are independent for $0 = t^i_0 < \cdots < t^i_{N^i} = T$. If we define

$$\mathcal{F}_t := \sigma((B^1_s, \ldots B^d_s), s \in [0, t]) \vee \{B \in \mathcal{F} : \mathbb{P}(B) = 0\},$$

then the filtration $(\mathcal{F}_t)$ is right-continuous [11, Section 2, Proposition 7.7], so that we use the stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0,T]})$ in the following. We recall

$$\mathcal{L}_2 :=$$
$$\left\{ L = (L_u)_{u \in [0,T]} : L \text{ progressively measurable }, [L]_2^2 = \mathbb{E} \int_0^T L_u^2 du < \infty \right\}.$$

Theorem 4.8.1 (Stochastic integral representation). *For $F \in L_2(\Omega, \mathcal{F}_T, \mathbb{P})$ there are $L^j \in \mathcal{L}_2$, $j = 1, \ldots, d$, such that*

$$F = \mathbb{E}F + \sum_{j=1}^d \int_0^T L^j_u dB^j_u.$$

Proof. The proof follows from [10, Theorem 56.2] and [18, Lemma V3.1]. We define

$$\mathcal{S} := \left\{ Z \in L_2(\Omega, \mathcal{F}_T, \mathbb{P}) : Z = \sum_{j=1}^d \int_0^T L^j_u dB^j_u,\ L^j \in \mathcal{L}_2,\ j = 1, \ldots, d \right\}.$$

We will verify our statement by proving that $\mathbb{R} \oplus \mathcal{S} = L_2(\Omega, \mathcal{F}_T, \mathbb{P})$. This follows from proving that for $Y \in L_2(\Omega, \mathcal{F}_T, \mathbb{P})$ with $\mathbb{E}Y = 0$ and $Y \perp \mathcal{S}$ one has $Y = 0$ a.s.

Step 1: If $K^j : [0, T] \to \mathbb{R}$ is Borel measurable with $\int_0^T (K^j_u)^2 du < \infty$, we set

$$\mathcal{E}_t(K) := \exp\left\{ \sum_{j=1}^d \int_0^t K^j_u dB^j_u - \frac{1}{2} \sum_{j=1}^d \int_0^t (K^j_u)^2 du \right\}.$$

By the Novikov condition we have that $(\mathcal{E}_t(K))_{t\in[0,T]}$ is a martingale. Moreover,

$$|\mathcal{E}_t(K)|^2 = \mathcal{E}_t(2K)\exp\left\{\sum_{j=1}^d \int_0^t (K_u^j)^2 du\right\} \tag{4.8}$$

so that

$$\mathbb{E}|\mathcal{E}_t(K)|^2 = \exp\left\{\sum_{j=1}^d \int_0^t (K_u^j)^2 du\right\}. \tag{4.9}$$

Our assumption implies that

$$\mathbb{E}\mathcal{E}_T(K)Y = 0 \quad \text{for all } K^1,\ldots,K^d \text{ with } \int_0^T (K_u^j)^2 du < \infty$$

because by Itô's formula one has $\mathcal{E}_T(K) - 1 \in \mathcal{S}$.

Step 2: Let $0 = t_0 < t_1 < \cdots < t_n = T$ and assume that

$$K^j = \sum_{m=1}^n \alpha_m^j \chi_{(t_{m-1},t_m]}, \quad \alpha_m^j \in \mathbb{R}.$$

Then

$$\mathcal{E}_T(K) = \exp\left(\sum_{j=1}^d \sum_{m=1}^n \alpha_m^j (B_{t_m}^j - B_{t_{m-1}}^j) - \frac{1}{2}\sum_{j=1}^d \int_0^T (K_u^j)^2 du\right).$$

Using the $\sigma(B_{t_1},\ldots,B_{t_n})$-measurability of $\mathcal{E}_T(K)$, we get that

$$\begin{aligned} 0 &= \mathbb{E}(\mathcal{E}_T(K)Y) \\ &= \mathbb{E}\mathbb{E}\left[\mathcal{E}_T(K)Y\middle|\sigma(B_{t_1},\ldots,B_{t_n})\right] \\ &= \mathbb{E}\left[\mathcal{E}_T(K)\mathbb{E}\left[Y\middle|\sigma(B_{t_1},\ldots,B_{t_n})\right]\right]. \end{aligned}$$

Step 3: We deduce that

$$\mathbb{E}\left[Y\middle|\sigma(B_{t_1},\ldots,B_{t_n})\right] = 0.$$

To show this, we set [1]

$$g(B_{t_1},\ldots,B_{t_n}) := \mathbb{E}\left[Y\middle|\sigma(B_{t_1},\ldots,B_{t_n})\right]$$

[1] The existence of a Borel function $g : (\mathbb{R}^d)^n \to \mathbb{R}$ such that $g(B_{t_1},\ldots,B_{t_n}) = \mathbb{E}[Y|\sigma(B_{t_1},\ldots,B_{t_n})]$ follows from the factorization theorem (see [1]).

and fix $\xi_1, \ldots, \xi_n \in \mathbb{R}^d$. For $z \in \mathbb{C}$ we set

$$f(z) := \mathbb{E}\left[\exp\left(z\sum_{m=1}^{n}\langle \xi_m, B_{t_m} - B_{t_{m-1}}\rangle\right) g(B_{t_1}, \ldots, B_{t_n})\right].$$

One can show that $f : \mathbb{C} \to \mathbb{C}$ is holomorphic and on $\mathbb{R}$ the function f is identically 0. By the identity theorem of holomorphic functions, $f \equiv 0$. But then especially

$$\mathbb{E}\left[\exp\left(i\sum_{m=1}^{n}\langle \alpha_m, B_{t_m} - B_{t_{m-1}}\rangle\right) g(B_{t_1}, \ldots, B_{t_n})\right] = 0$$

for all $\alpha_1, \ldots, \alpha_n \in \mathbb{R}^d$, so that by the uniqueness theorem for the Fourier transform,

$$g(B_{t_1}, \ldots, B_{t_n}) = 0 \text{ a.s.}$$

In fact, for $d = 1$ the precise argument is as follows: The above equation implies that

$$\int_{\mathbb{R}^n} e^{i\sum_{m=1}^{n}\alpha_m x_m} g(x_1, x_1+x_2, \ldots, x_1+x_2+\ldots+x_n) \\ \times h_{(B_{t_1},\ldots,B_{t_n}-B_{t_{n-1}})}(x_1, x_2, \ldots, x_n)dx_1\cdots dx_n \equiv 0$$

where $h_{(B_{t_1},\ldots,B_{t_n}-B_{t_{n-1}})}$ is the density of the law of the random vector $(B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}})$ with respect to the Lebesgue measure on $\mathbb{R}^n$. Hence

$$g(x_1, x_1+x_2, \ldots, x_1+x_2+\ldots+x_n)h_{(B_{t_1},\ldots,B_{t_n}-B_{t_{n-1}})}(x_1, x_2, \ldots, x_n) = 0 \text{ a.e.}$$

with respect to the Lebesgue measure on $\mathbb{R}^n$ and therefore

$$g(x_1, x_1+x_2, \ldots, x_1+x_2+\ldots+x_n) = 0 \text{ a.e.}$$

with respect to the Lebesgue measure on $\mathbb{R}^n$.

<u>Step 4</u>: To conclude with $Y = 0$ a.s. as we proceed as follows: We let

$$\mathcal{A} := \bigcup_{m=1}^{\infty}\ \bigcup_{0\le t_1<\cdots<t_m=T} \sigma(B_{t_0}, \ldots, B_{t_m})$$

which is an algebra that generates $\sigma(B_t : t \in [0,T])$. According to [7, Lemma 2.1.7] we have that for any $\varepsilon > 0$ and any $B \in \sigma(B_t : t \in [0,T])$ there exists

$A \in \mathcal{A}$ such that $\mathbb{P}(A\Delta B) < \varepsilon$. This remains true if $B \in \sigma(B_t : t \in [0,T])$ is replaced by $B \in \mathcal{F}_T^B$. Set $Y^+ := \max\{Y, 0\}$ and suppose that $\mathbb{E}Y^+\chi_B = c > 0$ for some $B \in \mathcal{F}_T^B$ with $B \subseteq \{Y \geq 0\}$. Then we find a set $A \in \mathcal{A}$ such that

$$|\mathbb{E}[Y(\chi_A - \chi_B)]| \leq \|Y\|_{L_2(\Omega,\mathcal{F}_T^B,\mathbb{P})}\mathbb{P}(A\Delta B)^{1/2} \leq \frac{c}{2}.$$

But since

$$\mathbb{E}Y\chi_A = \mathbb{E}Y(\chi_A - \chi_B) + \mathbb{E}Y\chi_B = \mathbb{E}Y(\chi_A - \chi_B) + \mathbb{E}Y^+\chi_B,$$

it follows that $\mathbb{E}[Y\chi_A] > 0$, which is a contradiction. The same argument can be applied to Y^-, so that $\mathbb{E}Y\chi_B = 0$ for all $B \in \mathcal{F}_T^B$, implying that $Y = 0$ a.s. Therefore, $\mathbb{R} \oplus \mathcal{S} = L_2(\Omega, \mathcal{F}_T^B, \mathbb{P})$. □

The following corollary shows that in this particular setting (called Wiener space setting) all square integrable martingales can be chosen to be continuous. This is **not true in general.**

Corollary 4.8.2. *Assume that $M = (M_t)_{t\in[0,T]} \subseteq L_2$ is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$. Then there exists a modification $\tilde{M} = (\tilde{M}_t)_{t\in[0,T]}$ (i.e. $\mathbb{P}(M_t = \tilde{M}_t) = 1$ for all $t \in [0,T]$) such that $t \mapsto \tilde{M}_t(\omega)$ is continuous for all $\omega \in \Omega$.*

Chapter 5

Backward stochastic differential equations (BSDEs)

5.1 Introduction

Let $T > 0$. We have considered SDEs

$$dX_t = a(t, X_t)dt + \sigma(t, X_t)dB_t, \quad X_0 = x.$$

A solution (in the strong sense) we defined as an adapted process $(X_t)_{t\in[0,T]}$ solving the equation

$$X_t = x + \int_0^t a(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s, \tag{5.1}$$

where we assumed that $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ (satisfying the usual conditions) and a Brownian motion $(B_t)_{t\in[0,T]}$ w.r.t. $(\mathcal{F}_t)_{t\in[0,T]}$ are given, also the initial condition $X_0 = x$. Could one, instead of the initial condition $X_0 = x$, demand a terminal condition $X_T = \xi \in L_2$? Re-writing this equation would give

$$X_t = \xi - \int_t^T a(s, X_s)ds - \int_t^T \sigma(s, X_s)dB_s, \quad t \in [0, T], \tag{5.2}$$

and $X_T = \xi$. However, equation (5.2) does not have in all cases an *adapted* solution. For example, take $\xi = 1$, $a \equiv 0$, and $\sigma \equiv 1$. This would yield that

$$X_t = 1 - \int_t^T dB_s = 1 + B_t - B_T,$$

so that X_t would not be $\mathcal{F}_t$-measurable. Instead of (5.2), we will consider

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s, \quad t \in [0,T], \tag{5.3}$$

and call (Y, Z) a solution, if $Y = (Y_t)_{t\in[0,T]}$ and $Z = (Z_t)_{t\in[0,T]}$ are progressively measurable processes satisfying further conditions specified below. In the above equation ξ is called *terminal condition*, f *generator*, and the pair (ξ, f) the data of the BSDE. So *given* the data (ξ, f), we look for an *adapted* solution (Y, Z).

5.2 Setting

For the terminal condition we use

(C_ξ) One has $\xi \in L_2$.

We use the following assumption (C_f) on the generator

$$f : [0,T] \times \Omega \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} :$$

(C_f1) $f(\,\cdot\,,\,\cdot\,, y, z)$ is progressively measurable for all $y, z \in \mathbb{R}$.

(C_f2) There exists an $L_f > 0$ such that

$$|f(t, \omega, y, z) - f(t, \omega, \bar{y}, \bar{z})| \le L_f\,(|y - \bar{y}| + |z - \bar{z}|)$$

for all $(t, \omega) \in [0,T] \times \Omega$ and $y, \bar{y}, z, \bar{z} \in \mathbb{R}$.

(C_f3) $\mathbb{E}\int_0^T f^2(t, 0, 0)dt < \infty$.

Moreover, we define and recall, respectively:

(i) The space $\mathcal{S}_2$ consists of all adapted and continuous processes $(X_t)_{t\in[0,T]}$ such that $\|X\|^2_{\mathcal{S}_2} := \mathbb{E}\sup_{0\le t\le T}|X_t|^2 < \infty$.

(ii) The space $\mathcal{L}_2$ consists of all progressively measurable $(X_t)_{t\in[0,T]}$ such that $[X]_2^2 = \mathbb{E}\int_0^T X_t^2 dt < \infty$.

(iii) For $\beta \ge 0$ and $X \in \mathcal{L}_2$ we let $[X]^2_{2,\beta} := \mathbb{E}\int_0^T X_t^2 e^{\beta t} dt$.

Notice that the norms $[\cdot]^2_{2,\beta}$ are equivalent for all $\beta \ge 0$.

5.3 A priori estimate

We start with a sufficient condition such that a local martingale is a martingale.

Lemma 5.3.1. *A (continuous) local martingale M which is bounded by an integrable bound G :*

$$\sup_t |M_t| \le G, \quad \mathbb{E}G < \infty$$

is a martingale.

(see [12, Theorem 7.21, p.196]).

Now we prove an *a priori estimate*, that can be seen as a *stability result* as well.

Proposition 5.3.2. *Assume that (C_ξ), $(C_{\bar{\xi}})$, (C_f), and $(C_{\bar{f}})$ hold for the data (ξ, f) and $(\bar{\xi}, \bar{f})$. Let (Y, Z) and $(\bar{Y}, \bar{Z})$ be in $\mathcal{S}_2 \times \mathcal{L}_2$ and be solutions to*

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s, \quad t \in [0, T],$$
$$\bar{Y}_t = \bar{\xi} + \int_t^T \bar{f}(s, \bar{Y}_s, \bar{Z}_s)ds - \int_t^T \bar{Z}_s dB_s, \quad t \in [0, T],$$

respectively. Then, for all $\beta \ge A + 2L_{\bar{f}} + 2L_{\bar{f}}^2 + \frac{1}{2}$, where $A > 0$, it holds that

$$[Y - \bar{Y}]_{2,\beta}^2 + [Z - \bar{Z}]_{2,\beta}^2 \le 2e^{\beta T}\mathbb{E}\left|\xi - \bar{\xi}\right|^2 + \frac{2}{A}[f(\,\cdot\,, Y_\cdot, Z_\cdot) - \bar{f}(\,\cdot\,, Y_\cdot, Z_\cdot)]_{2,\beta}^2.$$

Proof. Assume $\beta > 0$ and apply Itô's formula, in order to get

$$\begin{aligned}
e^{\beta T}(\xi - \bar{\xi})^2 &= e^{\beta T}(Y_T - \bar{Y}_T)^2 \\
&= e^{\beta t}(Y_t - \bar{Y}_t)^2 \\
&\quad + \int_t^T \beta e^{\beta s}(Y_s - \bar{Y}_s)^2 ds \\
&\quad - \int_t^T e^{\beta s} 2(Y_s - \bar{Y}_s)\left[f(s, Y_s, Z_s) - \bar{f}(s, \bar{Y}_s, \bar{Z}_s)\right] ds \\
&\quad + \frac{1}{2}\int_t^T 2e^{\beta s}(Z_s - \bar{Z}_s)^2 ds
\end{aligned}$$

$$+2\int_t^T e^{\beta s}(Y_s-\bar{Y}_s)(Z_s-\bar{Z}_s)dB_s.$$

We rearrange the terms to get

$$\begin{aligned}
e^{\beta t}\left|Y_t-\bar{Y}_t\right|^2 &+ \int_t^T e^{\beta s}\left|Z_s-\bar{Z}_s\right|^2 ds \\
&= e^{\beta T}\left|\xi-\bar{\xi}\right|^2 - \beta\int_t^T e^{\beta s}\left|Y_s-\bar{Y}_s\right|^2 ds \\
&\quad + 2\int_t^T e^{\beta s}(Y_s-\bar{Y}_s)\left(f(s,Y_s,Z_s)-\bar{f}(s,\bar{Y}_s,\bar{Z},s)\right) ds \\
&\quad - 2\int_t^T e^{\beta s}(Y_s-\bar{Y}_s)(Z_s-\bar{Z}_s)dB_s. \qquad (5.4)
\end{aligned}$$

Now we estimate the terms of the right-hand side of (5.4). If we let

$$I := \sup_{t\in[0,T]}\left|\int_t^T e^{\beta s}(Y_s-\bar{Y}_s)(Z_s-\bar{Z}_s)dB_s\right|,$$

then

$$\begin{aligned}
\mathbb{E}I &\leq C_1\mathbb{E}\left(\int_0^T |e^{\beta s}(Y_s-\bar{Y}_s)(Z_s-\bar{Z}_s)|^2 ds\right)^{\frac{1}{2}} \\
&\leq C_1\mathbb{E}\left[\left(\sup_t\left|Y_t-\bar{Y}_t\right|\right)\left(\int_0^T e^{2\beta s}\left|Z_s-\bar{Z}_s\right|^2 ds\right)^{1/2}\right] \\
&\leq C_1\|Y-\bar{Y}\|_{\mathcal{S}_2}[Z-\bar{Z}]_{2,\beta} < \infty
\end{aligned}$$

by the Burkholder-Davis-Gundy inequality (Proposition 4.3.1) and the Cauchy-Schwartz inequality. Therefore, the maximal function of the Itô integral in (5.4) is integrable, which implies that the Itô integral $\left(\int_0^t e^{\beta s}(Y_s-\bar{Y}_s)(Z_s-\bar{Z}_s)dB_s\right)_{t\geq 0}$ is a martingale by Lemma 5.3.1. Hence we get that

$$\mathbb{E}\int_t^T e^{\beta s}(Y_s-\bar{Y}_s)(Z_s-\bar{Z}_s)dB_s = 0.$$

For the second last term in (5.4) we use, for $A>0$ and $a,b\in\mathbb{R}$,

$$2ab \leq Aa^2 + \frac{1}{A}b^2$$

and

$$\left|\bar f(s,Y_s,Z_s)-\bar f(s,\bar Y_s,\bar Z_s)\right|\le L_{\bar f}\left(\left|Y_s-\bar Y_s\right|+\left|Z_s-\bar Z_s\right|\right)$$

in order to get

$$\begin{aligned}
& 2e^{\beta s}\left|Y_s-\bar Y_s\right|\left|f(s,Y_s,Z_s)-\bar f(s,\bar Y_s,\bar Z_s)\right| \\
\le\; & 2e^{\beta s}\left|Y_s-\bar Y_s\right|\left|f(s,Y_s,Z_s)-\bar f(s,Y_s,Z_s)\right| \\
& +2e^{\beta s}\left|Y_s-\bar Y_s\right|\left|\bar f(s,Y_s,Z_s)-\bar f(s,\bar Y_s,\bar Z_s)\right| \\
\le\; & Ae^{\beta s}\left|Y_s-\bar Y_s\right|^2+\frac{1}{A}e^{\beta s}\left|f(s,Y_s,Z_s)-\bar f(s,Y_s,Z_s)\right|^2 \\
& +2L_{\bar f}e^{\beta s}\left|Y_s-\bar Y_s\right|^2+2L_{\bar f}^2e^{\beta s}\left|Y_s-\bar Y_s\right|^2+\frac{1}{2}e^{\beta s}\left|Z_s-\bar Z_s\right|^2.
\end{aligned}$$

Here we used the above inequality with $a=\sqrt{2}L_{\bar f}\left|Y_s-\bar Y_s\right|$ and $b=\sqrt{1/2}\left|Z_s-\bar Z_s\right|$ to estimate

$$2e^{\beta s}L_{\bar f}\left|Y_s-\bar Y_s\right|\left|Z_s-\bar Z_s\right|\le 2L_{\bar f}^2e^{\beta s}\left|Y_s-\bar Y_s\right|^2+\frac{1}{2}e^{\beta s}\left|Z_s-\bar Z_s\right|^2.$$

Hence, for $\beta\ge A+2L_{\bar f}+2L_{\bar f}^2+\frac{1}{2}$, (5.4) implies that

$$\begin{aligned}
& \mathbb{E}e^{\beta t}\left|Y_t-\bar Y_t\right|^2+\frac{1}{2}\mathbb{E}\int_t^T e^{\beta s}\left|Z_s-\bar Z_s\right|^2ds \\
\le\; & \mathbb{E}e^{\beta T}\left|\xi-\bar\xi\right|^2-\frac{1}{2}\mathbb{E}\int_t^T e^{\beta s}\left|Y_s-\bar Y_s\right|^2ds \\
& +\frac{1}{A}\mathbb{E}\int_t^T e^{\beta s}\left|f(s,Y_s,Z_s)-\bar f(s,Y_s,Z_s)\right|^2ds.
\end{aligned}$$

We move the term with "$-$" to the left-hand side and consider the inequality for $t=0$. We derive

$$\begin{aligned}
[Y-\bar Y]_{2,\beta}^2+[Z-\bar Z]_{2,\beta}^2 \;\le\;& 2e^{\beta T}\mathbb{E}\left|\xi-\bar\xi\right|^2 \\
& +\frac{2}{A}[f(\,\cdot\,,Y_\cdot,Z_\cdot)-\bar f(\,\cdot\,,Y_\cdot,Z_\cdot)]_{2,\beta}^2.
\end{aligned}$$

□

Now we state and prove the fundamental existence and uniqueness theorem for BSDEs:

Theorem 5.3.3. *Assume that we have the filtered probability space $(\Omega, \mathcal{F}_T^B, \mathbb{P}; (\mathcal{F}_t^B)_{t\in[0,T]})$ and let (C_ξ) and (C_g) hold for the data (f, ξ), then*

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s, \quad t \in [0,T], \tag{5.5}$$

has a solution (Y, Z) which is unique in $\mathcal{S}_2 \times \mathcal{L}_2$.

Proof. We will define a map $F : \mathcal{L}_2 \times \mathcal{L}_2 \to \mathcal{L}_2 \times \mathcal{L}_2$ and show that F is a contraction, i.e. if $(\mathcal{Y}, \mathcal{Z}), (\bar{\mathcal{Y}}, \bar{\mathcal{Z}}) \in \mathcal{L}_2 \times \mathcal{L}_2$, $(Y, Z) = F(\mathcal{Y}, \mathcal{Z})$, and $(\bar{Y}, \bar{Z}) = F(\bar{\mathcal{Y}}, \bar{\mathcal{Z}})$, then

$$[Y - \bar{Y}]_{2,\beta}^2 + [Z - \bar{Z}]_{2,\beta}^2 \le c\left[[\mathcal{Y} - \bar{\mathcal{Y}}]_{2,\beta}^2 + [\mathcal{Z} - \bar{\mathcal{Z}}]_{2,\beta}^2\right]$$

for some constant $0 < c < 1$.

(a) Construction of the map F: Let $(\mathcal{Y}, \mathcal{Z}) \in \mathcal{L}_2 \times \mathcal{L}_2$. We look for a pair $(Y, Z) = F(\mathcal{Y}, \mathcal{Z})$ that solves

$$Y_t = \xi + \int_t^T f(s, \mathcal{Y}_s, \mathcal{Z}_s)ds - \int_t^T Z_s dB_s. \tag{5.6}$$

We get Z from Theorem 4.8.1 about the represenation property on the Wiener space. For this we let

$$M_t := \mathbb{E}\left[\xi + \int_0^T f(s, \mathcal{Y}_s, \mathcal{Z}_s)ds \middle| \mathcal{F}_t\right].$$

Then $M = (M_t)_{t\in[0,T]}$ is a square-integrable martingale and there exists a unique $Z \in \mathcal{L}_2$ such that

$$M_t = M_0 + \int_0^t Z_s dB_s \quad \text{for} \quad t \in [0,T] \text{ a.s.}$$

Now, letting

$$Y_t := \mathbb{E}\left[\xi + \int_t^T f(s, \mathcal{Y}_s, \mathcal{Z}_s)ds \middle| \mathcal{F}_t\right]$$

we get

$$Y_t = M_t - \int_0^t f(s, \mathcal{Y}_s, \mathcal{Z}_s)ds$$

$$
\begin{aligned}
&= M_0 + \int_0^t Z_s dB_s - \int_0^t f(s, \mathcal{Y}_s, \mathcal{Z}_s) ds \\
&= M_0 + \int_0^T Z_s dB_s - \int_0^T f(s, \mathcal{Y}_s, \mathcal{Z}_s) ds + \int_t^T f(s, \mathcal{Y}_s, \mathcal{Z}_s) ds - \int_t^T Z_s dB_s \\
&= \xi + \int_t^T f(s, \mathcal{Y}_s, \mathcal{Z}_s) ds - \int_t^T Z_s dB_s.
\end{aligned}
$$

Hence, we have

$$F(\mathcal{Y}, \mathcal{Z}) = (Y, Z).$$

(b) <u>F is a contraction</u>: Let $(\mathcal{Y}, \mathcal{Z}), (\bar{\mathcal{Y}}, \bar{\mathcal{Z}}) \in \mathcal{L}_2 \times \mathcal{L}_2$. Then, by the a priori estimate (assume, for the moment, that $\mathcal{Y}, \bar{\mathcal{Y}} \in \mathcal{S}_2$), for $\xi = \bar{\xi}$, setting

$$
\begin{aligned}
f_0(s, Y_s, Z_s) &:= f(s, \mathcal{Y}_s, \mathcal{Z}_s) \\
\bar{f}_0(s, \bar{Y}_s, \bar{Z}_s) &:= f(s, \bar{\mathcal{Y}}_s, \bar{\mathcal{Z}}_s),
\end{aligned}
$$

(notice that the r.h.s. is not depending on $Y, Z, \bar{Y}, \bar{Z}$) and for the solutions (Y, Z) and $(\bar{Y}, \bar{Z})$ we get

$$
\begin{aligned}
[Y - \bar{Y}]_{2,\beta}^2 + [Z - \bar{Z}]_{2,\beta}^2 &\leq \frac{2}{A} \mathbb{E} \int_0^T e^{\beta s} \left| f(s, \mathcal{Y}_s, \mathcal{Z}_s) - f(s, \bar{\mathcal{Y}}_s, \bar{\mathcal{Z}}_s) \right|^2 ds \\
&\leq \frac{4L_f^2}{A} \left[[\mathcal{Y} - \bar{\mathcal{Y}}]_{2,\beta}^2 + [\mathcal{Z} - \bar{\mathcal{Z}}]_{2,\beta}^2 \right]. \qquad (5.7)
\end{aligned}
$$

Now we choose A large enough so that $4L_g^2 < A$, and then a corresponding β according to Proposition 5.3.2.

(c) <u>The iteration</u>: Consider the following procedure: Start with $Y^0 \equiv 0, Z^0 \equiv 0$ and define $(Y^{k+1}, Z^{k+1}) = F(Y^k, Z^k)$. Then, because F is a contraction, there exists $(Y, Z) \in \mathcal{L}_2 \times \mathcal{L}_2$ such that

$$(Y^k, Z^k) \to (Y, Z) \text{ in } \mathcal{L}_2 \times \mathcal{L}_2,$$

provided that $Y^k \in \mathcal{S}_2$ for all $k = 1, 2, ...$

(d) <u>$Y^k \in \mathcal{S}_2$ and $Y \in \mathcal{S}_2$</u>: From the construction of the map F we know that

$$Y_t = M_t - \int_0^t f(s, \mathcal{Y}_s, \mathcal{Z}_s) ds.$$

Therefore,

$$\|Y\|_{\mathcal{S}_2}^2 = \mathbb{E} \sup_t |Y_t|^2$$

$$
\begin{aligned}
&\leq 2\,\mathbb{E}\sup_{0\leq t\leq T}|M_t|^2 + 2\mathbb{E}\left|\int_0^T |f(s,\mathcal{Y}_s,\mathcal{Z}_s)|ds\right|^2 \\
&\leq 8\,\mathbb{E}\,|M_T|^2 + 2\mathbb{E}\left|\int_0^T |f(s,\mathcal{Y}_s,\mathcal{Z}_s)|ds\right|^2 \\
&\leq 16\,\mathbb{E}\,|Y_T|^2 + 4\mathbb{E}\left|\int_0^T |f(s,\mathcal{Y}_s,\mathcal{Z}_s)|ds\right|^2 \\
&\leq 16\,\mathbb{E}\xi^2 + 4T\,\mathbb{E}\int_0^T |f(s,\mathcal{Y}_s,\mathcal{Z}_s) - f(s,0,0) + f(s,0,0)|^2\,ds \\
&\leq 16\,\mathbb{E}\xi^2 + 8T\,\mathbb{E}\int_0^T \left[L_f^2\left(\mathcal{Y}_s^2+\mathcal{Z}_s^2\right) + f^2(s,0,0)\right]ds \\
&< \infty
\end{aligned}
$$

where we used Doob's maximal inequality for $p=2$ and the assumptions $\mathbb{E}\xi^2<\infty$ and $(\mathcal{Y},\mathcal{Z})\in\mathcal{L}_2\times\mathcal{L}_2$. This shows that the $Y_k\in\mathcal{S}_2$, but the same argument also shows that $Y\in\mathcal{S}_2$ if we use the argument for $Y=\mathcal{Y}$ and $Z=\mathcal{Z}$.

(e) The Uniqueness follows from Proposition 5.3.2 because for $\xi=\bar{\xi}$ and $f=\bar{f}$ we have that

$$
\begin{aligned}
[Y-\bar{Y}]_{2,\beta}^2 + [Z-\bar{Z}]_{2,\beta}^2 &\leq 2e^{\beta T}\mathbb{E}\left|\xi-\bar{\xi}\right|^2 + \frac{2}{A}[f(\,\cdot\,,Y_\cdot,Z_\cdot)-\bar{f}(\,\cdot\,,Y_\cdot,Z_\cdot)]_{2,\beta}^2 \\
&= 0.
\end{aligned}
$$

□

Bibliography

[1] H. Bauer, Probability Theory, De Gruyter, 1996.

[2] A.N. Borodin and P. Salminen, Handbook of Brownian motion – facts and formulae, Birkhäuser, 1996.

[3] A.S. Cherny and H.J. Engelbert, Singular Stochastic Differential Equations. LNM 1858, Springer.

[4] K.L. Chung abd R.J. Williams, Introduction to Stochastic integration, 1983.

[5] S. Geiss. Quantitative approximation of certain stochastic integrals. *Stochastics and Stochastics Reports*, 73:241–270, 2002.

[6] S. Geiss and M. Hujo. Interpolation and approximation in $L_2(\gamma)$. *Journal of Approximation Theory*, 144:213–232, 2007.

[7] S. Geiss. Stochastic Processes in discrete time. *Lecture Notes 62*, University of Jyväskylä, 2009.

[8] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, 1989, second Edition.

[9] K. Itô and H.P. McKean, Diffusion processes and their sample paths, Springer, 1965.

[10] G. Kallianpur, P. Sundar, Stochastic Analysis and Diffusion Processes, Oxford University Press, 2014.

[11] I. Karatzas and S. Shreve, Brownian Motion and Stochastic Calculus, 1988, Springer.

[12] Klebaner, Introduction to Stochastics Calculus, Imperial College Press, 2nd Edition, 2005.

[13] P.E. Kloeden and E. Platen, Numerical Solution of Stochastic Differential Equations, Springer, 1992, 1995.

[14] D. Lamberton and B. Lapeyre, Introduction to Stochastic Calculus Applied to Finance, 2nd ed., Chapman & Hall, 2007.

[15] H.P. McKean, Stochastic Integrals, Academic Press, 1969.

[16] D. Nualart, The Malliavin Calculus and Related Topics, Springer, 1995.

[17] P. Protter, Stochastic integration and differential equations, Springer, second edition.

[18] D. Revuz and M. Yor, Continuous Martingales and Brownian Motion, Springer, 1994.

[19] L. C. G. Rogers and D. Williams, Diffusions, Markov Processes and Martingales, Foundations, Cambridge University Press, 2000.

[20] L. C. G. Rogers and D. Williams, Diffusions, Markov Processes and Martingales. Ito Calculus, Cambridge University Press, 2000.

[21] P. Sivula, Stochastic models for life insurance mathematics. Master Thesis, 2010.

[22] M. Steele, Stochastic Calculus and Financial Applications. Springer, 2001.