Chapter 1

Stochastic calculus

1.1 Probability

1.1.1 Measure theory

The basic concepts of measure theory that reader should be aware of are

- $(X, \mathcal{A})$ a measurable space: X is set, $\mathcal{A}$ its σ-algebra
- measurable function f, (positive) measure μ, integral $\int f \mathrm{d}\mu$
- Lebesgue measure on $\mathbb{R}^d$
- $L^p(\mu)$ space: Measurable f is in $L^p(\mu)$ if $\int |f|^p \mathrm{d}\mu < \infty$. Notation: $\|f\|_p = (\int |f|^p \mathrm{d}\mu)^{1/p}$.
- Product measures: If $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are measure spaces, then their product space is $(X \times Y, \mathcal{A} \times \mathcal{B}, \mu \times \nu)$ where $X \times Y$ is Cartesian product, $\mathcal{A} \times \mathcal{B}$ the σ-algebra generated by $A \times B$, $A \in \mathcal{A}$ and $B \in \mathcal{B}$, and $\mu \times \nu$ the unique extension of $A \times B \mapsto \mu(A)\nu(B)$. (Here we have to assume that both measures μ and ν are σ-finite in the sense that X can be written as $X = \bigcup_{k=1}^{\infty} X_k$ where X_k a measurable sets with finite μ-measure and the same holds for Y and ν.)

Here is a summary of some results of measure theory. For the details and proof see the books mentioned above.

- Monotone convergence theorem: If f_n are measurable functions such that $0 \leq f_n \nearrow f$, then $\int f \mathrm{d}\mu = \lim_{n\to\infty} \int f_n \mathrm{d}\mu$
- Dominated convergence theorem: If f_n are measurable functions and $f = \lim_{n\to\infty} f_n$ exists almost everywhere and $\exists g \geq 0$ such that $|f_n| \leq g$ for all n and $\int g \mathrm{d}\mu < \infty$, then $\int f \mathrm{d}\mu = \lim_{n\to\infty} \int f_n \mathrm{d}\mu$.
- Fubini's theorem: Assume that μ and ν are σ-finite. Let $f \in \mathcal{A} \times \mathcal{B}$. If $f \geq 0$ or $\int |f| \mathrm{d}(\mu \times \nu) < \infty$ then $\int_X (\int_Y f \mathrm{d}\nu) \mathrm{d}\mu = \int_{X \times Y} f \mathrm{d}(\mu \times \nu) = \int_Y (\int_X f \mathrm{d}\mu) \mathrm{d}\nu$.
- Radon–Nikodym theorem: If ν is a σ-finite signed measure and μ is a σ-finite measure on $(X, \mathcal{A})$ and ν is absolutely continuous with respect to μ, then exist $g \in \mathcal{F}$ such that $\nu(A) = \int_A g \, \mathrm{d}\mu$. Here ν is absolutely continuous with respect to μ, if $\nu(A) = 0$ whenever $\mu(A) = 0$, $A \in \mathcal{F}$. A notation: $f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ and it is called Radon–Nikodym derivative.

A notation which sometimes handy: $f \in \mathcal{A}$ where f is a function on X means that f is $\mathcal{A}$-measurable.

1.1.2 Probability theory

Probability theory is essentially measure theoretical formulation of probability. Therefore the basics of probability are easily accessible to anybody with background in mathematical analysis. Here is a list of basic facts about probability:

- A probability space is a measure space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{P}$ is a probability measure, i.e., $\mathbb{P}(\Omega) = 1$. Ω "outcomes", $\mathcal{F}$ "events"
- A random variable is a $\mathcal{F}$-measurable function $X : \Omega \to \mathbb{R}$. H-valued random variable is a measurable function $X : \Omega \to H$ (H is a measurable space).
- The expected value of X is $\mathbb{E}(X) = \int X \mathrm{d}\mathbb{P} \in [-\infty, \infty]$, which makes sense when $X \geq 0$ or when either $\int X^+ \mathrm{d}\mathbb{P} < \infty$ or $\int X^- \mathrm{d}\mathbb{P} < \infty$, where $X = X^+ - X^-$ is the decomposition of X into positive and negative part.
- $L^p(\mathbb{P})$ space: $\|X\|_p = (\mathbb{E}(|X|^p))^{1/p} < \infty$. By Hölder inequality, $\|X\|_p \leq \|X\|_q$ for $1 \leq p \leq q$ and hence $L^q(\mathbb{P}) \subset L^p(\mathbb{P})$ (a fact which isn't necessarily true for general measures).
- Independence: sub-σ-algebras $\mathcal{A}_1, \dots, \mathcal{A}_n$ of $\mathcal{F}$ are *independent* if

  $$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \ldots \cdot \mathbb{P}(A_n) \quad \text{for } A_k \in \mathcal{A}_k.$$

  Random variables $X_1, X_2, \dots, X_n$ are independent if the σ-algebras $\sigma(X_1), \sigma(X_2), \dots, \sigma(X_n)$ are independent

  $$\begin{aligned} \Leftrightarrow & \mathbb{P}(\{X_1 \in B_1\} \cap \{X_2 \in B_2\} \cap \ldots \cap \{X_n \in B_n\}) \\ & = \mathbb{P}(X_1 \in B_1) \cdot \mathbb{P}(X_2 \in B_2) \cdot \ldots \cdot \mathbb{P}(X_n \in B_n) \quad \text{for } B_k \in \mathcal{B}_{\mathbb{R}}. \end{aligned}$$

A couple of useful notations:

- $\mathbb{E}(X; E) = \int_E X \mathrm{d}\mathbb{P} = \int \mathbb{1}_E X \mathrm{d}\mathbb{P}$.
- A random variable X induces a measure on $\mathbb{R}$ by $\mu_X(B) = \mathbb{P}(X^{-1}(B))$ where $B \in \mathcal{B}_{\mathbb{R}}$ and $\mathcal{B}_{\mathbb{R}}$ is the Borel σ-algebra on $\mathbb{R}$. The measure μ_X is called *distribution* (or *law*) of X. When X and Y induce the same measure, we say that X and Y are equal in distribution and use the notation

  $$X \overset{\mathrm{d}}{=} Y.$$

1.2 Conditional expected value

Definition 1.2.1. Let X be a $L^1(\mathbb{P}, \mathcal{F})$ random variable and let $\mathcal{G} \subset \mathcal{F}$ be a σ-algebra. The *conditional expected value of X given $\mathcal{G}$* is defined to be any random variable Y such that Y is (i) $\mathcal{G}$-measurable and (ii) for any $E \in \mathcal{G}$

$$\int_E X \mathrm{d}\mathbb{P} = \int_E Y \mathrm{d}\mathbb{P}.$$

We then use the notation $\mathbb{E}(X|\mathcal{G})$ for the conditional expected value and any such Y is called a *version* of $\mathbb{E}(X|\mathcal{G})$.

Proposition 1.2.2. *The conditional expected value exists and is unique in the sense that if Y and Y' satisfy (i) and (ii) then $Y = Y'$ almost surely. Also the conditional expected value is integrable.*

Proof. Let $G = \{Y \geq 0\}$ which is $\mathcal{G}$-measurable. Then

$$\mathbb{E}(|Y|) = \int_G Y \mathrm{d}\mathbb{P} - \int_{G^c} Y \mathrm{d}\mathbb{P} = \int_G X \mathrm{d}\mathbb{P} - \int_{G^c} X \mathrm{d}\mathbb{P} \leq \|X\|_1,$$

where G^c is the complement $\Omega \setminus G$ of G. Therefore $\mathbb{E}(|Y|) < \infty$.

Existence follows from Radon–Nikodym theorem:

$$E \mapsto \int_E X \mathrm{d}\mathbb{P}$$

is a signed measure, which is absolutely continuous with respect to $\mathbb{P}$. Then the Radon-Nikodym derivative Y of that measure satisfies the properties of the conditional expected value.

Uniqueness: If Y and Y' are version of $\mathbb{E}(X|\mathcal{G})$, then let $E = \{Y > Y'\}$. Then if $\mathbb{P}(E) > 0$, $\int_E Y \mathrm{d}\mathbb{P} > \int_E Y' \mathrm{d}\mathbb{P}$ which is a contradiction. Hence $\mathbb{P}(\{Y = Y'\}) = 1$. □

Intuitively $\mathbb{E}(X|\mathcal{G})$ should be thought to the best guess of the value of X given the information contained in $\mathcal{G}$.

Example 1.2.3. (*Perfect information*) If X is $\mathcal{G}$ measurable then $\mathbb{E}(X|\mathcal{G}) = X$.

Example 1.2.4. (*No information*) If X is independent of $\mathcal{G}$ then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}X$.

Example 1.2.5. (*Relation to the usual conditional expected value*) Let $\Omega_1, \Omega_2, \dots$ be a finite or countably infinite disjoint partition of Ω into $\mathcal{F}$-measurable sets, each of which has positive probability. If $\mathcal{G}$ is the σ-algebra generated by $\Omega_1, \Omega_2, \dots$ then

$$\mathbb{E}(X|\mathcal{G}) = \frac{\mathbb{E}(X; \Omega_k)}{\mathbb{P}(\Omega_k)} \quad \text{on } \Omega_k.$$

Note that $\mathcal{G} = \{\bigcup_{k \in I} \Omega_k \ : \ I \subset \mathbb{N}\}$.

We list next some properties of conditional expected value.

Theorem 1.2.6. *Let X, Y be $L^1(\mathbb{P}, \mathcal{F})$ random variables and $a, b \in \mathbb{R}$ and $\mathcal{G}, \mathcal{G}_1, \mathcal{G}_2 \subset \mathcal{F}$ be σ-algebras. Then*

1. $\mathbb{E}(aX + bY|\mathcal{G}) = a\mathbb{E}(X|\mathcal{G}) + b\mathbb{E}(Y|\mathcal{G})$
2. $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$
3. $\mathbb{E}(X\,Y|\mathcal{G}) = Y\,\mathbb{E}(X|\mathcal{G})$ *if Y is $\mathcal{G}$-measurable*
4. (Tower property) $\mathbb{E}(\mathbb{E}(X|\mathcal{G}_2)|\mathcal{G}_1) = \mathbb{E}(X|\mathcal{G}_1)$ *if* $\mathcal{G}_1 \subset \mathcal{G}_2$
5. (Jensen's inequality) *If $\phi : \mathbb{R} \to \mathbb{R}$ is convex and $\mathbb{E}(|\phi(X)|) < \infty$ then $\phi(\mathbb{E}(X|\mathcal{G})) \leq \mathbb{E}(\phi(X)|\mathcal{G})$.*
6. $|\mathbb{E}(X|\mathcal{G})| \leq \mathbb{E}(|X||\mathcal{G})$ *and when* $\mathbb{E}(|X|^2) < \infty$, $|\mathbb{E}(X|\mathcal{G})|^2 \leq \mathbb{E}(|X|^2|\mathcal{G})$
7. *If $X_n \to X$ in $L^2(\mathbb{P}, \mathcal{F})$ then $\mathbb{E}(X_n|\mathcal{G}) \to \mathbb{E}(X|\mathcal{G})$ in $L^2(\mathbb{P}, \mathcal{F})$.*

The following notation is sometimes used: if X and Y are random variables and $\sigma(Y)$ is the σ-algebra generated by Y, then $\mathbb{E}(X|Y)$ means the same as $\mathbb{E}(X|\sigma(Y))$.

1.3 Stochastic processes

Let's use the following notation: $\mathbb{N} = \{1, 2, 3, \dots\}$, $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ and $\mathbb{R}_+ = [0, \infty)$.

Definition 1.3.1. A *stochastic process* is a collection of random variables X_t indexed by a ordered set I. A notation $(X_t)_{t \in I}$ is used for a stochastic process.

Almost always $I = \mathbb{R}_+$ or $I = \mathbb{Z}_+$. Since t is regarded as time, we call the process in those cases *continuous time stochastic process* and *discrete time stochastic process*, respectively. On this course usually $I = \mathbb{R}_+$.

The mapping $t \mapsto X_t(\omega)$ is called the *path* of $(X_t)_{t \in I}$. For continuous time processes the path regularity properties are usually essential already when defining the process (as in the definition of Brownian motion below).

Remember that X is a normally distributed with mean μ and variance σ^2 if and only if

$$\mathbb{P}(X \in A) = \int_A \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x$$

for any Borel subset A of $\mathbb{R}$.

Definition 1.3.2. A stochastic process $(B_t)_{t\geq 0}$ is called a *(standard one-dimensional) Brownian motion* if $B_0 = 0$ and

1. $B_{t_1} - B_{s_1}, B_{t_2} - B_{s_2}, \ldots, B_{t_n} - B_{s_n}$ are independent for any $n \in \mathbb{N}$ and for any $0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \ldots \leq s_n < t_n$.
2. For any $s, t \geq 0$, $B_{s+t} - B_s$ is normally distributed with mean 0 and variance t.
3. With probability one, $t \mapsto B_t$ is continuous.

Remark. We say that the process has *independent* and *stationary* increments, if the properties 1. and 2. hold, respectively.

Remark. The third property is best understood in the following way: Let P be a property that a function might or might not have, for example, continuity, differentiability etc. A process $(X_t)_{t\in\mathbb{R}_+}$ has property P with probability one (almost surely) if there exist $E \in \mathcal{F}$ such that $\mathbb{P}(E) = 1$ and $E \subset \{\omega \in \Omega : t \mapsto X_t(\omega)$ has property $P\}$. Note that $\{t \mapsto X_t(\omega)$ has property $P\}$ need not lie in $\mathcal{F}$.

The "canonical" probability space for Brownian motion is the space of continuous functions $C(\mathbb{R}_+)$ with a certain Borel probability measure $\mathbb{P}$ and where the Brownian motion is the coordinate map $B_t(\omega) = \omega_t$. As soon as Brownian motion exists in some probability space, its distribution in $C(\mathbb{R}_+)$ defines the 'canonical" Brownian motion.

Theorem 1.3.3. *Brownian motion exists.*

There are many ways to construct Brownian motion. One of them using so called *Brownian bridge* is left as an exercise, see Problem sheet 2. It's the same idea that was used in the original construction by Paul Lévy.

A *standard d-dimensional Brownian motion* is a $\mathbb{R}^d$-valued stochastic process $(B_t^{(1)}, \ldots, B_t^{(d)})$ where $B_t^{(1)}, \ldots, B_t^{(d)}$ are independent standard one-dimensional Brownian motions.

Exercise. Let $(B_t)_{t\in\mathbb{R}_+}$ be a standard d-dimensional Brownian motion and let $A : \mathbb{R}^d \to \mathbb{R}^d$ be a orthogonal transformation. Show that $A\,B_t$ is a standard d-dimensional Brownian motion. Show also that $(B_t)_{t\in\mathbb{R}_+}$ satisfies *Brownian scaling*: if $r > 0$ then $Y_t = r^{-1/2} B_{rt}$ is a standard d-dimensional Brownian motion.

The following theorem shows that the assumption that the increments are normal is partly redundant in the definition of Brownian motion.

Theorem 1.3.4. *If $(X_t)_{t\in\mathbb{R}_+}$ is a continuous stochastic process which has independent and stationary increments, then there exists a standard one-dimensional Brownian motion $(B_t)_{t\in\mathbb{R}_+}$ and real numbers $\alpha \geq 0$ and β such that $X_t = \alpha B_t + \beta t$.*

Remark. The process of the form $X_t = \alpha B_t + \beta t$ is called *Brownian motion with drift.*

We'll return to this characterization of Brownian motion later, because it will be an essential input for the motivation of studying SLEs.

Definition 1.3.5. A *filtration* on $(\Omega, \mathcal{F})$ is a collection $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ of sub-σ-algebras $\mathcal{F}_t \subset \mathcal{F}$ such that for each $0 \leq s < t$, $\mathcal{F}_s \subset \mathcal{F}_t$.

A filtration can be though as increasing information on the probability space. For example, the σ-algebras generated by a Brownian motion $(B_t)_{t\in\mathbb{R}_+}$, i.e. $\mathcal{F}_t^B = \sigma(B_s,\, 0 \leq s \leq t)$, form a filtration $(\mathcal{F}_t^B)_{t\in\mathbb{R}_+}$.

Definition 1.3.6. A stochastic process $(X_t)_{t\in\mathbb{R}_+}$ on $(\Omega, \mathcal{F})$ is *adapted* to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ if for each $t \geq 0$, X_t is $\mathcal{F}_t$-measurable.

We will make the following more restrictive definition of Brownian motion.

Definition 1.3.7. A process $(B_t)_{t\geq 0}$ is called a *(standard one-dimensional) Brownian motion with respect to the filtration* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ if it is adapted to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$, $B_0 = 0$ and

1. $B_t - B_s$ are independent from $\mathcal{F}_s$ for any $0 \leq s < t$,

2. $B_t - B_s$, $0 \le s < t$, is normally distributed with mean 0 and variance $t - s$
3. With probability one, $t \mapsto B_t$ is continuous.

Remark. We can always weaken this to $(\mathcal{F}_t^B)_{t\in\mathbb{R}_+}$, the filtration generated by the Brownian motion itself, and therefore any Brownian motion in the sense of Definition 1.3.7 is also a Brownian motion in the sense of Definition 1.3.2. But this definition is useful for instance when we are considering two (independent) Brownian motions $B^{(1)}$ and $B^{(2)}$ in the same probability space and $(\mathcal{F}_t) = (\mathcal{F}_t^{(B^{(1)},B^{(2)})})$.

Definition 1.3.8. Let $p \ge 1$. Define the *p'th variation* of a process $(X_t)_{t\in\mathbb{R}_+}$ as the process

$$V_X^{(p)}(t) = \lim_{\mathrm{mesh}(\pi)\to 0} \sum_{k=0}^{m(\pi)-1} |X_{t_{k+1}} - X_{t_k}|^p$$

where π is a partitions of $[0,t]$ of the form $\pi = \{0 = t_0 < t_1 < \ldots < t_{m(\pi)} = t\}$ and the limit is the *limit in probability* as $\mathrm{mesh}(\pi) = \max_k(t_{k+1} - t_k) \to 0$ in the sense that that for each $\varepsilon > 0$ there exist $\delta > 0$ such that

$$\mathbb{P}\left(\left|\sum_{k=0}^{m(\pi)-1} |X_{t_{k+1}} - X_{t_k}|^p - V_X^{(p)}(t)\right| \ge \varepsilon\right) < \varepsilon$$

when $\mathrm{mesh}(\pi) < \delta$. We call the first variation ($p = 1$) as *total variation* and the second variation ($p = 2$) as *quadratic variation.*

Proposition 1.3.9. *The quadratic variation of a Brownian motion exist and* $V_B^{(2)}(t) = t$.

Proof. Let $\varepsilon > 0$ and π be a partitioning with $\mathrm{mesh}(\pi) < (2t)^{-1}\varepsilon^3$. Let $\Delta_k = (B_{t_{k+1}} - B_{t_k})^2 - (t_{t+1} - t_k)$

$$\begin{aligned}
&\mathbb{E}\left(\left(\sum_{k=0}^{m(\pi)-1} (B_{t_{k+1}} - B_{t_k})^2 - t\right)^2\right) = \mathbb{E}\left(\left(\sum_{k=0}^{m(\pi)-1} \Delta_k\right)^2\right)\\
=&\mathbb{E}\left(\sum_k \Delta_k^2\right) + 2\mathbb{E}\left(\sum_{j<k} \Delta_j\Delta_k\right) = \sum_k \mathbb{E}\left(\Delta_k^2\right) + 2\sum_{j<k} \underbrace{\mathbb{E}\left(\Delta_j\Delta_k\right)}_{\text{by inpendence, }=0}\\
=&\mathbb{E}((N^2-1)^2)\sum_k (t_{k+1} - t_k)^2 \le 2\mathrm{mesh}(\pi)t.
\end{aligned}$$

Here $N \sim N(0,1)$ and we used the scaling property of Brownian motion. Hence

$$\mathbb{P}\left(\left|\sum_{k=0}^{m(\pi)-1} (B_{t_{k+1}} - B_{t_k})^2 - t\right| \ge \varepsilon\right) \le \frac{2\mathrm{mesh}(\pi)t}{\varepsilon^2} < \varepsilon \tag{1.1}$$

□

The above proof and Borel–Cantelli lemma will give that almost surely Brownian motion is not a finite variation process.

Lemma 1.3.10 (Borel–Cantelli)**.** *Let A_k, $k \in \mathbb{N}$ be a sequence of events. Define $\{\omega : \omega \in A_k \text{ i.o.}\}$, where i.o. stands for* infinitely often*, as the event $\bigcap_{k=1}^\infty \bigcup_{j=k}^\infty A_j$. If $\sum_{k=1}^\infty \mathbb{P}(A_k) < \infty$ then*

$$\mathbb{P}(A_n \text{ i.o.}) = 0.$$

The total variation of a Brownian motion is almost surely infinite in the sense that if take the limit along the sequence of dyadic partitionings of $[0,t]$

$$\pi_n = \{t\,k\,2^{-n} : k = 0, 1, 2, \ldots, 2^n\} = \{t_0 < t_1 < \ldots < t_{2^n}\},$$

then

$$\lim_{n\to\infty} \sum_{t_k\in\pi_n, k\leq 2^n-1} |B_{t_{k+1}} - B_{t_k}| = \infty$$

almost surely. Namely, if we denote $\mathbb{P}(E(\pi))$ the left-hand side of (1.1), then $\sum_n \mathbb{P}(E(\pi_n)) < \infty$ and hence

$$\sum_{t_k\in\pi_n, k\leq 2^n-1} (B_{t_{k+1}} - B_{t_k})^2 \to t$$

almost surely. Take any ω for which this convergence happens. Then

$$\underbrace{\sum_{t_k\in\pi_n, k\leq 2^n-1} (B_{t_{k+1}}(\omega) - B_{t_k}(\omega))^2}_{\to t} \leq \underbrace{\text{mesh}(\pi_n)}_{\to 0} \sum_{t_k\in\pi_n, k\leq 2^n-1} |B_{t_{k+1}}(\omega) - B_{t_k}(\omega)|$$

which implies that the total variation is infinite for such ω.

1.4 Stochastic integration

1.4.1 Motivation of stochastic integral

The goal of this section is to define a process X_t which can be interpreted as the integral

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

It is important because of the following reasons:

- It is tool for generating new stochastic processes out of Brownian motion.
- Coordinate changes such as $f(B_t)$ turn out to have extremely useful representation using the above intagral.
- Appears in many applications, since $\mathrm{d}B_t$ represents some kind of independent and stationary noise.

The integral doesn't exist pathwise as a Riemann-Stieltjes (or similar) integral even for a continuous f, because the Brownian motion doesn't have finite total variation. For example we will see that

$$\int_0^t B_s \mathrm{d}B_s \neq \frac{1}{2}B_t^2$$

and therefore the usual integration by parts formula can't hold.

1.4.2 Stochastic integral

In this section $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is a filtration and $(B_t)_{t\in\mathbb{R}_+}$ is a standard one-dimensional Brownian motion with respect to $\mathcal{F}_t$.

First we need to define the correct set of integrands f.

Definition 1.4.1. A stochastic process $(X_t)_{t\in\mathbb{R}_+}$ is *measurable* if the mapping $(t,\omega) \mapsto X_t(\omega)$ is $\mathcal{B}_\mathbb{R} \times \mathcal{F}$-measurable.

Definition 1.4.2. Let $T > 0$. We define $\mathcal{L}^2$ to be the set of measurable, adapted processes f that satisfy

$$\mathbb{E}\left(\int_0^T f(t,\cdot)^2 \mathrm{d}t\right) < \infty \tag{1.2}$$

and we call $f \in \mathcal{L}^2$ *simple* if f can be written in the form

$$f(t,\omega) = \sum_{k=0}^{n-1} X_k(\omega) \mathbb{1}_{[t_k,t_{k+1})}(t) \tag{1.3}$$

where $0 \leq t_0 < t_1 < t_2 \ldots < t_n \leq T$ and X_k is a $\mathcal{F}_{t_k}$-measurable, square integrable random variable.

Remark. The above class could be called as $\mathcal{L}^2(T)$ and then we could set $f \in \mathcal{L}^2$ if and only if $f \in \mathcal{L}^2(T)$ for any $T > 0$. However, we don't make a big difference between $\mathcal{L}^2(T)$ and $\mathcal{L}^2$, and consequently, we use the notation $\mathcal{L}^2$ for both classes.

Remark. Note that $\mathcal{L}^2$ is a closed subspace of $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$.

We would like to define a mapping $f \mapsto I[f]$ which we later denote by

$$I[f](\omega) = \int_0^T f(t,\omega)\mathrm{d}B_t(\omega).$$

If that notation makes any sense we have to define

$$I[\mathbb{1}_{[s,t)}] = B_t - B_s$$

for any $0 \le s < t \le T$. Therefore for any f which is of the form (1.3) it is natural to define by linearity that

$$I[f] = \sum_{k=0}^{n-1} X_k(B_{t_{k+1}} - B_{t_k}).$$

Therefore I is now defined for any simple $f \in \mathcal{L}^2$ and it turns out that there is a unique L^2-continuous extension of it to the whole $\mathcal{L}^2$. Namely, we first observe the following isometry.

Proposition 1.4.3 (Itô isometry for simple processes)**.** *For any bounded, simple $f \in \mathcal{L}^2$*

$$\mathbb{E}(I[f]^2) = \mathbb{E}\left(\int_0^T f(t,\cdot)^2\mathrm{d}t\right).$$

Proof. Let's calculate both sides explicitly for a bounded, simple $f \in \mathcal{L}^2$ of the form (1.3). Notice that $f^2 = \sum_{k=0}^{n-1} X_k^2 \mathbb{1}_{[t_k,t_{k+1})}$ and hence

$$\mathbb{E}\left(\int_0^T f(t,\cdot)^2\mathrm{d}t\right) = \sum_{k=0}^{n-1}\mathbb{E}(X_k^2)(t_{k+1}-t_k).$$

On the other hand

$$\mathbb{E}(I[f]^2) = \sum_k \mathbb{E}(X_k^2)\mathbb{E}\left((B_{t_{k+1}} - B_{t_k})^2\right) + \sum_{k<l}\mathbb{E}(X_kX_l(B_{t_{k+1}} - B_{t_k})(B_{t_{l+1}} - B_{t_l}))$$

which gives the claim after we notice that

$$\mathbb{E}(X_kX_l(B_{t_{k+1}} - B_{t_k})(B_{t_{l+1}} - B_{t_l})) = \mathbb{E}(X_kX_l(B_{t_{k+1}} - B_{t_k}))\mathbb{E}(B_{t_{l+1}} - B_{t_l}) = 0$$

for $k < l$ and that $\mathbb{E}\left((B_{t_{k+1}} - B_{t_k})^2\right) = t_{k+1} - t_k$. □

Next we prove that the simple processes are dense in $\mathcal{L}^2$.

Proposition 1.4.4. *For each $f \in \mathcal{L}^2$, there exist a sequence of bounded, simple $f_n \in \mathcal{L}^2$ such that*

$$\mathbb{E}\left(\int_0^T (f(t,\cdot) - f_n(t,\cdot))^2\mathrm{d}t\right) \to 0,$$

i.e. f_n converges to f in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$.

Remark. We divide the proof in three steps. The first and last steps are the most important for us, because we will mostly only consider continuous processes as integrands.

Proof. **Bounded continuous** $f \in \mathcal{L}^2$: Take any sequence of partitions π_n such that $\text{mesh}(\pi_n) \to 0$ as $n \to \infty$ and define a sequence of bounded, simple processes $f_n \in \mathcal{L}^2$ as

$$f_n(t,\omega) = \sum_{k=0}^{m(\pi)-1} f(t_k,\omega)\mathbb{1}_{[t_k,t_{k+1})}(t)$$

when π_n is $0 = t_0 < t_1 < \ldots < t_{m(\pi_n)} = T$. Then

$$\sup_{t\in[0,T]} |f(t,\omega) - f_n(t,\omega)| \leq \sup_{s,t\in[0,T]\,:\,|s-t|\leq \text{mesh}(\pi_n)} |f(t,\omega) - f(s,\omega)|$$

By continuity the right-hand side goes to zero almost surely. Since $|f| \leq C < \infty$ for some constant C, we can apply the dominated convergence theorem (DCT) to show that the right-hand side goes to zero also in $L^2(\mathrm{d}\mathbb{P})$.

Hence

$$\mathbb{E}\left(\int_0^T |f(t,\cdot) - f_n(t,\cdot)|^2 \mathrm{d}t\right) \leq \mathbb{E}\left(T \sup_{t\in[0,T]} |f(t,\cdot) - f_n(t,\cdot)|^2\right) \to 0.$$

Bounded $g \in \mathcal{L}^2$: Take a sequence of continuous functions $\psi_n : \mathbb{R} \to \mathbb{R}$ such that

1. $\psi_n \geq 0$
2. $\psi_n(x) = 0$ when $x \notin (-1/n, 0)$
3. $\int_{-\infty}^{\infty} \psi_n(x) = 1$

Define a sequence of bounded, continuous processes $g_n \in \mathcal{L}^2$ as

$$g_n(s,\omega) = \int_0^t \psi_n(s-t)g(s,\omega)\mathrm{d}s$$

The sequence (ψ_n) forms a *approximate identity* and by standard properties of such sequences,

$$\int_0^T (g_n(s,\omega) - g(s,\omega))^2 \mathrm{d}s \to 0.$$

The measurability requirements of $\mathcal{L}^2$ are slightly tricky, and we omit such details here. By DCT, $g_n \to g$ in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$.

General $h \in \mathcal{L}^2$: Define a sequence of bounded processes $h_n \in \mathcal{L}^2$ as

$$h_n(t,\omega) = \begin{cases} -n & \text{if } h(t,\omega) < -n \\ h(t,\omega) & \text{if } h(t,\omega) \in [-n,n] \\ n & \text{if } h(t,\omega) > n \end{cases}$$

Then by DCT, $h_n \to h$ in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$. □

If $f_n \in \mathcal{L}^2$ is a sequence of simple, bounded processes converging to f, then f_n is a Cauchy sequence in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$ and hence by the isometry property $I[f_n]$ is a Cauchy sequence if $L^2(\mathrm{d}P)$ and hence it converges. Therefore we can define $I[f] = \lim_n I[f_n]$. Note that this limit doesn't depend on the choice of f_n: if f_n and f'_n are two such sequences, then $f_n - f'_n$ goes to zero in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$ and hence by isometry, $\lim_n I(f_n) = \lim_n I(f'_n)$ almost surely. This is summarized in the following definition.

Definition 1.4.5. For any $f \in \mathcal{L}^2$, the *stochastic integral* (or *Itô integral*) is defined to be

$$\int_0^T f(t,\omega)\mathrm{d}B_t(\omega) = I[f](\omega) = (\lim_n I[f_n])(\omega) \tag{1.4}$$

where the limit is in $L^2(\mathbb{P})$ and $f_n \in \mathcal{L}^2$ is any sequence of bounded, simple processes converging to f in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$. The integral is defined almost surely.

Corollary 1.4.6 (Itô isometry for $\mathcal{L}^2$)**.** *For any $f \in \mathcal{L}^2$*

$$\mathbb{E}\left(\left(\int_0^T f(t,\cdot)\mathrm{d}B_t\right)^2\right) = \mathbb{E}\left(\int_0^T f(t,\cdot)^2\mathrm{d}t\right).$$

Corollary 1.4.7. *If $f_n \in \mathcal{L}^2$, $f \in \mathcal{L}^2$ and $f_n \to f$ in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$ then $\int_0^T f_n \mathrm{d}B_t \to \int_0^T f \mathrm{d}B_t$ in $L^2(\mathbb{P})$.*

Example 1.4.8. We'll show that

$$\int_0^t B_s \mathrm{d}B_s = \frac{1}{2}B_t - \frac{1}{2}t$$

Let π_n be a sequence of partitions of $[0,t]$ such that $\mathrm{mesh}(\pi_n) \to 0$. By the above, the sequence of processes $f_n(s,\omega) = \sum_{t_j\in\pi_n} B_{t_j}(\omega)\mathbb{1}_{[t_j,t_{j+1})}(s)$ is a reasonable choice for a discretization of the integrand. Since

$$\mathbb{E}\left(\int_0^t (B_s - f_n(s,\cdot))\mathrm{d}s\right) = \mathbb{E}\left(\sum_j \int_{t_j}^{t_{j+1}} (B_s - B_{t_j})^2 \mathrm{d}s\right) = \sum_j \frac{1}{2}(t_{j+1}-t_j)^2 \to 0$$

as $n \to \infty$, then by Corollary 1.4.7, $\int_0^t B_s \mathrm{d}B_s = \lim \int_0^t f_n \mathrm{d}B_s = \lim \sum_j B_{t_j}(B_{t_{j+1}} - B_{t_j})$. Now notice that

$$B_{t_{j+1}}^2 - B_{t_j}^2 = (B_{t_{j+1}} - B_{t_j})^2 + 2B_{t_j}(B_{t_{j+1}} - B_{t_j})$$

and thus

$$\sum_j B_j(B_{t_{j+1}} - B_{t_j}) = \frac{1}{2}B_t^2 - \frac{1}{2}\sum_j (B_{t_{j+1}} - B_{t_j})^2$$

and the second term on the right converges in L^2 to the quadratic variation of Brownian motion which we already showed to be t.

The following proposition states some properties of the stochastic integral. Those properties hold for the simple processes and hence hold also for any limit of a sequence of simple processes.

Proposition 1.4.9. *Let $f, g \in \mathcal{L}^2$ and let $0 \le S < U < T$. Then*

1. $\int_S^T f \mathrm{d}B_t = \int_S^U f \mathrm{d}B_t + \int_U^T f \mathrm{d}B_t$
2. $\int_S^T (af + bg)\mathrm{d}B_t = a\int_S^T f \mathrm{d}B_t + b\int_S^T g \mathrm{d}B_t$
3. $\mathbb{E}(\int_S^T f \mathrm{d}B_t) = 0$
4. $\int_S^T f \mathrm{d}B_t$ *is $\mathcal{F}_T$-measurable*

1.4.3 Martingales

The following concept will be extremely useful during this course.

Definition 1.4.10. A stochastic process $(M_t)_{t\in\mathbb{R}_+}$ is called a *(continuous-time) martingale* with respect to a filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ if

1. M_t is $\mathcal{F}_t$-measurable for each $t \ge 0$,
2. $\mathbb{E}(|M_t|) < \infty$ for each $t \ge 0$,
3. $\mathbb{E}(M_t|\mathcal{F}_s) = M_s$ for each $0 \le s < t$.

If the last property is replaced by $\mathbb{E}(M_t|\mathcal{F}_s) \ge M_s$, the process is called *submartingale*, and if the last property is replaced by $\mathbb{E}(M_t|\mathcal{F}_s) \le M_s$, the process is called *supermartingale*.

Quite many results for martingales are proved using discrete-time martingales.

Definition 1.4.11. A *discrete-time filtration* on $(\Omega, \mathcal{F})$ is a collection $(\mathcal{F}_t)_{t\in\mathbb{Z}_+}$ of sub-σ-algebras $\mathcal{F}_t \subset \mathcal{F}$ such that for each $t \in \mathbb{Z}_+$, $\mathcal{F}_t \subset \mathcal{F}_{t+1}$.

A stochastic process $(M_t)_{t\in\mathbb{Z}_+}$ is called a *(discrete-time) martingale* with respect to a filtration $(\mathcal{F}_t)_{t\in\mathbb{Z}_+}$ if

1. M_t is $\mathcal{F}_t$-measurable for each $t \in \mathbb{Z}_+$,
2. $\mathbb{E}(|M_t|) < \infty$ for each $t \in Z_+$,
3. $\mathbb{E}(M_{t+1}|\mathcal{F}_t) = M_t$ for each $t \in \mathbb{Z}_+$.

If the last property is replaced by $\mathbb{E}(M_{t+1}|\mathcal{F}_t) \geq M_t$, the process is called *submartingale*, and if the last property is replaced by $\mathbb{E}(M_{t+1}|\mathcal{F}_t) \leq M_t$, the process is called *supermartingale*.

Example 1.4.12. Let $X \in L^1(\mathbb{P}, \mathcal{F})$ and let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration. Then $M_t = \mathbb{E}(X|\mathcal{F}_t)$ is a martingale: 1 holds by the definition of conditional expected value, 2 holds by items 2 and 6 of Theorem 1.2.6 and 3 holds by item 4 of Theorem 1.2.6.

Example 1.4.13. Let $X_0, X_1, X_2, \ldots$ be a sequence of independent integrable random variables such that $\mathbb{E}(X_k) = 0$ for each k and let $\mathcal{F}_n = \sigma(X_0, X_1, X_2, \ldots, X_n)$. Then $(M_n)_{n\in\mathbb{Z}_+}$ defined by

$$M_n = \sum_{k=0}^{n} X_k$$

is a martingale with respect to $(\mathcal{F}_n)_{n\in\mathbb{Z}_+}$.

Example 1.4.14 (The origin of the name *martingale*)**.** There is a *gambling strategy* called martingale. Consider a gambler that is playing roulette, where the outcome is either red or black with probability 1/2 each. After a loss the gambler always doubles his bet and keeps playing until the first time when he wins. After that he stops playing. If the first bet is x, then the gambler is sure to win x by this strategy! Do you see any problem with the martingale strategy? This is related to the previous example when we consider $X_0, X_1, \ldots$ such that $X_0 = 0$, $X_1 = \hat{X}_1$ and

$$X_k = \hat{X}_k \, \mathbb{1}_{\{\text{no wins during rounds } 1,2,\ldots,k-1\}}$$

for $k \geq 2$, where $\hat{X}_k$ are independent and $\mathbb{P}(\hat{X}_k = \pm x\, 2^k) = 1/2$. Then M_n is the wealth of the gambler after n rounds relative to the wealth at time zero.

Example 1.4.15. Let $X_0, X_1, X_2, \ldots$ be a sequence of independent integrable random variables such that $\mathbb{E}(X_k) = 1$ for each k and let $\mathcal{F}_n = \sigma(X_0, X_1, X_2, \ldots, X_n)$. Then $(M_n)_{n\in\mathbb{Z}_+}$ defined by

$$M_n = \prod_{k=0}^{n} X_k$$

is a martingale with respect to $(\mathcal{F}_n)_{n\in\mathbb{Z}_+}$.

Example 1.4.16. There are many martingales related to Brownian motions. During the course we will check the following formulas

$$\mathbb{E}(B_t \,|\, \mathcal{F}_s) = B_s$$
$$\mathbb{E}(B_t^2 - t \,|\, \mathcal{F}_s) = B_s^2 - s$$
$$\mathbb{E}\left(\exp\left(\theta B_t - \frac{\theta^2}{2} t\right) \,\middle|\, \mathcal{F}_s\right) = \exp\left(\theta B_s - \frac{\theta^2}{2} s\right).$$

The first result that we need about martingales for the Itô integral is the next inequality. Its proof is given in the exercises.

Theorem 1.4.17 (Doob's maximal inequality)**.** *Suppose that $(M_t)_{t\in\mathbb{R}_+}$ is a continuous martingale. Then for each $p \geq 1$, $T > 0$*

$$\mathbb{P}\left(\sup_{0\leq s\leq t} |M_s| \geq \lambda\right) \leq \frac{1}{\lambda^p} \mathbb{E}(|M_T|^p).$$

1.4.4 Itô integral as a process

By the above, we try to define a process X_t such that $X_t = \int_0^t f(\cdot, s)\mathrm{d}B_s$ for every t. The problem in just defining $X_t = I[f\mathbb{1}_{[0,t]}]$ is that the for each fixed t, X_t is defined in a set of probability one, say, in Ω_t, but it is possible that the probability of the uncountable intersection $\bigcap_t \Omega_t$ is strictly less than 1 or even that $\bigcap_t \Omega_t$ is not an event (a measurable set). Therefore we define X_t that way in a countable set of t and then extend by continuity to other t's as in the following theorem.

Theorem 1.4.18. *For each $f \in \mathcal{L}^2$ there exists a continuous square integrable martingale $(X_t)_{t\in\mathbb{R}_+}$ such that for each t, $X_t = \int_0^t f(\cdot, s)\mathrm{d}B_s$ almost surely.*

Remark. The process $(X_t)_{t\in\mathbb{R}_+}$ is unique in the sense that if there is another process $(X'_t)_{t\in\mathbb{R}_+}$ with the same properties, then almost surely $X_t = X'_t$ for all t.

Proof. Fix some $T > 0$. Take a sequence of simple (and bounded) $f_n \in \mathcal{L}^2$ such that $f_n \to f$ in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P}, [0,T] \times \Omega)$ and define $X_t^{(n)} = I[f_n\mathbb{1}_{[0,t]}]$ which is well-defined in whole Ω. If $f_n = \sum a_k \mathbb{1}_{[t_k,t_{k+1})}$, then for $t_l \le t < t_{l+1}$ we have an explicite formula

$$X_t^{(n)} = a_l \cdot (B_t - B_{t_l}) + \sum_{k=0}^{l-1} a_k \cdot (B_{t_{k+1}} - B_{t_k}). \tag{1.5}$$

Clearly $t \mapsto X_t$ is continuous. To show that it is a martingale, notice first that it is adapted because all the random variables on the right of (1.5) are $\mathcal{F}_t$-measurable. Also $\mathbb{E}|X_t^{(n)}| < \infty$, because it is a finite sum of integrable random variables (you can also use Itô isometry), and for $0 \le s < t \le T$ we can assume that $s = t_l$ and $t = t_m$ for some l and m (redefine the "partitioning" of f_n again if necessary) and then

$$\begin{aligned}
\mathbb{E}(X_t^{(n)}|\mathcal{F}_s) &= \mathbb{E}(X_s^{(n)}|\mathcal{F}_s) + \mathbb{E}(\sum_{k=l}^{m-1} a_k \cdot (B_{t_{k+1}} - B_{t_k})|\mathcal{F}_s) \\
&= X_s^{(n)} + \sum_{k=l}^{m-1} \mathbb{E}(a_k \cdot (B_{t_{k+1}} - B_{t_k})|\mathcal{F}_s) \\
&= X_s^{(n)}
\end{aligned}$$

because

$$\begin{aligned}
\mathbb{E}(a_k \cdot (B_{t_{k+1}} - B_{t_k})|\mathcal{F}_s) &= \mathbb{E}(\mathbb{E}(a_k \cdot (B_{t_{k+1}} - B_{t_k})|\mathcal{F}_{t_k})|\mathcal{F}_s) \\
&= \mathbb{E}(a_k \cdot \mathbb{E}((B_{t_{k+1}} - B_{t_k})|\mathcal{F}_{t_k})|\mathcal{F}_s) = 0.
\end{aligned} \tag{1.6}$$

Now since $X_t^{(n)} - X_t^{(m)}$ is a martingale, by Doob's maximal inequality

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\in[0,T]} \left|X_t^{(n)} - X_t^{(m)}\right| \ge \varepsilon\right) &\le \frac{1}{\varepsilon^2}\mathbb{E}(|X_T^{(n)} - X_T^{(m)}|^2) \\
&= \frac{1}{\varepsilon^2}\|f_n - f_m\|^2_{L^2(\mathrm{d}t\times\mathrm{d}\mathbb{P})}
\end{aligned}$$

for any $\varepsilon > 0$. Choose a subsequence n_k such that $\|f_{n_{k+1}} - f_{n_k}\|^2_{L^2(\mathrm{d}t\times\mathrm{d}\mathbb{P})} \le 2^{-3k}$ and use the previous estimate for $\varepsilon = 2^{-k}$ to get

$$\mathbb{P}\left(\sup_{t\in[0,T]} \left|X_t^{(n_{k+1})} - X_t^{(n_k)}\right| \ge 2^{-k}\right) \le 2^{-k}$$

By the Borel–Cantelli lemma, there exist random variable N which is almost surely finite and for $k \ge N(\omega)$

$$\sup_{t\in[0,T]} \left|X_t^{(n_{k+1})} - X_t^{(n_k)}\right| < 2^{-k}.$$

Hence the sequence of the continuous processes $(X_t^{(n_k)})$ converges almost surely uniformly to a continuous process (X_t). Since for any fixed t, $\lim X_t^{(n_k)}$ in $L^2(\mathbb{P})$ is $\int_0^t f \mathrm{d}B_s$ then

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

almost surely. This also shows that (X_t) is adapted and square integrable.

Finally the martingale property of $(X_t^{(n)})$, for any $0 \le s < t \le T$

$$X_s^{(n)} = \mathbb{E}(X_t^{(n)}|\mathcal{F}_s).$$

Since the random variables $X_s^{(n)}$ and $X_s^{(n)}$ converge in $L^2(\mathbb{P})$ to X_s and X_t, respectively, then by Theorem 1.2.6

$$X_s = \mathbb{E}(X_t|\mathcal{F}_s).$$

for any $0 \le s < t \le T$. For the whole $\mathbb{R}_+$, the claim follows from the above by taking a countable sequence $T \nearrow \infty$ and using the uniqueness. □

Remark. The property that we used in (1.6) could be reformulated in the following way: if $(M_t)_{t\in\mathbb{R}_+}$ is a martingale and if $0 \le s \le t \le u$ and Y is a $\mathcal{F}_t$-measurable bounded random variable, then

$$\mathbb{E}(Y(M_u - M_t)\,|\mathcal{F}_s) = 0.$$

We say that the martingale increments $M_u - M_t$ is orthogonal to $\mathcal{F}_t$.

Definition 1.4.19. For any $f \in \mathcal{L}^2$, the *stochastic integral* (or *Itô integral*) is redefined to be the continuous *version* $X_t(\omega)$ of $\int_0^t f(s,\omega)\mathrm{d}B_s(\omega)$ constructed in the previous theorem.

Remark. The processes $(X_t)_{t\in\mathbb{R}_+}$ and $(Y_t)_{t\in\mathbb{R}_+}$ are versions of each other if $\mathbb{P}(X_t = Y_t) = 1$ for each t.

Definition 1.4.20. For any process $X_t = \int_0^t f \mathrm{d}B_s$, define the *quadratic variation process* as

$$\langle X\rangle_t(\omega) = \int_0^t f(s,\omega)^2 \mathrm{d}t.$$

We will later prove that $\langle X\rangle$ is the quadratic variation in the sense of Definition 1.3.8, but before that we state and prove a couple more theorems about the Itô integral as a process.

Theorem 1.4.21. *Let $f \in \mathcal{L}^2$, $X_t = \int_0^t f \mathrm{d}B_s$ and $\langle X\rangle_t$ as above. Then $X_t^2 - \langle X\rangle_t$ is a martingale.*

Proof. We leave as an exercise to check this for bounded, simple $f \in \mathcal{L}^2$. In the general case take a sequence of bounded, simple $f_n \in \mathcal{L}^2$ and define $X_t^{(n)} = \int_0^t f_n \mathrm{d}B_s$. The claim follows easily from the $L^1(\mathbb{P})$ convergence of $(X_t^{(n)})^2 - \langle X^{(n)}\rangle_t$ which implies the $L^1(\mathbb{P})$ convergence of $\mathbb{E}[(X_t^{(n)})^2 - \langle X^{(n)}\rangle_t \,|\, \mathcal{F}_s]$ by the properties 2. and 6. in Theorem 1.2.6. □

Next we define a stopping time which can be taught as the time when something happens such that for each time instant, the question whether this event already occurred or not before or at that time is a "measurable question".

Definition 1.4.22. A random variable $\tau : \Omega \to [0,\infty]$ is called a *stopping time* with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ if for all $t \ge 0$, $\{\omega : \tau(\omega) \le t\} \in \mathcal{F}_t$.

One way to describe the following result is that by that proposition, the pathwise interpretation of the Itô integral makes sense: if two integrands have the same paths up to a stopping time, then the integrals also agree up to that stopping time.

Proposition 1.4.23. *If τ is a stopping time and $f \in \mathcal{L}^2$ and $g \in \mathcal{L}^2$ processes such that $f(t,\omega) = g(t,\omega)$ for any (t,ω) such that $t \le \tau(\omega)$, then for almost all ω*

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

for all $t \le \tau(\omega)$.

Proof. Let $X_t = \int_0^t f(s,\cdot)\mathrm{d}B_s$. It is clearly enough to prove that if τ is a stopping time and $f(t,\omega) = 0$ for $t \le \tau(\omega)$, then for almost all ω, $X_t(\omega) = 0$ for all $t \le \tau(\omega)$.

Assume for a moment that $|f| \le K$. Pick a sequence of simple $f_n \in \mathcal{L}^2$ converging to f in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$. We can assume that $|f_n| \le K$. Write

$$f_n(t,\omega) = \sum_{k=0}^{m_n-1} a_k^{(n)}(\omega)\mathbb{1}_{\left[t_k^{(n)},t_{k+1}^{(n)}\right)}(t)$$

Since it is possible that $f_n(t,\omega) \neq 0$ for some (t,ω) satisfying $t \le \tau(\omega)$, we modify f_n by setting

$$\tilde{f}_n(t,\omega) = \sum_{k=0}^{m_n-1} a_k^{(n)}(\omega)\mathbb{1}_{\left\{\tau<t_k^{(n)}\right\}}(\omega)\mathbb{1}_{\left[t_k^{(n)},t_{k+1}^{(n)}\right)}(t).$$

Notice that $\tilde{f}_n \in \mathcal{L}^2$ (here we need that τ is a stopping time). Now since $f_n\mathbb{1}_{[\tau,\infty)} \to f\mathbb{1}_{[\tau,\infty)} = f$ in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$, to check that $\tilde{f}_n \to f$ in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$ we have to show that $\tilde{f}_n - f_n\mathbb{1}_{[\tau,\infty)} \to 0$ in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$.

Now

$$\begin{aligned}\left|\tilde{f}_n(t,\cdot) - f_n(t,\cdot)\mathbb{1}_{[\tau,\infty)}(t)\right| \le& K\sum_{k=0}^{m-1}\left|\mathbb{1}_{\left\{\tau<t_k^{(n)}\right\}} - \mathbb{1}_{\{\tau\le t\}}\right|\mathbb{1}_{\left[t_k^{(n)},t_{k+1}^{(n)}\right)}(t)\\ \le& K\sum_{k=0}^{m-1}\mathbb{1}_{\left\{t_k^{(n)}\le\tau<t_{k+1}^{(n)}\right\}}\mathbb{1}_{\left[t_k^{(n)},t_{k+1}^{(n)}\right)}(t)\end{aligned}$$

and therefore

$$\begin{aligned}\mathbb{E}\left(\int_0^T\left(\tilde{f}_n - f_n\mathbb{1}_{[\tau,\infty)}\right)^2\mathrm{d}t\right) &\le K^2\sum_{k=0}^{m-1}\mathbb{P}\left(t_k^{(n)} \le \tau < t_{k+1}^{(n)}\right)\int_0^T \mathbb{1}_{\left[t_k^{(n)},t_{k+1}^{(n)}\right)}(t)\mathrm{d}t\\ &\le K^2\,\mathrm{mesh}(\pi_n)\end{aligned}$$

where $\pi_n = \{t_0^{(n)},\dots,t_{m_n}^{(n)}\}$. We can assume that $\mathrm{mesh}(\pi_n) \to 0$ by adding points to the partition if necessary. Since $\tilde{f}_n \to f$ in $L^2(\mathrm{d}t \times \mathrm{d}\mathbb{P})$ and $\int_0^t \tilde{f}_n(s,\cdot)\mathrm{d}B_s = 0$ for $t \le \tau$, and since by the proof of Theorem 1.4.18 there is a subsequence of $\int_0^t \tilde{f}_n(s,\cdot)\mathrm{d}B_s$ that converges almost surely uniformly on $[0,T]$ for $T > 0$, $X_t(\omega) = 0$ almost surely for any $t \le \tau$. □

1.4.5 More general integrands and localization

Often useful notations: $s \wedge t = \min\{s,t\}$ and $s \vee t = \max\{s,t\}$.

At this point, we have the Itô integral defined for any measurable, adapted process f such that

$$\mathbb{E}\left(\int_0^T f^2\mathrm{d}t\right) < \infty$$

for any $T \in (0,\infty)$. However, we would like to have a larger class of processes that includes at least all the continuous processes such as $f(t,\omega) = \exp(B_t(\omega)^3)$ which is an example of a process that doesn't belong to $\mathcal{L}^2$.

Definition 1.4.24. $\mathcal{L}^2_{\mathrm{loc}}$ is definied to be the set of measurable, adapted process f such that

$$\int_0^T f(t,\cdot)^2\,\mathrm{d}t < \infty$$

almost surely for any $T \in (0,\infty)$.

Fix some $f \in \mathcal{L}^2_{\text{loc}}$. Define a stopping time

$$\tau_n(\omega) = \inf\left\{ t \in \mathbb{R}_+ \,:\, \int_0^t f(s,\omega)^2 \mathrm{d}s \geq n \right\}.$$

It follows from $f \in \mathcal{L}^2_{\text{loc}}$, that $\tau_n \nearrow \infty$ almost surely as $n \to \infty$.

Let $f_n(t,\omega) = f(t,\omega)\mathbb{1}_{t \leq \tau_n}$. Then $f_n \in \mathcal{L}^2$ and we can define the Itô integral $X_t^{(n)} = \int_0^t f_n \mathrm{d}B_s$. Since $f_n(t,\omega) = f_m(t,\omega)$ for all (t,ω) such that $t \leq (\tau_n \wedge \tau_m)(\omega)$ and since $\tau_n \wedge \tau_m$ is also a stopping time, by Proposition 1.4.23 for almost all ω,

$$X_t^{(n)}(\omega) = X_t^{(m)}(\omega)$$

for $t \leq (\tau_n \wedge \tau_m)(\omega)$. For each fixed ω, this is really strong convergence: there is a finite $n_0(\omega)$ such that $X_t^{(n)}(\omega) = X_t^{(m)}(\omega)$ for any $n, m \geq n_0(\omega)$. Define now a process $(X_t)_{t\in\mathbb{R}_+}$ on the event $\{\tau_n \nearrow \infty\}$

$$X_t(\omega) = X_t^{(n)}(\omega)$$

where $n \in \mathbb{N}$ is any number satisfying $\tau_n(\omega) \geq t$. The complement of $\{\tau_n \nearrow \infty\}$ has zero probability and there we can define $X_t = 0$ identically, say. Note that: if for any $t > 0$, $X_t(\omega) = X_t^{(n)}(\omega)$ for some n, then $X_s(\omega) = X_s^{(n)}(\omega)$ for all $s \in [0,t]$. Therefore $s \mapsto X_s(\omega)$ is continuous.

Definition 1.4.25. The Itô integral of $f \in \mathcal{L}^2_{\text{loc}}$ is defined as

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

where $n \in \mathbb{N}$ is any number satisfying $\tau_n(\omega) \geq t$ and $X_t^n(\omega)$ is as above,

We will conclude this section by stating a theorem that lists the properties of Itô integral For any continuous process $(X_t)_{t\in\mathbb{R}_+}$ and for any stopping time τ, define a stopped process $(X_t^\tau)_{t\in\mathbb{R}_+}$ by $X_t^\tau = X_{t\wedge\tau}$. The continuity guarantees that X_t^τ is measurable.

Definition 1.4.26. A continuous process $(M_t)_{t\in\mathbb{R}_+}$ adapted to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is called *local martingale* if there exist a sequence of stopping times $0 \leq \tau_1 \leq \tau_2 \leq \ldots$ such that $\mathbb{P}(\tau_k \nearrow \infty) = 1$ and for each k, M^{τ_k} is a martingale. It is a *local square integrable martingale*, if each $(M_t^{\tau_k})_{t\in\mathbb{R}_+}$ is a square integrable martingale.

Theorem 1.4.27. *For any $f \in \mathcal{L}^2_{\text{loc}}$, $X_t = \int_0^t f(s,\cdot)\,\mathrm{d}B_s$ is a continuous local square integrable martingale and $X_t^2 - \langle X\rangle_t$ is a continuous local martingale.*

1.4.6 Quadratic variation of Itô integrals[1]

Remember that above we noticed that $X_t^2 - \langle X\rangle_t$ is a martingale for any $X_t = \int_0^t f \mathrm{d}B_s$, $f \in \mathcal{L}^2$. We can use this property to show that $\langle X\rangle_t$ is the quadratic variation $V_X^{(2)}(t)$ in the sense of Definition 1.3.8. This proposition has a version for $f \in \mathcal{L}^2_{\text{loc}}$ and also then $V_X^{(2)} = \langle X\rangle$.

Theorem 1.4.28. *For any $f \in \mathcal{L}^2$, the Itô integral $X_t = \int_0^t f \mathrm{d}B_s$ has finite quadratic variation and*

$$V_X^{(2)}(t) = \langle X\rangle_t$$

almost surely for any t.

Proof. Assume first that $f \in \mathcal{L}^2$ is such that the Itô integral $X_t = \int_0^t f \mathrm{d}B_s$ and the quadratic variation $\langle X\rangle_t$ are bounded processes, that is, there exists a constant K such that for almost all ω and for all t, $|X_t(\omega)| \leq K$ and $\langle X\rangle_t \leq K$.

Let $t > 0$ and $\pi = \{0 = t_0 < t_1 < \ldots < t_m = t\}$. Define

$$\Delta_k = (X_{t_{k+1}} - X_{t_k})^2 - \langle X\rangle_{t_{k+1}} + \langle X\rangle_{t_k}$$

[1] We didn't go through this proof in the lectures. But it is included here for completeness.

and note that

$$\sum_{k=0}^{m-1} \Delta_k = \sum_{k=0}^{m-1} (X_{t_{k+1}} - X_{t_k})^2 - \langle X \rangle_t.$$

Notice also that for any $0 \leq u \leq t_k$, $\mathbb{E}[\Delta_k|\mathcal{F}_u] = 0$ by the martingale increment orthogonality. Therefore

$$\mathbb{E}\left(\left(\sum_{k=0}^{m-1} \Delta_k\right)^2\right) = \sum_{k=0}^{m-1} \mathbb{E}\left(\Delta_k^2\right).$$

By the inequality $(a+b)^2 \leq 2(a^2+b^2)$,

$$\begin{aligned}\mathbb{E}\left(\left(\sum_{k=0}^{m-1} \Delta_k\right)^2\right) \leq & 2\sum_{k=0}^{m-1} \mathbb{E}\left((X_{t_{k+1}} - X_{t_k})^4\right) \\ & + 2\mathbb{E}(\langle X \rangle_t \sup\{|\langle X \rangle_s - \langle X \rangle_{s'}| : 0 \leq s, s' \leq t, |s-s'| \leq \text{mesh}(\pi)\}).\end{aligned}$$

The second term goes to zero as $\text{mesh}(\pi) \to 0$ by boundedness and continuity of $\langle X \rangle_t$. So it remains to show that

$$\sum_{k=0}^{m-1} \mathbb{E}\left((X_{t_{k+1}} - X_{t_k})^4\right) \to 0$$

as $\text{mesh}(\pi) \to 0$.

We will first show that

$$\mathbb{E}\left(\left(\sum_{k=0}^{m-1} (X_{t_{k+1}} - X_{t_k})^2\right)^2\right) \leq 6K^4 \tag{1.7}$$

By using the martingale property of $(X_t)_{t\in\mathbb{R}_+}$

$$\sum_{k=j}^{m-1} \mathbb{E}\left((X_{t_{k+1}} - X_{t_k})^2 \middle| \mathcal{F}_{t_j}\right) = \sum_{k=0}^{m-1} \mathbb{E}\left(X_{t_{k+1}}^2 - X_{t_k}^2 \middle| \mathcal{F}_{t_j}\right) \leq \mathbb{E}\left(X_{t_m}^2 \middle| \mathcal{F}_{t_j}\right) \leq K^2$$

and therefore

$$\begin{aligned}&\sum_{j=0}^{m-1} \sum_{k=j+1}^{m-1} \mathbb{E}\left((X_{t_{j+1}} - X_{j_k})^2 (X_{t_{k+1}} - X_{t_k})^2\right) \\ &= \sum_{j=0}^{m-1} \mathbb{E}\left((X_{t_{j+1}} - X_{j_k})^2 \sum_{k=j+1}^{m-1} \mathbb{E}\left((X_{t_{k+1}} - X_{t_k})^2 \middle| \mathcal{F}_{t_{j+1}}\right)\right) \\ &\leq K^2 \sum_{j=0}^{m-1} \mathbb{E}\left((X_{t_{j+1}} - X_{j_k})^2\right) \leq K^4\end{aligned}$$

We also have

$$\sum_{k=0}^{m-1} \mathbb{E}\left((X_{t_{k+1}} - X_{t_k})^4\right) \leq 4K^2 \sum_{k=0}^{m-1} \mathbb{E}\left((X_{t_{k+1}} - X_{t_k})^2\right) \leq 4K^4.$$

The inequality (1.7) follows directly from the last two inequalities.

Now by the Cauchy–Schwarz inequality

$$\begin{aligned}&\mathbb{E}\left(\sum_{k=0}^{m-1} (X_{t_{k+1}} - X_{t_k})^4\right) \\ \leq & \mathbb{E}\left(\sup\left\{|X_s - X_{s'}|^2 : 0 \leq s, s' \leq t, |s-s'| \leq \text{mesh}(\pi))^2\right\} \sum_{k=0}^{m-1} (X_{t_{k+1}} - X_{t_k})^2\right) \\ \leq & \left[\mathbb{E}\left(\sup\left\{|X_s - X_{s'}|^2 : 0 \leq s, s' \leq t, |s-s'| \leq \text{mesh}(\pi))^2\right\}^2\right) \mathbb{E}\left(\left(\sum_{k=0}^{m-1} (X_{t_{k+1}} - X_{t_k})^2\right)^2\right)\right]^{\frac{1}{2}}\end{aligned}$$

And the right-hand side goes to zero by continuity of X_t and by the estimate (1.7).

We have now show that when X_t and $\langle X\rangle_t$ are bounded processes then the quadratic variation exists and $V_X^{(2)}(s) = \langle X\rangle_s$. To complete the proof for a general $f \in \mathcal{L}^2$, let

$$\tau_n = \inf\{t \geq 0 \,:\, |X_t| \geq n \text{ or } \langle X\rangle_t \geq n\}$$

and use the above argument for $f_n = f\mathbb{1}_{[0,\tau_n]}$ and $X_t^{(n)} = \int_0^t f_n(s,\cdot)\mathrm{d}B_s$. Notice that $X_t^{(n)} = X_t^{\tau_n}$ and that $\tau_n \nearrow \infty$ almost surely. $\square$

1.5 Itô's formula

Also in this section $(B_t)_{t\in\mathbb{R}_+}$ and $(B_t^{(1)},\ldots,B_t^{(m)})_{t\in\mathbb{R}_+}$ denote Brownian motions adapted to a filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.

1.5.1 Itô's formula for a Brownian motion

Theorem 1.5.1 (Itô's formula for a Brownian motion). *Let $F : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be a continuous function such that $\dot F, F', F''$ exist and are continuous, where*

$$\dot F(t,x) = \frac{\partial F}{\partial t}(t,x), \quad F'(t,x) = \frac{\partial F}{\partial x}(t,x) \quad \text{and } F''(t,x) = \frac{\partial^2 F}{\partial x^2}(t,x).$$

Then almost surely

$$F(t,B_t) = F(0,B_0) + \int_0^t \dot F(s,B_s)\mathrm{d}s + \int_0^t F'(s,B_s)\mathrm{d}B_s + \frac{1}{2}\int_0^t F''(s,B_s)\mathrm{d}s \tag{1.8}$$

for any $t \in \mathbb{R}_+$. For the previous equation we will use the shorthand notation

$$\mathrm{d}F(t,B_t) = \dot F(t,B_t)\mathrm{d}t + F'(t,B_t)\mathrm{d}B_t + \frac{1}{2}F''(t,B_t)\mathrm{d}t.$$

Proof. We'll prove this claim in the case when $\dot F, F', F''$ are compactly supported. The general case follows from this when we set $F_n = F\,h_n$ where $0 \leq h_n \leq 1$ is a sequence of smooth functions such that $h_n = 1$ in $[0,n]\times[-n,n]$ and 0 in the complement of $[0,n+1]\times[-n-1,n+1]$.

Take a partition π of $[0,t]$ and write a telescoping sum

$$F(t,B_t) - F(0,B_0) = \sum_{k=0}^{m(\pi)-1} (F(t_{k+1},B_{t_{k+1}}) - F(t_k,B_{t_k})).$$

By the mean value theorem

$$\begin{aligned} &F(t_{k+1},B_{t_{k+1}}) - F(t_k,B_{t_k}) = [F(t_{k+1},B_{t_{k+1}}) - F(t_k,B_{t_{k+1}})] + [F(t_k,B_{t_{k+1}}) - F(t_k,B_{t_k})] \\ &= \underbrace{[F(t_{k+1},B_{t_{k+1}}) - F(t_k,B_{t_{k+1}})]}_{=a_k} + \underbrace{F'(t_k,B_{t_k})(B_{t_{k+1}} - B_{t_k})}_{=b_k} + \frac{1}{2}\underbrace{F''(t_k,\eta_k)(B_{t_{k+1}} - B_{t_k})^2}_{=c_k} \end{aligned}$$

where η_k is a $\mathcal{F}_{t_{k+1}}$-measurable random variable that lies between B_{t_k} and $B_{t_{k+1}}$. Take a sequence of partitions π_n such that $\mathrm{mesh}(\pi_n) \to 0$ as $n \to \infty$. The claim is that the sums $\sum a_k$, $\sum b_k$ and $\sum c_k$ will converge to each of the three integrals in (1.8), respectively. The convergence will be almost sure along suitable subsequences of π_n.

Define for any $g : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ the following quantities measuring sizes of oscillations

$$\begin{aligned} O_{(B)}(\delta) &= \sup\{|B_s - B_{s'}| \,:\, 0 \leq s,s' \leq t \text{ s.t. } |s-s'| \leq \delta\} \\ O_g(\delta,\delta') &= \sup\{|g(s,x) - g(s',x')| \,:\, 0 \leq s,s' \leq t \text{ s.t. } |s-s'| \leq \delta \text{ and } x,x' \in \mathbb{R} \text{ s.t. } |x-x'| \leq \delta'\} \\ O_{g,B}(\delta) &= O_g\left(\delta, O_{(B)}(\delta)\right). \end{aligned}$$

Note first that by the mean value theorem

$$F(t_{k+1}, B_{t_{k+1}}) - F(t_k, B_{t_{k+1}}) = \dot{F}(\rho_k, B_{t_{k+1}})(t_{k+1} - t_k)$$

where $\rho_k \in (t_k, t_{k+1})$ is a random variable. Now

$$\left| \dot{F}(\rho_k, B_{t_{k+1}}) - \dot{F}(t_k, B_{t_k}) \right| \le O_{\dot{F},B}(\mathrm{mesh}(\pi_n))$$

and therefore

$$\left| \sum_k \dot{F}(\rho_k, B_{t_{k+1}})(t_{k+1} - t_k) - \sum_k \dot{F}(t_k, B_{t_k})(t_{k+1} - t_k) \right| \le t\, O_{\dot{F},B}(\mathrm{mesh}(\pi_n))$$

which goes to zero almost surely as $\mathrm{mesh}(\pi_n) \to 0$. By almost sure continuity of $t \mapsto \dot{F}(t, B_t)$,

$$\sum_k \dot{F}(t_k, B_{t_k})(t_{k+1} - t_k) \to \int_0^t \dot{F}(s, B_s)\mathrm{d}s$$

almost surely as $\mathrm{mesh}(\pi_n) \to 0$ and hence

$$\sum_k \dot{F}(\rho_k, B_{t_{k+1}})(t_{k+1} - t_k) \to \int_0^t \dot{F}(s, B_s)\mathrm{d}s$$

almost surely as $\mathrm{mesh}(\pi_n) \to 0$ and we have shown the almost sure convergence of $\sum a_k$.

We know from the definition of Itô integral that

$$\sum F'(t_k, B_{t_k})(B_{t_{k+1}} - B_{t_k}) \to \int_0^t F'(s, B_s)\mathrm{d}B_s \tag{1.9}$$

in L^2. Choose a subsequence of π_n (denoted for simplicity still by π_n) such that this convergence is almost sure. This gives the claim for $\sum b_k$.

Finally,

$$\left| \sum (F''(t_k, \eta_k) - F''(t_k, B_{t_k})) \cdot (B_{t_{k+1}} - B_{t_k})^2 \right| \le O_{F'',B}(\mathrm{mesh}(\pi_n)) \sum (B_{t_{k+1}} - B_{t_k})^2$$

Take a subsequence such that $\sum (B_{t_{k+1}} - B_{t_k})^2$ goes to t almost surely as $\mathrm{mesh}(\pi_n) \to 0$. Then the right-hand side goes to zero almost surely. Now the same calculation as for the quadratic variation of Brownian motion shows that

$$\begin{aligned}
&\mathbb{E}\left(\left(\sum F''(t_k, B_{t_k}) \cdot ((B_{t_{k+1}} - B_{t_k})^2 - (t_{k+1} - t_k)) \right)^2 \right) \\
&= \sum \mathbb{E}\left(F''(t_k, B_{t_k})^2 \cdot ((B_{t_{k+1}} - B_{t_k})^2 - (t_{k+1} - t_k))^2 \right) \\
&\le \|F''\|_\infty^2 \sum \mathbb{E}\left(((B_{t_{k+1}} - B_{t_k})^2 - (t_{k+1} - t_k))^2 \right) \\
&= 2\|F''\|_\infty^2 \sum (t_{k+1} - t_k)^2
\end{aligned}$$

which goes to zero. Now take yet another subsequence such that

$$\sum F''(t_k, B_{t_k}) \cdot ((B_{t_{k+1}} - B_{t_k})^2 - (t_{k+1} - t_k)) \to 0$$

almost surely. Then on the event that $t \mapsto F''(t, B_t)$ is continuous,

$$\sum F''(t_k, B_{t_k})((B_{t_{k+1}} - B_{t_k})^2 \to \int_0^t F''(s, B_s)\mathrm{d}s$$

almost surely. Hence along the chosen subsequence

$$\sum F''(t_k, \eta_k)(B_{t_{k+1}} - B_{t_k})^2 \to \int_0^t F''(s, B_s)\mathrm{d}s \tag{1.10}$$

almost surely giving the claim for $\sum c_k$.

Now we have shown that for fixed t, Itô's formula (1.8) holds almost surely. Therefore it holds almost surely for all rational t. Finally, by continuity of both sides in t, it holds almost surely for all t. □

1.5.2 Itô's formula for semimartingales

From now on, we'll write the time parameter of the integrands explicitly. Let

$$X_t = \int_0^t f_s \mathrm{d}B_s, \quad Y_t = \int_0^t g_s \mathrm{d}B_s$$

where $f, g \in \mathcal{L}^2_{\text{loc}}$. Then their *(quadratic) covariation process* is defined as

$$\langle X, Y\rangle_t = \int_0^t f_s g_s \mathrm{d}s.$$

Note that it satisfies

$$4\langle X, Y\rangle_t = \langle X+Y\rangle_t - \langle X-Y\rangle_t.$$

Hence $X_tY_t - \langle X, Y\rangle_t$ is a local martingale and along partitions of $[0,t]$

$$\lim_{\text{mesh}(\pi)\to 0} \sum_{k=0}^{m(\pi)-1} (X_{t_{k+1}} - X_{t_k})(Y_{t_{k+1}} - Y_{t_k}) = \langle X, Y\rangle_t$$

in probability. This claim and similar claims below are not verified here in detail.

If $(B^{(1)}, B^{(2)})$ is a standard two-dimensional Brownian motion and

$$X_t = \int_0^t f_s \,\mathrm{d}B^{(1)}_s, \quad Y_t = \int_0^t g_s \,\mathrm{d}B^{(2)}_s$$

where $f, g \in \mathcal{L}^2_{\text{loc}}$, then X_tY_t is a local martingale and the covariation process is naturally

$$\langle X, Y\rangle_t = 0.$$

This can be seen by showing first for bounded, simple processes (f_t) and (g_s) that X_tY_t is a martingale and then extend to other cases.

In the most general case, let $(B^{(1)}_t, B^{(2)}_t, \ldots, B^{(m)}_t)$ be a standard m-dimensional Brownian motion. Let

$$X_t = X_0 + \int_0^t f_s \mathrm{d}s + \sum_{k=1}^m \int_0^t g^{(k)}_s \,\mathrm{d}B^{(k)}_s \tag{1.11}$$

$$Y_t = Y_0 + \int_0^t \hat{f}_s \mathrm{d}s + \sum_{k=1}^m \int_0^t \hat{g}^{(k)}_s \,\mathrm{d}B^{(k)}_s$$

where X_0 and Y_0 are $\mathcal{F}_0$-measurable random variables, $g^{(k)}, \hat{g}^{(k)} \in \mathcal{L}^2_{\text{loc}}$ and $f, \hat{f}$ are measurable, adapted to $(F_t)_{t\in\mathbb{R}_+}$ and satisfy

$$\mathbb{P}\left(\int_0^t |f_s| \mathrm{d}s < \infty \text{ for all } t \in \mathbb{R}_+\right) = 1.$$

Then since integrals $\int_0^t f_s \mathrm{d}s$ have (locally) finite total variation, by the above it is natural to define

$$\langle X\rangle_t = \sum_{k=1}^m \int_0^t \left(g^{(k)}_s\right)^2 \mathrm{d}s, \qquad \langle Y\rangle_t = \sum_{k=1}^m \int_0^t \left(\hat{g}^{(k)}_s\right)^2 \mathrm{d}s, \qquad \langle X, Y\rangle_t = \sum_{k=1}^m \int_0^t g^{(k)}_s \hat{g}^{(k)}_s \,\mathrm{d}s$$

which are the quadratic variation and covariation processes also in the sense of Definition 1.3.8.

Definition 1.5.2. We call a process of the form (1.11) a *semimartingale* and use a shorthand notation

$$\mathrm{d}X_t = f_t \mathrm{d}t + \sum_{k=1}^m g^{(k)}_t \,\mathrm{d}B^{(k)}_t.$$

Remark. This is a slight abuse of standard termilogy. More generally semimartingale is a process that is sum of an adapted finite variation process and a local martingale.

Theorem 1.5.3 (Itô's formula for semimartingales)**.** *Let $1 \le l \le n$. Let $X_t^{(j)}$ be semimartingales*

$$\mathrm{d}X_t^{(j)} = f_t^{(j)}\,\mathrm{d}t + \sum_{k=1}^{m} g_t^{(j,k)}\,\mathrm{d}B_t^{(k)}$$

for $1 \le j \le n$ where $f^{(j)}$ and $g^{(j,k)}$ are as above. Assume that $g^{(j,k)} = 0$ identically for $j > l$. Let $F : \mathbb{R}^n \to \mathbb{R}$ be continuous function such that $\partial_{x_i} F$ exists and is continuous for all $1 \le i \le n$ and that $\partial_{x_i x_j} F$ exists and is continuous for all $1 \le i, j \le l$.

Then $Y_t = F(X_t^{(1)}, \ldots, X_t^{(n)})$ is a semimartingale and almost surely

$$\begin{aligned}\mathrm{d}Y_t =& \sum_{j=1}^{n} \left\{ \partial_{x_j} F(X_t^{(1)}, \ldots, X_t^{(n)}) f_t^{(j)}\,\mathrm{d}t + \sum_{k=1}^{m} \partial_{x_j} F(X_t^{(1)}, \ldots, X_t^{(n)}) g_t^{(j,k)}\,\mathrm{d}B_t^{(k)} \right\} \\ &+ \frac{1}{2} \sum_{i,j=1}^{l} \sum_{k=1}^{m} \partial_{x_i,x_j} F(X_t^{(1)}, \ldots, X_t^{(n)})\, g_t^{(i,k)} g_t^{(j,k)}\,\mathrm{d}t\end{aligned}$$

for all $t \in \mathbb{R}_+$. This is written shortly as

$$\mathrm{d}Y_t = \sum_{j=1}^{n} (\partial_j F)\ \mathrm{d}X_t^{(j)} + \frac{1}{2} \sum_{i,j=1}^{l} (\partial_{ij} F)\ \mathrm{d}\left\langle X^{(i)}, X^{(j)} \right\rangle_t .$$

Remark. Note that the theorem includes the case when F depends explicitly on time: let $l < n$ and take $X_t^{(n)} = t$. Theorem 1.5.1 is a special case of Theorem 1.5.3.

1.5.3 Rules for stochastic calculus

Let $Y_t = F(X_t^{(1)}, \ldots, X_t^{(n)})$. Then the reader can memorize Itô's formula for Y_t by writing formally $Z_{t+\mathrm{d}t} = Z_t + \mathrm{d}Z_t$ for any semimartingale Z_t and then take the Taylor expansion of F at $(X_t^{(1)}, \ldots, X_t^{(n)})$ and then use the rules

$$\mathrm{d}t^2 = 0, \qquad \mathrm{d}t\,\mathrm{d}B_t^{(i)} = 0, \qquad \mathrm{d}B_t^{(i)}\,\mathrm{d}B_t^{(j)} = \delta_{ij}\mathrm{d}t.$$

1.5.4 Examples

Example 1.5.4. Let $F(x) = x^2/2$ and let $(B_t)_{t\in\mathbb{R}_+}$ be a one-dimensional Brownian motion with $B_0 = 0$, then by Theorem 1.5.1

$$\frac{1}{2}B_t^2 = \int_0^t B_s \mathrm{d}B_s + \frac{1}{2}\int_0^t \mathrm{d}s$$

and hence after rearranging the terms

$$\int_0^t B_s \mathrm{d}B_s = \frac{1}{2}B_t^2 - \frac{1}{2}t$$

which is in agreement with the result we obtained by directly applying the definition of Itô integral.

Example 1.5.5. Let $(B_t^{(1)}, \ldots, B_t^{(m)})$ be m-dimensional standard Brownian motion, $m \ge 2$, started from $(B_0^{(1)}, \ldots, B_0^{(m)}) \ne 0$ and let $F(x_1, \ldots, x_m) = \left(\sum_{k=1}^m x_k^2\right)^{1/2}$. Then by Itô's formula $Y_t = F(B_t^{(1)}, \ldots, B_t^{(n)})$ satisfies

$$\mathrm{d}Y_t = \sum_k \frac{B_t^{(k)}\mathrm{d}B_t^{(k)}}{Y_t} + \frac{m-1}{2Y_t}\mathrm{d}t$$

as shown in the exercises.

1.6 Further topics in stochastic analysis

1.6.1 Usual conditions

Let's comment on some assumptions usually assumed in textbooks on stochastic analysis. If we are given a probability space $(\Omega, \mathcal{F}', \mathbb{P})$ and a filtration $(\mathcal{F}'_t)_{t\in\mathbb{R}}$, then we can complete $\mathcal{F}'$ by including all null sets and use the *usual augmentation* of $(\mathcal{F}'_t)_{t\in\mathbb{R}}$ which is defined by including all the null sets in the filtration and making the filtration right continuous:

$$\begin{aligned}
\mathcal{N} &= \{A \subset \Omega \,:\, A \subset E \text{ for some } E \in \mathcal{F} \text{ s.t. } \mathbb{P}(E) = 0\} \\
\mathcal{F} &= \sigma(\mathcal{F}' \cup \mathcal{N}) \\
\overline{\mathcal{F}}_t &= \sigma(\mathcal{F}'_t \cup \mathcal{N}) \\
\mathcal{F}_t &= \bigcap_{s>t} \overline{\mathcal{F}}_s.
\end{aligned}$$

The filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ constructed in this way is *right-continuous* in the sense that $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$.

We will now assume that $\mathcal{F}$ is complete and that $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ satisfies the *usual conditions*, i.e., it is complete and right-continuous. The right-continuity of the filtration affects the set of stopping times. Here is an example result.

Lemma 1.6.1. *If $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is right-continuous and $(X_t)_{t\in\mathbb{R}_+}$ is a continuous, adapted $\mathbb{R}^d$-valued process, then the* hitting-time *of a open or closed set $H \subset \mathbb{R}^d$*

$$\tau_H = \inf\{t \in \mathbb{R}_+ \,:\, X_t \in H\}$$

is a stopping time.

1.6.2 Optional stopping

In this section we present second martingale tool which we need for our theory.

Definition 1.6.2. If τ is a stopping time with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$, define the *stopping time σ-algebra* as

$$\mathcal{F}_\tau = \{A \in \mathcal{F} \,:\, A \cap \{\tau \le t\} \in \mathcal{F}_t \text{ for all } t \in \mathbb{R}_+\}$$

In the same way, as $\mathcal{F}_t$ can be thought as the information available at time t, a stopping time σ-algebra $\mathcal{F}_\tau$ can be thought as the information available at a random time τ. The following set of results extends the martingale property to random times.

Theorem 1.6.3. *Let $(M_t)_{t\in\mathbb{R}_+}$ be a continuous martingale and τ and σ stopping times with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. Then for each $t \in \mathbb{R}_+$*

$$\mathbb{E}(M_{t\wedge\tau}|\mathcal{F}_\sigma) = M_{t\wedge\sigma\wedge\tau}$$

Remark. As seen below, we have to care about the integrability of quantities such as M_τ. Here the non-random number t in $M_{t\wedge\tau}$ guarantees that $\mathbb{E}|M_{t\wedge\tau}| < \infty$.

Corollary 1.6.4. *Let $(M_t)_{t\in\mathbb{R}_+}$ be a continuous martingale and τ be a stopping time with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. Then the process $(M^\tau_t)_{t\in\mathbb{R}_+}$ defined by*

$$M^\tau_t = M_{t\wedge\tau}$$

is a continuous martingale with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.

Remark. Stopped local martingales are local martingales by the same argument.

Definition 1.6.5. A collection $\mathcal{C}$ of random variables is said to be *uniformly integrable* if

$$\lim_{m\to\infty} \sup_{X\in\mathcal{C}} \mathbb{E}(|X|\,;\,|X| \ge m) = 0.$$

Corollary 1.6.6. *Let $(M_t)_{t\in\mathbb{R}_+}$ be a continuous martingale and τ and σ almost surely finite stopping times with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. Assume that $\sigma \le \tau$. Then*

$$\mathbb{E}(M_\tau|\mathcal{F}_\sigma) = M_\sigma$$

under any of the following conditions:

- *For some constant $C > 0$, $\sigma \le \tau \le C$*
- *For some constant $C > 0$, $|M_t| \le C$ for all t.*
- *The collection of random variables M_t, $t \in \mathbb{R}_+$, is uniformly integrable.*

Remark. In some sense, the first two cases are special cases of the last case.

Remark. In the last case, $M_\sigma = \mathbb{E}(M_\tau|\mathcal{F}_\sigma) = \mathbb{E}(M_\infty|\mathcal{F}_\sigma)$ for some random variable M_∞ and $M_t \to M_\infty$ in L^1.

Example 1.6.7. The martingale strategy presented in Example 1.4.14 is not uniformly integrable and $\lim_{n\to\infty} M_n$ doesn't exist in L^1 although $\lim_{n\to\infty} M_n = 1$ almost surely.

The reader can find more about optional stopping from Section 3.1 of Seppäläinen's book and Section II.3 of Revuz&Yor: "Continuous martingales and Brownian motion."

1.6.3 Time-change of local martingales

The following theorem is an application of Itô's formula. It is a special case of more general result that *any continuous local martingale is a time-change of a Brownian motion.* The proof of the general result would follow the same lines if we had established the theory of the stochastic integral with respect to local martingales and we had corresponding Itô's formula available.

Theorem 1.6.8. *Let $(X_t)_{t\in\mathbb{R}_+}$ be a local martingale defined by*

$$X_t = \sum_{k=1}^m \int_0^t g_s^{(k)}\,\mathrm{d}B_s^{(k)}$$

where $g_t^{(k)} \in \mathcal{L}^2_{\mathrm{loc}}$. Let $(\sigma_r)_{r\in\mathbb{R}_+}$ be the set of stopping times

$$\sigma_r = \inf\{t \ge 0 \,:\, \langle X\rangle_t \ge r\}$$

where

$$\langle X\rangle_t = \sum_{k=1}^m \int_0^t \left(g_s^{(k)}\right)^2\,\mathrm{d}s$$

is the quadratic variation process as before. Assume that almost surely $\langle X\rangle_t \to \infty$ as $t \to \infty$. Then the process $(Y_t)_{t\in\mathbb{R}_+}$ defined by

$$Y_t = X_{\sigma_t}$$

is a standard one-dimensional Brownian motion with respect to the filtration $(\mathcal{F}_{\sigma_t})_{t\in\mathbb{R}_+}$.

Proof. Since $\langle X\rangle_t \to \infty$ as $t \to \infty$, each σ_r is almost surely finite. By the continuity of the mapping $t \mapsto \langle X\rangle_t$, we have that $\langle X\rangle_{\sigma_r} = r$.

Let

$$M_t = \exp\left(i\theta X_t + \frac{\theta^2}{2}\langle X\rangle_t\right).$$

By Itô's formula $(M_t)_{t\in\mathbb{R}_+}$ is a continuous local martingale. Note that $(M_t)_{t\in\mathbb{R}_+}$ is a complex valued process, but this causes no problems: we can apply Itô's formula separately for its real and imaginary parts. The statement that it is a local martingale means that both its real and imaginary parts are local martingales. Since $M_t^{\sigma_r} = M_{t\wedge\sigma_r}$ is bounded, $(M_t^{\sigma_r})_{t\in\mathbb{R}_+}$ is a martingale. Namely, if τ_n is the localizing sequence of $(M_t)_{t\in\mathbb{R}_+}$, then $(M_t^{\sigma_r\wedge\tau_n})_{t\in\mathbb{R}_+}$ is a martingale. Hence by boundedness of $(M_t^{\sigma_r})_{t\in\mathbb{R}_+}$ and by the fact that $\tau_n \nearrow \infty$ almost surely as $n\to\infty$,

$$\mathbb{E}(\underbrace{M_t^{\sigma_r\wedge\tau_n}}_{\to M_t^{\sigma_r} \text{ in } L^1}|\mathcal{F}_s) = \underbrace{M_s^{\sigma_r\wedge\tau_n}}_{\to M_s^{\sigma_r} \text{ in } L^1, \text{ as } n\to\infty}$$

and therefore

$$\mathbb{E}(M_t^{\sigma_r}|\mathcal{F}_s) = M_s^{\sigma_r}$$

Hence $(M_t^{\sigma_r})_{t\in\mathbb{R}_+}$ is a continuous bounded martingale.

Now we can apply the optional stopping theorem for stopping times $\sigma_s \le \sigma_r$, $0\le s\le r$, to show that

$$\mathbb{E}(M_{\sigma_r}|\mathcal{F}_{\sigma_s}) = M_{\sigma_s}.$$

This implies that for any $0\le s\le r$ and for any $\theta\in\mathbb{R}$,

$$\mathbb{E}\left(\exp\left(i\theta(X_{\sigma_r}-X_{\sigma_s})\right)|\mathcal{F}_{\sigma_s}\right) = \exp\left(-\frac{\theta^2}{2}(r-s)\right).$$

The right-hand side of this equation is the characteristic function of a normal random variable with mean 0 and variance $r-s$. The left-hand side is a conditional version of characteristic function of $X_{\sigma_r}-X_{\sigma_s}$. That characteristic function is now constant as a $\mathcal{F}_{\sigma_s}$-measurable random variable. Therefore the fact that the characteristic function determines the distribution uniquely shows that $X_{\sigma_r}-X_{\sigma_s}$ is independent from $\mathcal{F}_{\sigma_s}$ and also that $X_{\sigma_r}-X_{\sigma_s}$ is normally distributed with mean 0 and variance $r-s$. □

1.6.4 Strong Markov property

For the sake of completeness, let's state the following property of Brownian motion which extends the Markov property of Brownian motion (the property that for each $s\in\mathbb{R}_+$, the process $Y_t = B_{t+s}-B_s$ is a standard Brownian motion independent from $\mathcal{F}_s$).

Theorem 1.6.9 (Strong Markov property)**.** *For any stopping time τ which is almost surely finite, the process $(Y_t)_{t\in\mathbb{R}_+}$ defined by*

$$Y_t = B_{\tau+t} - B_\tau$$

is a standard Brownian motion independent of $\mathcal{F}_\tau$.

Remark. Note that in the independence property, an "infinitesimal peek to the future" is allowed because the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is right-continuous and hence $\mathcal{F}_\tau = \bigcap_{h>0}\mathcal{F}_{\tau+h}$.

1.6.5 Stochastic differential equations

Let X_t be an $\mathbb{R}^n$ valued continuous stochastic process and let B_t be a standard m-dimensional Brownian motion. We say that X_t satisfies the *stochastic differential equation* (SDE)

$$\mathrm{d}X_t = F(t,X_t)\mathrm{d}t + G(t,X_t)\mathrm{d}B_t$$

with initial condition $X_0 = Z$ if for each t

$$X_t = Z + \int_0^t F(s,X_s)\,\mathrm{d}s + \int_0^t G(s,X_s)\,\mathrm{d}B_s$$

Here $G(s,X_s)\,\mathrm{d}B_s$ is understood as a matrix product so that the i'th component, $1\le i\le n$, is $\sum_{j=1}^m G^{(i,j)}(s,X_s)\,\mathrm{d}B_s^{(j)}$.

Theorem 1.6.10. *Let B_t be m-dimensional Brownian motion and let*

$$F : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$$
$$G : [0,T]\times\mathbb{R}^n \to \mathbb{R}^{n\times m}$$

be measurable maps. Let Z be $\mathbb{R}^n$-valued square integrable random variable which is independent from $\sigma(B_t, t \in \mathbb{R}+)$. Suppose that

$$|F(t,x)| + |G(t,x)| \le C(1+|x|)$$
$$|F(t,x) - F(t,y)| + |G(t,x) - G(t,y)| \le D\,|x-y|$$

where for any matrix A, $|A| = \sqrt{\sum_{i,j} A_{i,j}}$.

Then there exist a unique continuous solution $(X_t)_{t\in[0,T]}$ to the stochastic differential equation

$$X_0 = Z, \qquad \mathrm{d}X_t = F(t,X_t)\mathrm{d}t + G(t,X_t)\mathrm{d}B_t, \qquad t \in [0,T],$$

with the property that X_t is adapted to the filtration $\mathcal{F}_t^{(B,Z)}$ generated by Z and B_s, $s \in [0,t]$. Furthermore

$$\mathbb{E}\left[\int_0^T |X_t|^2 \mathrm{d}t\right] < \infty.$$

Remark. In the time-homogeneous case, $F(t,x) = F(x)$ and $G(t,x) = G(x)$, these solutions X_t are called *diffusions*. Another viewpoint to diffusions is that it is a family of processes with one element for each starting point $x \in \mathbb{R}^n$. The uniqueness of the solution together with the strong Markov property of Brownian motion imply that diffusions have the following *strong Markov property*: $X_{\tau+t}$ conditioned on $\mathcal{F}_\tau$ is distributed as a independent copy of the diffusion $\tilde{X}_t$ send from X_τ.

1.7 Conformal invariance of two-dimensional Brownian motion

As usual complex number z is represented in terms of its real and imaginary parts as $z = x + iy$, similarly complex valued function of a complex variable is divided into its real and imaginary parts as $f(z) = u(z) + i\,v(z)$. Define as usual the following partial differential operators

$$\partial = \frac{1}{2}(\partial_x - i\partial_y), \qquad \overline{\partial} = \frac{1}{2}(\partial_x + i\partial_y).$$

Let U be a open set in the complex plane $\mathbb{C}$ and let $z_0 \in U$. The basic result of complex analysis is that the following statements about a function $f : U \to \mathbb{C}$ are equivalent:

- The function f is *holomorphic* near z_0: the complex derivative

  $$f'(z) = \lim_{h\to 0} \frac{f(z+h) - f(z)}{h}$$

  exists and is continuous in a neighborhood of z_0. This is equivalent to that the statement that f has continuous partial derivatives $\partial_x f$, $\partial_y f$ and satisfies $\overline{\partial} f(z) = 0$ in a neighborhood of z_0. Hence the complex derivative $f'(z) = \lim_{h\to 0} h^{-1}(f(z+h) - f(z))$ satisfies $f'(z) = \partial f(z) = \partial_x f(z) = -i\partial_y f(z)$.

- The real and imaginary parts of f satisfies *Cauchy–Riemann equations* near z_0:

  $$\partial_x u = \partial_y v, \qquad \partial_x v = -\partial_y u$$

- The function f is *(complex) analytic* at z_0: $f(z) = \sum_{n=0}^{\infty} c_n (z-z_0)^n$ which converges absolutely when $|z - z_0| \le r$ for some $r > 0$.

Remember that u and v are *harmonic*: $\Delta u = 0 = \Delta v$ (say, by the Cauchy–Riemann equations $u_{xx} + u_{yy} = v_{xy} - v_{xy} = 0$).

Define a complex Brownian motion send from $z_0 \in \mathbb{C}$ as

$$B_t = B_t^{\mathbb{C}} = z_0 + B_t^{(1)} + i\,B_t^{(2)}$$

The complex Brownian motion is conformally invariant (up to a time-change) as shown next.

Theorem 1.7.1. *Let $U \subset \mathbb{C}$ be a domain (non-empty connected open set) and let $f : U \to \mathbb{C}$ be analytic. Let $z_0 \in U$ and let B_t be a complex Brownian motion send from $z_0 \in \mathbb{C}$. Let $\tau = \inf\{t \geq 0 : B_t \notin U\}$. Let $Z_t = f(B_{\sigma(t)})$ for $0 \leq t < S(\tau)$ where $\sigma(t) = S^{-1}(t)$ and*

$$S(t) = \int_0^t |f'(B_s)|^2 \mathrm{d}s$$

for $0 \leq 0 < \tau$. Then Z_t is a complex Brownian motion send from $f(z_0)$ and stopped at $S(\tau)$.

Proof. As above write $f = u + i\,v$. Define

$$X_t = u(B_t), \qquad Y_t = v(B_t)$$

Since u and v are harmonic and satisfy the Cauchy–Riemann equations,

$$\begin{aligned} \mathrm{d}X_t &= u_1(B_t)\mathrm{d}B_t^{(1)} + u_2(B_t)\mathrm{d}B_t^{(2)} \\ \mathrm{d}Y_t &= -u_2(B_t)\mathrm{d}B_t^{(1)} + u_1(B_t)\mathrm{d}B_t^{(2)} \end{aligned}$$

by Itô's formula, where $u_1 = \partial_x u$ and $u_2 = \partial_y u$ are the partial derivatives of u. Therefore $(X_t)_{t\in\mathbb{R}_+}$ and $(Y_t)_{t\in\mathbb{R}_+}$ are local martingales. Now

$$\langle X\rangle_t = \langle Y\rangle_t = \int_0^t u_1(B_s)^2 + u_2(B_s)^2\,\mathrm{d}s = \int_0^t |f'(B_s)|^2 \mathrm{d}s$$

and $\langle X, Y\rangle_t = 0$. Here we used the fact that $f'(z) = u_1(z) - i\,u_2(z)$. A slight modification of the proof of Theorem 1.6.8 shows that for any $\theta_1, \theta_2 \in \mathbb{R}$

$$\exp\left(i\theta_1 X_t + \frac{\theta_1^2}{2}\langle X\rangle_t\right) \exp\left(i\theta_2 Y_t + \frac{\theta_2^2}{2}\langle Y\rangle_t\right)$$

is a local martingale and that $(X_{\sigma_t})_{t\in\mathbb{R}_+}$ and $(Y_{\sigma_t})_{t\in\mathbb{R}_+}$ are independent Brownian motions. □

Remark. In the previous proof, it was crucial that $(X_t)_{t\in\mathbb{R}_+}$ and $(Y_t)_{t\in\mathbb{R}_+}$ had the same quadratic variation. There is no general time-change result for multidimensional continuous local martingales of the form of Theorem 1.6.8.