Stochastic Differential Equations
Summary

Romke Rozendaal
4094255

June 5, 2015

1 Measures, Integrals, and Foundations of Probability Theory

1.1 Measure theory and Integration

Definition 1. A family $\mathcal{F}$ of subsets of Ω is called a $\sigma-$**algebra** if:

1. $\Omega \in \mathcal{F}$ and $\emptyset \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
3. $A_1, A_2, \dots \in \mathcal{F} \to \cup_{n=1}^{\infty} A_n \in \mathcal{F}$

Example 1. Some examples of σ-algebra's:

- $\{\emptyset, \Omega\}$ is a trivial σ-algebra.
- The power set 2^{Ω}, which is the collection of all subsets of A is a σ-algebra.

Example 2. Given a family of sets A, there is a smallest σ-algebra which contains A. Notation: $\sigma(A)$, called the σ-algebra generated by A.

Example 3. The Borel σ-algebra of $\mathbb{R}^d$, (notation $\mathcal{B}(\mathbb{R}^d)$) is the σ-algebra generated by all open sets in $\mathbb{R}^d$.

Example 4. Let $f : \Omega \to \mathbb{R}$ be a function. Let $\{f \in B\} = \{\omega \in \Omega : f(\omega) \in B\}$. The collection $\mathcal{O}(f) := \{\{f \in B\} : B \in \mathcal{B}(\mathbb{R})\}\}$ is a σ-algebra in Ω. It is called the σ-algebra generated by f.

Let $(\Omega, \mathcal{F})$ be a measurable space. $f : \Omega \to \mathbb{R}$ is called **measurable/Borel measurable** if $\forall B \in \mathcal{B}$ it holds that $\{f \in B\} \in \mathcal{F}$.

- Sums, product, etc. of measurable functions are measurable.
- Limits, countable suprema and infima are measurable.

Definition 2. A mapping: $\mu : f \to [0, \infty]$ is called a **measure** if

1. $\mu(\emptyset) = 0$

2. $\forall$ disjoint $A_1, A_2, \dots \in \mathcal{F}$ then $\mu\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$

Caratheodory extension Theorem:

Definition 3. For a given set Ω, we may define a *ring* R as a subset of the powerset of Ω which has the following properties

- $\emptyset \in R$
- For all $A, B \in R$ we have $A \cup B \in R$
- For all $A, B \in R$ we have $A \setminus B \in R$

This theorem states that if there exists a measure μ on a ring R then there exists a measure μ^* on the sigma algebra of that ring such that μ^* is an extension of μ (That is, $\mu^*|_R = \mu$)

Dynkin uniqueness of measure

Definition 4. Let Ω be a nonempty set, and let D be a collection of subsets of Ω. Then D is a λ-system if

1. $\Omega \in D$
2. If $A, B \in D$ and $A \subseteq B$, then $B \setminus A \in D$.
3. If $A_1, A_2, A_3, \dots$ is a sequence of subsets in D and $A_n \subseteq A_{n+1}$ for all $n \geq 1$ then $\bigcup_{n=1}^{\infty} A_n \in D$

Equivalently, D is a π-system if

1. $\Omega \in D$
2. If $A \in D$ then $A^c \in D$.
3. If $A_1, A_2, A_3, \dots$ is a sequence of subsets in D and $A_i \cap A_j = \emptyset$ for all $i \neq j$ then $\bigcup_{n=1}^{\infty} A_n \in D$

An important fact is that a λ-system which is also a π-system (i.e. closed under finite intersection) is a σ-algebra.

Theorem 1 (Dynkin's $\pi - \lambda$ theorem)**.** *If P is a π-system and D is a λ-system with $P \subseteq D$ then $\sigma(P) \subseteq D$. In other words the σ-algebra generated by P is contained in D.*

Completion of measure There are certain technical benefits to having the following property in a measure space $(X, \mathcal{F}, \mu)$ called *completion*: if $N \in \mathcal{F}$ satisfies $\mu(N) = 0$, then every subset of N is measurable and then of course has measure zero.
It turns out that this can always be arranged by a simple enlargement of the σ-algebra. Let

$$\bar{\mathcal{F}} = \{F \in X : \text{ there exists } B, N \in \mathcal{F} \text{ and } F \subseteq N \text{ such that } \mu(N) = 0 \text{ and } A = B \cup F\}$$

1.2 Lebesgue measure

There exists a measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ which satisfies $\mu([a_1, b_1) \times \cdots \times [a_d, b_d) = \prod_{n=1}^{d}(b_n - a_n)$.
Integration: $f = \sum_i c_i \mathbf{1}_{A_i}$ then $\int f \mathrm{d}\mu = \sum_i c_i \mu(A_i)$.
The power of Lebesgue-integration lies in the fact that one can prove convergence theorems such as monotone convergence and dominated convergence.

Theorem 2 (Monotone convergence theorem)**.** *Let f_n be nonnegative measurable functions, ans assume $f_n \leq f_{n+1}$ almost everywhere, for each n. Let $f = \lim_{n\to\infty} f_n$. This limit exists at least almost everywhere. Then.*

$$\int f \, d\mu = \lim_{n\to\infty} f_n \, d\mu$$

Theorem 3 (Dominated convergence theorem)**.** *Let f_n be measurable functions, and assume the limit $f = \lim_{n\to\infty} f_n$ exists almost everywhere. Assume there exists a function $g \geq 0$ such that $|f_n| \leq g$ almost everywhere for each n and $\int g \, d\mu < \infty$. Then*

$$\int f \, d\mu = \lim_{n\to\infty} f_n \, d\mu$$

L^p-spaces: For a Borel-measurable function $f : \Omega \to \mathbb{R}$ let $||f||_{L^p} = \left(\int |f|^p \mathrm{d}\mu\right)^{\frac{1}{p}}$.
Let $\mathcal{L}^p(\Omega, \mathcal{F}, \mu) = \{f : \Omega \to \mathbf{R} \text{ measurable } : ||f||_p < \infty\}$. Then $\mathcal{L}^p$ is a vector space. $||.||_{L^p}$ is not a norm because $||f||_{L^p} = 0 \not\Rightarrow f = 0$. Let $f \sim g$ if $f = g$ almost everywhere, which is an equivalence relation. Then $L^p = \mathcal{L}^p \backslash \sim$ becomes a normed space. Moreover L^p is a complete space.
Hölder's inequality: $||f \cdot g||_{L^1} \leq ||f||_{L^p} \cdot ||g||_{L^q}$ for $\frac{1}{p} + \frac{1}{q} = 1$

Theorem 4 (Fubini's theorem)**.** *Let $f \in L^1(\mu \otimes \nu)$. Then $f_x \in L^1(\nu)$ for μ-almost every x, $f_y \in L^1(\mu)$ for ν-almost every y, $g \in L^1(\mu)$ and $h \in L^1(\nu)$. Iterated integration as follows, is valid:*

$$\begin{aligned}
\int_{X\times Y} f \, d(\mu \otimes \nu) &= \int_X \left\{ \int_Y f(x,)\nu(dy) \right\} \mu(dx) \\
&= \int_Y \left\{ \int_X f(x, y)\mu(dx) \right\} \nu(dy)
\end{aligned}$$

1.3 Probability spaces

We call $(\Omega, \mathcal{F}, P)$ a probability space if $P(\Omega) = 1$.

Definition 5. $X : \Omega \to \mathbb{R}$ is called a *random variable* if it is measurable.

Definition 6. σ-algebras $\mathcal{F}_1, \mathcal{F}_2, \ldots$ are *independent* if

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i) \qquad \forall A_i \in \mathcal{F}_i \qquad \forall i \leq n \qquad \forall n \in \mathbb{N}$$

Definition 7. $X_1, X_2, \ldots, : \Omega \to \mathbb{R}$ are *independent* if $\sigma(X_1), \sigma(X_2), \ldots,$ are independent.

Image measure: $X : \Omega \to \mathbb{R}^d$, $\mu_X(B) = P(X \in B), B \in \mathcal{B}(\mathbb{R}^d)$
Expectation: $\mathbb{E}[X] = \int_\Omega X \mathrm{d}P$

Theorem 5. *$X_1, \ldots, X_n : \Omega \to \mathbb{R}$ are independent $\iff$ the distribution of $(X_1, \ldots, X_n)$ is $\mu = \mu_{X_1} \times \cdots \times \mu_{X_n}$*

Theorem 6. *If X and Y are independent, then $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$ and $X \in L^p, Y \in L^{p'}$ then $\frac{1}{p} + \frac{1}{p'} = 1$*

Proof. $\mu_X(B) = P(X \in B), \mu_Y(B) = P(Y \in B)$ then

$$\begin{aligned}
\mathbb{E}[X] \cdot \mathbb{E}[Y] &= \int \int xy \mathrm{d}\mu_X(x) \mathrm{d}\mu_Y(y) \\
&\underbrace{=}_{\text{Fubini}} \int \int xy \mathrm{d}\mu_X \times \mu_Y(x, y) \\
&\underbrace{=}_{\text{independence}} \quad \mathbb{E}[XY]
\end{aligned}$$

□

Definition 8. *Almost surely (a.s.)* means with probability 1

Definition 9. Let $\{X_n\}$ be a sequence of random variables and X a random variable, all real valued.

1. $X_n \to X$ *almost surely* if
$$P\left\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\} = 1$$
2. $X_n \to X$ *in probability* if for every $\epsilon > 0$
$$\lim_{n\to\infty} P\{\omega : |X_n(\omega) - X(\omega)| \geq \epsilon\} = 0$$
3. $X_n \to X$ *in* L^p for $1 \leq p < \infty$ if
$$\lim_{n\to\infty} \mathbb{E}[|X_n(\omega) - X(\omega)|^p] = 0$$
4. $X_n \to X$ *in distribution* (also called *weakly*) if
$$\lim_{n\to\infty} P(X_n \leq x) = P(X \leq x)$$
for each x at which $F(x)$ is continuous.

Theorem 7 (Theorem 1.21)**.** *Let $\{X_n\}$ and X be real-valued random variables on a common probability space.*

1. *If $X_n \to X$ almost surely or in L^p for some $1 \leq p < \infty$, then $X_n \to X$ in probability.*
2. *If $X_n \to X$ in probability, then $X_n \to X$ weakly.*
3. *If $X_n \to X$ in probability, then there exists a subsequence X_{n_k} such that $X_{n_k} \to X$ almost surely.*
4. *Suppose $X_n \to X$ in probability. Then $X_n \to X$ in L^1 if and only if $\{X_n\}$ is uniformly integrable.*

1.4 Conditional Expectations

Example 5. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $x_1, \ldots, x_m, z, \ldots, z_n \in \mathbb{R}$ be distinct. Now let $X : \Omega \to \{x_1, \ldots, x_m\}, Z : \Omega \to \{z_1, \ldots, z_n\}$. Recall: $P(X = x_i | Z = z_j) \overset{\text{def}}{=} \frac{P(X=x_i, Z=z_j)}{P(Z=z_j)}$ and $\mathbb{E}[X|Z = z_j] = \sum_{i=1}^m x_i P(X = x_i | Z = z_j) = \frac{1}{P(Z=z_j)} \int_{\{Z=z_j\}} X \mathrm{d}P$.
A possible definition of $Y = \mathbb{E}[X|Z]$ could be $Y : \Omega \to \mathbb{R}$, $Y = \sum_{j=1}^n Y_j \mathbf{1}_{\{Z=z_j\}}$, where $Y_j = \mathbb{E}[X|Z = z_j]$.
How to extend this to general X? Let $A = \sigma(Z)$
Observation 1: Y is constant on sets $\{Z = z_j\}$ thus Y is $\mathcal{A}$-measurable.
Observation 2: $\int Y \mathrm{d}P = y_j \cdot P(Z = z_j) = \int_{\{Z=z_j\}} X \mathrm{d}P$. Thus $\forall G \in \mathcal{G}$: $\int_G Y \mathrm{d}P = \int_G X \mathrm{d}P$

Definition 10. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $X \in L^1(P)$ and let $\mathcal{A} \subseteq \mathcal{F}$ be a sub-σ-algebra.
We say that $Y : \Omega \to \mathbb{R}$ is *the conditional expectation of X given $\mathcal{A}$* if:

1. Y is $\mathcal{A}$-measurable.
2. $Y \in L^1(P)$ and $\forall A \in \mathcal{A}$ $\int_A Y \mathrm{d}P = \int_A x \mathrm{d}P$

Notation: $Y(\omega) = \mathbb{E}[X|\mathcal{A}](\omega)$ or $\mathbb{E}[X|\mathcal{A}]$

Note that $\mathbb{E}[\mathbb{E}[X|\mathcal{A}]] = \mathbb{E}[X]$

Theorem 8 (Uniqueness)**.** *If Y and $\tilde{Y}$ are both conditional expectations of X given $\mathcal{A}$ then $Y = \tilde{Y}$ a.s.*

Proof. Let $\Delta Y = Y - \tilde{Y}$. Then ΔY is $\mathcal{A}$-measurable and $\forall A \in \mathcal{A} : \int_A \Delta Y \mathrm{d}P = 0$
Let $A_1 = \{\Delta Y \geq 0\}$ and $A_2 = \{\Delta Y < 0\}$. Then $\mathbb{E}[|\Delta Y|] = \int_{A_1} \Delta Y \mathrm{d}P - \int_{A_2} \Delta Y \mathrm{d}P = 0 - 0 = 0$. Thus $|\Delta Y| = 0$ a.s., thus $Y = \tilde{Y}$ a.s. $\square$

Definition 11. In this case Y and $\tilde{Y}$ are called *versions* of $\mathbb{E}[X|\mathcal{A}]$

Theorem 9. *Properties of conditional expectation Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $X, Y \in L^1(P), \mathcal{A}, \mathcal{B} \subseteq \mathcal{F}$ be sub-σ-fields. Then:*

1. $\mathbb{E}[\mathbb{E}[X|\mathcal{A}]] = \mathbb{E}[X]$
2. *(Linearity)* $\mathbb{E}[\alpha X + \beta Y|\mathcal{A}] = \alpha \mathbb{E}[X|\mathcal{A}] + \beta \mathbb{E}[Y|\mathcal{A}]$, $\alpha, \beta \in \mathbb{R}$
3. *(Positivity) If $X \geq Y$ then* $\mathbb{E}[X|\mathcal{A}] \geq \mathbb{E}[Y|\mathcal{A}]$.
4. *If X is $\mathcal{A}$-measurable then* $\mathbb{E}[X|\mathcal{A}] = X$.
5. *(Taking out what is known). If X is $\mathcal{A}$-measurable and $XY \in L^1(P)$, then* $\mathbb{E}[XY|\mathcal{A}] = X\mathbb{E}[Y|\mathcal{A}]$
6. *(Independence) If X and $\mathcal{A}$ are independent, then* $\mathbb{E}[X|\mathcal{A}] = \mathbb{E}[X]$
7. *(Tower property) If $\mathcal{A} \subseteq \mathcal{B}$, then* $\mathbb{E}[\mathbb{E}[X|\mathcal{B}]|\mathcal{A}] = \mathbb{E}[X|\mathcal{A}]$ *and also* $\mathbb{E}[\mathbb{E}[X|\mathcal{A}]|\mathcal{B}] = \mathbb{E}[X|\mathcal{A}]$ *by 4.*
8. *If $\mathcal{A} \subseteq \mathcal{B}$ and $\mathbb{E}[X|\mathcal{B}]$ is $\mathcal{A}$-measurable, then* $\mathbb{E}[X|\mathcal{B}] = \mathbb{E}[X|\mathcal{A}]$.

9. *(Jensen's inequality) Let* $f : (a,b) \to \mathbb{R}$ *be convex,* $-\infty \leq a < b <\leq \infty$*. Assume that* $a < X < b$*. a.s. and* $f(X) \in L^1(P)$ *Then:* $f(\mathbb{E}[X|\mathcal{A}] \leq \mathbb{E}[f(X)|\mathcal{A}]$

Proof. Simple exercises: 1,2,4,6,8
Good exercises: 3,5,7
Too difficult: 9,10 $\square$

2 Stochastic Processes

Let $(\Omega, \mathcal{F}, P)$ be a probability space. From now on we will assume that $\mathcal{F}$ is complete, i.e. if $N \in \mathcal{A}$ satisfies $\mu(N) = 0$, then every subset of N is measurable (and then of course has measure zero).

Definition 12. A *filtration* on $(\Omega, \mathcal{F}, P)$ is a family of σ-fields $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}, \forall 0 \leq s < t < \infty$.

Definition 13. A *process* $X : \mathbb{R} \times \Omega \to \mathbb{R}$ if $\mathcal{B}_{\mu_F} \times \mathcal{F}$-measurable.
Notation: $(X_t)_{t\geq 0}$, $(t, \omega) \to X_t(\omega)$ or $X(t, \omega)$

Example 6. $(X_t)_{t\geq 0}$ a stock price. A possible filtration $\mathcal{F}_t^X = \sigma(X_s : s \in [0, t])$, our knowledge at time t.

Convention: $\mathcal{F}_t$ contains all null sets of $\mathcal{F}$ otherwise replace $\mathcal{F}_t$ by $\bar{\mathcal{F}}_t = \{\mathcal{B} \in \mathcal{F} : \exists \mathcal{A} \in \mathcal{F}_t \text{ s.t. } P(\mathcal{A}\Delta\mathcal{B}) = 0\}$ where $\mathcal{A}\Delta\mathcal{B}$ is the symmetric difference.

Definition 14. $(X_t)_{t\geq 0}$ is called *adapted* to $(\mathcal{F}_t)_{t\geq 0}$ if $\forall t \geq 0 : \omega \to X_t(\omega)$ is $\mathcal{F}_t$-measurable.

Definition 15. $(X_t)_{t\geq 0}$ is called *progressively measurable* if $\forall T \geq 0$ X restricted to $[0, T] \times \Omega$ is $\mathcal{B}_{[0,T]}$

Observation: X progressively measurable $\Rightarrow$ X is adapted.

Definition 16. $(X_t)_{t\geq 0}, (Y_t)_{t\geq 0}$ are called *modifications or versions* if $\forall t \geq 0, P(X_t = Y_t) = 1$.
$(X_t)_{t\geq 0}, (Y_t)_{t\geq 0}$ are called *indistinguishable* if $P(X_t = Y_t, \forall t \geq 0) = 1$.

Theorem 10. *Assume X is adapted to $(\mathcal{F}_t)_{t\geq 0}$ and X is left or right-continuous, then X is progressively measurable.*

Definition 17. X is called *cadlag* if it has right-continuous paths and $\forall \omega \in \Omega$: $\forall t > 0 : \lim_{s\uparrow t} X_s(\omega)$ exists.
caglad left-continuous and right limits exists.

Theorem 11. *Assume X, Y are right-continuous. Assume: $S \subseteq \mathbb{R}_+$ is dense and countable. If $\forall t \in S$: $P(X_t = Y_t) = 1$, then X and Y are indistinguishable. Similar for left-continuous if $0 \in S$.*

Proof. Let $\forall s \in S$: $V_s = \{X_s = Y_s\}$. Then $P(V_s) = 1$. Let $\Omega_0 = \bigcap_{s\in S} V_s$, then $P(\Omega_0) = 1$.
Claim: $\forall \omega \in \Omega_0, \forall t > 0$ $X_t = Y_t$ thus $P(X_t = Y_t, \forall t > 0) = P(\Omega_0) = 1$. $\square$

Definition 18. $\tau : \Omega \to [0, \infty]$ is called a *stopping time* if $\forall t \in (0, \infty) : \{\tau < t\} \in \mathcal{F}_t$

Example 7. First time a stock price is > 100.
First time a stock price is lower than the price a week before.

Theorem 12. *X adapted and continuous, $H \in \mathbb{R}$ is closed. Define: $\tau_H(\omega) = \inf\{\tau \geq 0 : X_t(\omega) \in H\}$, then τ_H is a stopping time.*

2.1 Quadratic variation

We start with *bounded variation* from section 1.1.9.
Given $F : [a, b] \to \mathbb{R}$, define: $V_F(t) := \sup\{\sum_{i=1}^{n} |F(S_i) - F(S_{i-1})| : a = S_0 < S_1 < \cdots < S_n = b\}$. F has *bounded variation* if $V_F(b) < \infty$.
Observation: $V_F(0) = 0$, V_f is non-decreasing.
Notation: $\mathrm{BV}[a, b]$ is space of functions of bounded variation.

Theorem 13. *$F \in BV[a, b] \iff F$ is the difference of two nondecreasing functions: $F = F_1 - F_2$.*

Lebesgue-Stieltjes integral: F increasing on $[a, b]$ then $\Lambda_f(u, v] = F(v) - F(u)$ extends to a positive Borel measure Λ_F on $[a, b]$, which is called the Lebesgue-Stieltjes measure.
Notation: $\int_{(a,b]} g \mathrm{d}\Lambda_F$ or $\int_{(a,b]} g(x)\mathrm{d}F(x)$ for the Lebesgue-Stieltjes integral.
Careful if F has a jump in t, then $\Lambda_F(\{t\}) = F(t) - F(t-)$.

An idea for quadratic variation is $\sum(F(S_i) - F(S_{i-1}))^2$, but we want more.
Given $\pi(t) = \{0 = t_0, \ldots, t_m = t\}$ a mesh on $[0, t]$ and process Y. Let $V_y^2(\pi(t)) = \sum_{i=0}^{m-1} \left|Y_{t_{i+1}}(\omega) - Y_{t_i}(\omega)\right|^2$.
We say that V_Y^2 converges in probability to process Z if $\forall \epsilon > 0 \exists \delta > 0 : \forall t > 0, \forall \pi(t), \mathrm{mesh}(\pi) < \delta \Rightarrow P(|V_Y^2(\pi(t)) - Z_t| > \epsilon) < \epsilon$
Notation: $[Y]_t = \lim_{\mathrm{mesh}(\pi)\to 0} V_Y^2(\pi(t))$ in probability.

Definition 19. $[Y] = ([Y]_t)_{t\geq 0}$ is called the *quadratic variation process of* Y *if*

- the limit exists.
- There exists a version of $[Y]$ s.t. $\forall\omega : t \to [Y]_t(\omega)$ is nondecreasing.

Definition 20. $[X, Y] = \frac{1}{4}[X + Y] - \frac{1}{4}[X - Y]$ if the right hand side exists.

$$\lim_{\mathrm{mesh}\to 0} \sum_i (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i}) = [X, Y]_t$$

where we use the fact that $\frac{1}{4}(a + b)^2 - \frac{1}{4}(a - b)^2 = ab$
Also: $[X, Y]_t = \frac{1}{2}([X + Y]_t - [X] - [Y])$

Theorem 14. *If X, Y are cadlag and $[X, Y]$ exists then $[X, Y]$ has a cadlag modification and $\Delta[X, Y]_t = (\Delta X_t)(\Delta Y_t)$. Here $\Delta Z_t = Z_t - Z_{t-}$ for Z cadlag.*

Theorem 15. $|[X, Y]_t - [X, Y]_s| \leq ([X]_t - [X]_s)^{\frac{1}{2}}([Y]_t - [Y]_s)^{\frac{1}{2}}$

Theorem 16 (Kunita-Watanabe inequality)**.** *Assume that $[X], [Y], [X, Y]$ exist and are right-continuous. Then for bounded and measurable functions $G, H : [0, T] \times \Omega \to \mathbb{R}$*

$$\left|\int_{[0,T]} G(t,\omega)H(t,\omega)d[X,Y]_t(\omega)\right| \leq \left(\int_{[0,T]} G(t,\omega)^2 d[X]_t(\omega)\right)^{1/2} \left(\int_{[0,T]} H(t,\omega)^2 d[Y]_t(\omega)\right)^{1/2}$$

Remark: by a Radon-Nikodym derivative this result also holds iwth

$$\left| \int_{[0,T]} G(t,\omega) H(t,\omega) |\Lambda_{[X,Y](\omega)}| \mathrm{d}t \right|$$
$$\leq \left(\int_{[0,T]} G(t,\omega)^2 \mathrm{d}[X]_t(\omega) \right)^{1/2} \left(\int_{[0,T]} H(t,\omega)^2 \mathrm{d}[Y]_t(\omega) \right)^{1/2}$$

3 Brownian motion

Definition 21. Let $(\Omega, \mathcal{F}, P)$ be a probability space with filtration $(\mathcal{F}_t)_{t\geq 0}$. A process $(B_t)_{t\geq 0}$ is called a *one-dimensional Brownian motion w.r.t.* $(\mathcal{F}_t)_{t\geq 0}$ if:

1. For almost all $\omega \in \Omega : t \to B_t(\omega)$ is continuous.
2. $\forall 0 \leq s \leq t$, $B_t - B_s$ is independent of $\mathcal{F}_s$ and has a normal distribution with $\mathbb{E}\left[B_t - B_s\right] = 0$ and $\mathbb{E}\left[(B_t - B_S)^2\right] = t - s$

If additionally 3. $B_0 = 0$ a.s. then B is called a standard Brownian motion.

Theorem 17. *Assume* $(\Omega, \mathcal{F}, P)$ *is rich enough. Then there exists a process* $(B_t)_{t\geq 0}$ *such that* $(B_t)_{t\geq 0}$ *is a standard Brownian Motion w.r.t.* $(\mathcal{F}_t)_{t\geq 0}$

Two pages about the construction of Brownian Motion - Not relevant I think.

Theorem 18. *Let* $(B_t)_{t\geq 0}$ *be a Brownian Motion w.r.t.* $(\mathcal{F}_t)_{t\geq 0}$*. Then* $\forall s \leq t$ *we have that* $\mathbb{E}\left[B_t|\mathcal{F}_s\right] = B_s$ *and* $\mathbb{E}\left[B_t^2 - t|\mathcal{F}_s\right] = B_s^2 - s$

Proof. We start with noticing that $\mathbb{E}\left[B_t - B_s|\mathcal{F}_s\right] = \mathbb{E}\left[B_t - B_s\right] = 0$. Therefore $\mathbb{E}\left[B_t|\mathcal{F}_s\right] = \mathbb{E}\left[B_t - B_s + B_s|\mathcal{F}_s\right] = B_s$. And $\mathbb{E}\left[(B_t - B_s)^2|\mathcal{F}_s\right] = t - s$ thus $\mathbb{E}\left[B_t^2 - 2B_tB_s + B_s^2|\mathcal{F}_s\right] = t - s$ and $\mathbb{E}\left[B_tB_s|\mathcal{F}_s\right] = B_s\mathbb{E}\left[B_t|\mathcal{F}_s\right] = B_s^2$. Conclusion: $\mathbb{E}\left[B_t^2|\mathcal{F}_s\right] - B_s^2 = t - s$ $\square$

Theorem 19. $[B]_t = t$*, moreover for all partitions* π *we have that*

$$\mathbb{E}\left[\left(\sum_{i=0}^{m(\pi)-1}(B_{t_{i+1}} - B_{t_i})^2 - t\right)^2\right] \leq 2t mesh(\pi)$$

Thus $\sum_{i=0}^{m(\pi)-1}(B_{t_{i+1}} - B_{t_i})^2 \to t$ *in* $L^2(p)$ *and in* P *as* $mesh(\pi) \to 0$.

Theorem 20. *Almost surely for all* $T > 0$*, the path* $t \mapsto B_t(\omega)$ *is not a member of* $BV[0, T]$.

4 Uniform integrability and Martingales

4.1 Uniform integrability

Definition 22. A collection C of random variables is called *uniformly integrable (UI)* if

$$\lim_{r\to\infty} \sup_{Z\in C} \int_{\{|Z|>r\}} |Z| \mathrm{d}P = 0$$

Example 8. If $X \in L^1$, then $C = \{X\}$ is UI.

Example 9. If $X \in L^1$ then $C = \{Z : \Omega \to \mathbb{R} : |Z| \leq |X| \text{ a.s. } \}$ is UI.

Theorem 21. *Let $p > 1$. If $C \subseteq L^p$ and $K := \sup_{Z\in C} ||Z||_{L^p} < \infty$ then C is UI.*

Example 10. $\Omega = [0, 1]$, P is Lebesgue-measure. $X_n = n\mathbf{1}_{[0,\frac{1}{n}]}$, $n \geq 1$. Then $C = \{X_n : n \in \mathbb{N}\}$ is not UI.
Indeed, given $r > 0$ choose $n > r$. Then $\int_{\{|X_n|>r\}} |X_n| \mathrm{d}P = \int |X_n| \mathrm{d}P = 1$. Thus $\sup_{X\in C} \int_{\{|X_n|>r\}} |X_n| \mathrm{d}P = 1$ for all $r > 0$.

Theorem 22. *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $X \in L^1(P)$ and define $C := \{\mathbb{E}[X|\mathcal{G}] : \mathcal{G} \subseteq \mathcal{F}\}$. Then C is uniformly integrable.*

Theorem 23 (Bounded convergence theorem)**.** *Assume $X_n \to X$ in probability. Assume $\exists K > 0 : \forall n \in \mathbb{N}, \forall \omega \in \Omega |X_n(\omega)| \leq K$, then $X_n \to X$ in L^1*

Theorem 24. *Let $X_n, X \in L^1$.*

$$X_n \to X \text{ in } L^1 \iff \begin{cases} X_n \to X \text{ in probability.} \\ \{X_n : n \geq 1\} \text{ is UI.} \end{cases}$$

4.2 Martingales

Definition 23. $(M_t)_{t\geq 0}$ is called a *martingale* w.r.t. $(\mathcal{F}_t)_{t\geq 0}$ if

1. $M_t \in L^1(P)$
2. (M_t) is $(\mathcal{F})_t$-adapted.
3. $\forall 0 \leq s < t : \mathbb{E}[M_t|\mathcal{F}_S] = M_s$ almost surely

Submartingale: Replace 2. by $\mathbb{E}[M_t|\mathcal{F}_S] \geq M_s$
Supermartingale: Replace 2. by $\mathbb{E}[M_t|\mathcal{F}_S] \leq M_s$

Note that $\mathbb{E}[M_t|\mathcal{F}_s] \geq M_s \iff \forall A \in \mathcal{F}_s \mathbb{E}[\mathbf{1}_A M_t] \geq \mathbb{E}[\mathbf{1}_A M_s]$
M is called *square integrable* if $\forall t \geq 0 : \mathbb{E}\left[M_t^2\right] < \infty$. The discrete definition is analogue.

Theorem 25. *If $(M_t)_{t\geq 0}$ is a martingale and ϕ is convex and $\forall t > 0 : \phi(M_t) \in L^1$ then $\phi(M_t)$ is a submartingale.*

Proof. Jensen's inequality for $s < t$: $\mathbb{E}[\phi(M_t)|\mathcal{F}_s] \geq \phi(\mathbb{E}[M_t|\mathcal{F}_s]) = \phi(M_s)$. $\square$

4.3 Optional stopping

We extend the times used in the definition of martingales to stopping times.
Notation: $x \wedge y = \min\{x, y\}$ and $x \vee y = \max\{x, y\}$.
First the discrete case:

Theorem 26 (Lemma 3.4). *Let M be a submartingale. Assume that τ and σ are stopping times whose values lie in an ordered countable set $\{s_1 < s_2 < s_3 < \dots\} \cup \{\infty\}$ where $s_n \to \infty$. Then for any $T < \infty$,*

$$\mathbb{E}\left[M_{\tau \wedge T} | \mathcal{F}_\sigma\right] = M_{\sigma \wedge \tau \wedge T}$$

Theorem 27 (Lemma 3.5). *Let M be a submartingale with right-continuous paths and $T < \infty$. Let ρ be a stopping time with $P(\rho < T) = 1$. Then:*

$$\mathbb{E}\left[M_\rho\right] \leq 2\mathbb{E}\left[M_T^+\right] - \mathbb{E}\left[M_0\right]$$

so $M_\rho \in L^1$.

Theorem 28. *Let M be a right-continuous submartingale. Let σ, τ be stopping times, $T < \infty$. Then $\mathbb{E}\left[M_{\tau \wedge T} | \mathcal{F}_\sigma\right] \geq M_{\sigma \wedge \tau \wedge T}$. Note the integrability by lemma 3.5*

Theorem 29 (Corollary 3.7). *Suppose $(M_t)_{t \geq 0}$ is a right-continuous (sub)martingale and τ is a stopping time. Then $M^\tau = (M_{t \wedge \tau})_{t \geq 0}$ is a right-continuous (sub)martingale. If M is an L^2 martingale, then M^τ is as well.*

Theorem 30 (Corollary 3.8). *Suppose M is a right-continuous submartingale. Let $\{\sigma(u) : u \geq 0\}$ be nondecreasing, $[0, \infty)$-values process such that $\sigma(u)$ is a bounded stopping time for each u. Then $\{M_{\sigma(u)} : u \geq 0\}$ is a submartingale with respect to the filtration $\{\mathcal{F}_{\sigma(u)} : u \geq 0\}$*

5 Further investigating martingales

5.1 Inequalities and limits

Towards Doob's inequality:

Theorem 31 (Lemma 3.9)**.** *Let M be a submartingale, $0 < T < \infty$ and H a finite subset of $[0,T]$. Then for all $r > 0$*

$$P(\{\max_{t\in H} M_t \geq r\}) \leq r^{-1}\mathbb{E}\left[M_T^+\right]$$

and

$$P(\{\min_{t\in H} M_t \leq r\}) \leq r^{-1}(\mathbb{E}\left[M_T^+\right] - \mathbb{E}\left[M_0\right])$$

Theorem 32 (Doobs mean)**.** *Let M be a right-continuous submartingale and $0 < T < \infty$. Then for all $r > 0$:*

$$P(\{\sup_{t\in H} M_t \geq r\}) \leq r^{-1}\mathbb{E}\left[M_T^+\right]$$

and

$$P(\{\inf_{t\in H} M_t \leq r\}) \leq r^{-1}(\mathbb{E}\left[M_T^+\right] - \mathbb{E}\left[M_0\right])$$

Theorem 33 (Doob's Inequality)**.** *Let M be a nonnegative, right-continuous submartingale and $0 < T < \infty$. Then for $1 < p < \infty$*

$$\mathbb{E}\left[\sup_{0\leq t\leq T} M_t^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[M_T^p\right]$$

$$P\left(\sup_{0\leq t\leq T} M_t \geq C\right) \leq \frac{\mathbb{E}\left[M_T^p\right]}{C^p}$$

Example 11. For example if (N_t) is a right-continuous martingale, we can apply Doob's inequality on $M_t = |N_t|$.

Most important cases of martingale convergence: M_t is a martingale with $\sup_{t<\infty}\mathbb{E}\left[|M_t|\right] < \infty$ then $M_\infty = \lim_{t\to\infty} M_t$ exists almost surely and $M_\infty \in L^1$. Convergence need not be in L^1. This holds if and only if $\{M_t : t \geq 0\}$ is uniformly integrable.

5.2 Local martingales and semimartingales

Notation: For process X, τ a stopping time we denote with $X_t^\tau = X_{t\wedge\tau}$. X^τ is called the stopped process.

Definition 24. M_t is called a *local martingale* if

1. M_t is $(\mathcal{F}_t)$ adapted.

2. There exists a sequence of stopping times $(\tau_k)_{k=1}^{\infty}$ such that $\tau_1 \leq \tau_2 \leq \ldots, \tau_k \to \infty$ a.s. and $\forall k : M^{\tau_k}$ is a martingale.

$(\tau_k)_k$ is called a *localizing sequence* for M.
M is called a *local square integrable* martingale if 1., 2. and $M^{\tau_k} \in L^2$ for all k.

Remark: If M has continuous paths, we can take $\tau_k = \inf\{t \geq 0 : |M_t| \geq k\}$ as a localizing sequence. Moreover $|M_t^{\tau_k}| \leq k$

Definition 25. A cadlag process Y is called a *semimartingale* if there exists a local martingale M with $M_0 = 0$ and there exists a finite variation process V with $V_0 = 0$ such that $Y_t = M_t + V_t + Y_0$ for all $t \geq 0$.

Continuous semimartingale: if additionally M, V are continuous.

5.3 Quadratic variation for Semimartingales

Remember that $[B]_t = t$ for a Brownian Motion and $[B, Y]_t = 0$ if B, Y are independent Brownian Motions.

Theorem 34 (Theorem 3.26)**.** *Let M be a right-continuous local martingale, then $[M]$ exists and there is a version of $[M]$ which is:*

- *real-valued (so no ∞)*
- *right-continuous*
- *nondecreasing*
- *adapted*
- $[M]_0 = 0$

If M is an L^2-martingale then $\lim_{mesh(\pi)\to 0} \sum_{i=0}^{m(\pi)-1} |M_{t_{i+1}} - M_{t_i}|^2 \to [M]_t$ *is in L^1 and* $\mathbb{E}\left[[M]_t\right] = \mathbb{E}\left[M_t^2 - M_0^2\right]$
If M is continuous, then $[M]$ has a version, which is continuous.

Theorem 35 (Lemma 3.27)**.** *Let M be a right-continuous local martingale. Let τ be a stopping time. Then $[M^\tau] = [M]^\tau$. This means that for all $t \geq 0$: $[M^\tau] - T = [M]_{\tau \wedge t}$*

Theorem 36 (Theorem 3.28)**.** *If M is a right-continuous (local) L^2-martingale then $M^2 - [M]$ is as well.*

If M, N are right-continuous (local) L^2-martingales then $[M, N]$ also exists and $[M^\tau, N] = [M^\tau, N^\tau] = [M, N]^\tau$.
Moreover $MN - [M, N]$ is also a (local) L^2-martingale again.

Theorem 37 (Corallary 3.31)**.** *Let M be a cadlag local martingale, V a cadlag FV process $M_0 = V_0 = 0$, and $Y = Y_0 + M + V$ the cadlag semimartingale. Then $[Y]$ exists and is given by:*

$$[Y]_t = [M]_t + 2[M, V]_t + [V]_t$$

Furthermore, $[Y^\tau] = [Y]^\tau$

6 Spaces of martingales and Stochastic Integration

6.1 Spaces of martingales

From now on only continuous L^2-martingales $\mathcal{M}_2^C$ and sometimes local $\mathcal{M}_{2,loc}^C$.
Remind from analysis: $C[a,b]$ with $||f||_\infty = \sup_{t\in[a,b]} |f(t)|$ is complete. Furthermore $L^2(p)$ is complete. $||X||_{L^2} = \left(\mathbb{E}\left[|X|^2\right]\right)^{\frac{1}{2}}$
Possible norm on martingales on $[0,T]$ would be $||M_T||_{L^2}$. But note that for all $t \in [0,T] ||M_t||_{L^2} \leq ||M_T||_{L^2}$, even more: $|| \sup_{t\in[0,T]} |M_t| ||_{L^2} \leq 2||M_T||_{L^2}$
Thus $(M^{(n)})_{n\geq 1}$ sequence such that $M_T^{(n)}$ is Cauchy in $L^2(p)$ implies $\forall \epsilon > 0$

$$P\left(\sup_{t\in[0,T]} |M_t^{(n)} - M_t^{(m)}| \geq \epsilon\right) \leq \frac{\mathbb{E}\left[|M_T^{(n)} - M_T^{(m)}|^2\right]}{\epsilon^2}$$

by Doob's inequality. This is called $(M^{(n)})_{n\geq 1}$ is uniformly Cauchy in probability. After some calculations we find that $||M_T||_{L^2}$ could become ∞ for $T \to \infty$. Therefore we define

$$||M||_{\mathcal{M}_2^C} := \sum_{k=1}^{\infty} 2^{-k}(1 \wedge ||M_k||_{L^2})$$

but there are many other equivalent choices possible.
This is not a norm because $||aM||_{\mathcal{M}_2^C} \neq |a| \cdot ||M||_{\mathcal{M}_2^C}$ but $\mathrm{d}_{M_2}(M,N) = ||M - N||_{\mathcal{M}_2^C}$ is a metric.

Theorem 38 (Theorem 3.40)**.** *Let $(\mathcal{F}_t)$ be complete. Then $\mathcal{M}_2^C$ is a complete metric space under the metric $d_{\mathcal{M}_2}$.*

Theorem 39. *If $M^{(n)} \to M$ in $\mathcal{M}_2^C$, then:*

$$\forall T < \infty, \forall \epsilon > 0 : \lim_{n\to\infty} P(\sup_{t\in[0,T]} |M_t^{(n)} - M_t| \geq \epsilon) = 0$$

This is called uniform convergence on compact intervals.
Furthermore there exists a subsequence $(M^{(n_k)})$ and $\Omega_0 \subseteq \Omega$ such that $P(\Omega_0) = 1$ and for each $\omega \in \Omega_0, \forall T < \infty$

$$\lim_{n\to\infty} \sup_{0\leq t\leq T} |M_t^{(n_k)}(\omega) - M_t(\omega)| = 0$$

6.2 Stochastic integration of predictable processes

We only consider $\int X \mathrm{d}Y$ with Y continuous to simplify the presentation in the lectures.

Definition 26. ρ is the smallest σ-algebra which contains $(s,t] \times F$ with $0 \leq s < t < \infty, F \in \mathcal{F}_s$ and $\{0\} \times F_0$ with $F_0 \in \mathcal{F}_0$
ρ is called *predictable σ-algebra*
$(s,t] \times F$ is called *predictable rectangle.*

Theorem 40 (Lemma 5.1). *A process is ρ-measurable if and only if it can be approximated by (left)-continuous adapted processes*

Proof. We proof that a left-continuous adapted process X is ρ-measurable.
Rewrite $X_n(t,\omega) = X_0(\omega)\mathbf{1}_{\{0\}} + \sum_{i=0}^{\infty} X_{i2^{-n}}\mathbf{1}_{[i2^{-n},(i+1)2^{-n}]}(t)$
Now $\{X_n \in \mathcal{B}\} = \underbrace{\{0\} \times \{X_0 \in \mathcal{B}\}}_{\in\rho} \cup \bigcup_{i=0}^{\infty} \underbrace{(i2^{-n},(i+1)2^{-n}] \times \{X_{i2^{-n}} \in \mathcal{B}\}}_{\in\rho}$. Thus $\{X_n \in \mathcal{B}\} \in \rho$, thus X_n is ρ-measurable.
Also by left continuity $X_n \to X$ on $[0,\infty) \times \Omega$ thus X is ρ-measurable. □

Remarks: Not all right-continuous adapted processes are predictable.
$X : [0,\infty) \to \mathbb{R}$ with the Borel-measure is predictable.

Doleans measure: μ_M on ρ Let $M \in \mathcal{M}_2^C$ then Doleans measure is defined as:

$$\mu_M(A) = \int_\Omega \int_{[0,\infty)} \mathbf{1}_A(t,\omega)\mathrm{d}[M]_t(\omega)\mathrm{d}P(\omega)$$

The meaning of this formula is that first, for each fixed ω, the function $t \mapsto \mathbf{1}_A(t,\omega)$ is integrated by the Lebesgue-Stieltjes measure $\Lambda_{[M](\omega)}$ of the function $t \mapsto [M]_t(\omega)$. The resulting integral is a measurable function of ω, which is then averaged over the probability space.
Convention: $\Lambda_{[M](\omega)}(\{0\}) = 0$.
Note: $\mu_M([0,T] \times \Omega) = \mathbb{E}\left[[M]_t - [M]_0\right] = \mathbb{E}\left[M_t^2\right] - \mathbb{E}\left[M_0^2\right] < \infty$ thus μ_M is a σ-finite measure.

Example 12. Assume $(B_t)_t$ is a standard Brownian Motion and $\mu_B = m \otimes p$ where m is the Lebesgue measure. Indeed: $\mu_B(B) = \int_\Omega \int_{[0,\infty)} \mathbf{1}_A(t,\omega)\mathrm{d}t\mathrm{d}P(\omega) = m \otimes P(A)$

Definition 27. For $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ predictable:

$$||X||_{\mu_M,T} = \left(\int_{[0,T]\times\Omega} |X|^2 \mathrm{d}\mu_M\right)^{\frac{1}{2}} = \mathbb{E}\left[\int_{[0,T]} |X(t)|^2 \mathrm{d}[M]_t\right]$$

$\mathcal{L}_2 = \mathcal{L}_2(M,P)$ is the set of all predictable X such that $\forall T < \infty : ||X||_{\mu_M,T} < \infty$
A metric on $\mathcal{L}_2$ is defined as:

$$\mathrm{d}_{\mathcal{L}_2}(X,Y) = ||X - Y||_{\mathcal{L}_2}$$

with

$$||X||_{\mathcal{L}_2} = \sum_{k=1}^{\infty} 2^{-k}(1 \wedge ||X||_{\mu_M,k})$$

Here we identify processes which are μ_M almost everywhere equal.

Example 13. Let $(B_t)_{t\geq 0}$ be a Brownian Motion and X a predictable process. Then we have that $X \in \mathcal{L}_2$ if and only if

$$\forall T < \infty : X \in L^2((0,T] \times \Omega)$$

Example 14. Let $M \in \mathcal{M}_2^C$. If $\forall T < \infty \exists C_T, \forall \omega, t |X_t(\omega)| \leq C_T$ and X predictable, then $X \in \mathcal{L}(M, P)$.
Indeed,

$$\begin{aligned}
\mathbb{E}\left[\int_{[0,T]} |X(s)|^2 \mathrm{d}[M]_s\right] &\leq \mathbb{E}\left[\int_{[0,T]} C_T^2 \mathrm{d}[M]_S\right] \\
&= C_T^2 \mathbb{E}\left[[M]_T - [M]_0\right] \\
&= C_T^2 \mathbb{E}\left[M_T^2 - M_0^2\right] < \infty
\end{aligned}$$

6.3 Construction of the stochastic integral

Our goal is to define $(X \cdot M)_t := \int_{(0,t]} X \mathrm{d}M$ for $X \in \mathcal{L}_2(M, P)$

Step 1 $X \in \mathcal{S}_2$ a simple predictable process.

Step 2 Prove L^2-isometry for $X \cdot M$

$$\mathbb{E}\left[|(X \cdot M)_T|^2\right] = ||X||_{\mu_M, T} \text{ for } X \in \mathcal{S}_2$$

Step 3 Approximation/density argument for $X \in \mathcal{L}_2(M, P)$. Here completeness of $\mathcal{M}_2^C$ plays a crucial role.

Step 4 Localization: no integrability conditions on Ω

Step 5 Extension to continuous semimartingales.

Definition 28. A process X of the form:

$$\begin{cases} X_t(\omega) = \xi_0(\omega)\mathbf{1}_{\{0\}}(t) + \sum_{i=1}^{n-1} \xi_i(\omega)\mathbf{1}_{(t_i, t_{i+1}]}(t) \\ \text{with } 0 = t_0 < t_1 < \cdots < t_n \text{ and } \xi_i \text{ is } \mathcal{F}_{t_i}\text{-measurable.} \end{cases}$$

is called a simple predictable process, notation $X \in \mathcal{S}_2$

7 Stochastic Integration

7.1 Step 1,2 and 3

Definition 29. A process X of the form:

$$\begin{cases} X_t(\omega) = \xi_0(\omega)\mathbf{1}_{\{0\}}(t) + \sum_{i=1}^{n-1} \xi_i(\omega)\mathbf{1}_{(t_i,t_{i+1}]}(t) \\ \text{with } 0 = t_0 < t_1 < \cdots < t_n \text{ and } \xi_i \text{ is } \mathcal{F}_{t_i}\text{-measurable.} \end{cases}$$

is called a simple predictable process, notation $X \in \mathcal{S}_2$

Theorem 41 (Lemma 5.6)**.** *X of the form is indeed predictable*

Proof. By linearity it suffices to consider $\xi\mathbf{1}_{(a,b]}$ with ξ $\mathcal{F}$-measurable. Now approximate ξ by simple random variables to get predictable rectangles. Similarly for $\xi\mathbf{1}_{\{0\}}$ □

Definition 30. For X a simple predictable process and $M \in \mathcal{M}_2^C$ we define the *stochastic integral* to be:

$$(X \cdot M)_t(\omega) = \sum_{i=1}^{n-1} \xi_i(\omega)\left(M_{t_{i+1}\wedge t}(\omega) - M_{t_i\wedge t}(\omega)\right)$$

Remarks: The value at zero of X and M are irrelevant. Adding a $\mathcal{F}_0$-measurable random variable to M does not change the stochastic integral.
Two other notations: $\int_0^t X \mathrm{d}M$ and $I(X)$ for $X \cdot M$.

Theorem 42 (Lemma 5.8)**.**

1. *The stochastic integral does not depend on its representation.*
2. *The integral is linear.*

Theorem 43. *Let $X \in \mathcal{S}_2$, $M \in \mathcal{M}_2^C$, then $X \cdot M \in M_2^C$ and the following L^2-isometries hold:*

$$||(X \cdot M)_t||_{L^2(\Omega,P)} = ||X||_{L^2((0,t)\times\Omega,\mu_M)} \tag{1}$$
$$||X \cdot M||_{\mathcal{M}_2^C} = ||X||_{\mathcal{L}_2} \tag{2}$$

Now we continue with step 3:

Theorem 44 (Lemma 5.10)**.** *For any $X \in \mathcal{L}_2$ there exists a sequence $(X_n)_{n\geq 1} \in \mathcal{S}_2$ such that $\lim_{n\to\infty} ||X - X_n||_{\mathcal{L}_2} = 0$*

Definition 31. Take $M \in \mathcal{M}_2^C$ and $X \in \mathcal{L}_2(M)$. Choose $(X_n)_{n\geq 1} \in \mathcal{S}_2$ such that $||X - X_n||_{\mathcal{L}_2} \to 0$. Now we define the *stochastic integral* for X to be

$$(X \cdot M)_t = \lim_{n\to\infty} (X_n \cdot M)_t$$

Existence of limit. $(X_n)_{n\geq 1}$ exists by lemma 5.10. Also:

$$\begin{aligned} ||X_n \cdot M - X_m \cdot M||_{\mathcal{M}_2^C} &= ||(X_n - X_m) \cdot M||_{\mathcal{M}_2^C} \\ &= ||X_n - X_m||_{\mathcal{L}_2} \\ &\leq ||X_n - X||_{\mathcal{L}_2} + ||X - X_m||_{\mathcal{L}_2} \to 0 \end{aligned}$$

Thus $(X_n \cdot M)_{n\geq 1}$ is a Cauchy sequence in M_2^C hence converges by the completeness of M_2^C. Thus $\lim_{n\to\infty} X_n \cdot M$ exists in $\mathcal{M}_2^C$
Uniqueness: Take $Z_n \in \mathcal{S}_2$ such that $Z_n \to X$ in $\mathcal{L}_2$. Then

$$\begin{aligned}||X_n \cdot M - Z_n \cdot M||_{\mathcal{M}_2^C} &= ||(X_n - Z_n) \cdot M||_{\mathcal{M}_2^C} \\ &= ||X_n - Z_n||_{\mathcal{L}_2} \\ &\leq ||X_n - X||_{\mathcal{L}_2} + ||Z_n - X||_{\mathcal{L}_2} \to 0\end{aligned}$$

Thus $(Z_n \cdot M)_{n\geq 1}$ has the same limit as $(X_n \cdot M)_{n\geq 1}$ in $\mathcal{M}_2^C$. Thus $(X \cdot M)_t$ is unique up to indistinguishability.

Theorem 45 (Proposition 5.12)**.** *Let $M \in \mathcal{M}_2^C$, $X \in \mathcal{L}_2(M)$ then $\forall t < \infty$ $||(X \cdot M)_t||_{L^2(\Omega,P)} = ||X||_{L^2((0,t)\times\Omega,\mu_M}$ and $||X \cdot M||_{\mathcal{M}_2^C} = ||X||_{\mathcal{L}_2(M)}$
In particular, if $X = Y$, μ_M-almost surely, then $X \cdot M$ and $Y \cdot M$ are indistinguishable.*

Proof. Just take limits in lemma 5.9. Als use the reverse triangle inequality:

$$|||\phi|| - ||\psi||| \leq ||\phi - \psi||$$

□

Properties of the stochastic integral

Theorem 46 (Proposition 5.14)**.** *This proposition gives some properties of the stochastic integral:*

1. *Linearity:*
$$(\alpha X + \beta B) \cdot M = \alpha(X \cdot M) + \beta(Y \cdot M)$$

2. *For any $0 \leq u \leq v$,*
$$\int_{(0,t]} \mathbf{1}_{[0,v]} X \, dM = \int_{(0,v\wedge t]} X \, dM$$
and
$$\int_{(0,t]} \mathbf{1}_{(u,v]} X \, dM = (X \cdot M)_{v\wedge t} - (X \cdot M)_{u\wedge t} = \int_{(u\wedge t, v\wedge t]} X \, dM$$

3. *For $s < t$ we have a condition form of the isometry:*
$$\mathbb{E}\left[((X \cdot M)_t - (X \cdot M)_s)^2 | \mathcal{F}_s\right] = \mathbb{E}\left[\int_{(s,t]} X_u^2 d[M]_u | \mathcal{F}_s\right]$$

Theorem 47 (Proposition 5.19)**.** *Let $M, N \in \mathcal{M}_2$, $\alpha, \beta \in \mathbb{R}$, and $X \in \mathcal{L}_2(M,P) \cap \mathcal{L}_2(N,P)$. Then $X \in \mathcal{L}_2(\alpha M + \beta N, P)$ and*

$$X \cdot (\alpha M + \beta N) = \alpha(X \cdot M) + \beta(X \cdot N)$$

8 Stochastic Integration

8.1 Step 4 and 5

Last time we considered $M \in \mathcal{M}_2^C$, the continuous L^2-martingale and $(X \cdot M) \in \mathcal{M}_2^C$ for $X \in \mathcal{L}^2(M)$.
Here $X \in \mathcal{L}_2(M) \iff \forall T < \infty X \in L^2((0,T) \times \Omega, \mathrm{d}\mu_M)$

Theorem 48 (Proposition 5.16)**.**

$$((\mathbf{1}_{[0,\tau]}X) \cdot M)_t = (X \cdot M)_{\tau \wedge t} = (X \cdot M^\tau)_t$$

Today we only want to assume;

- $M \in \mathcal{M}_{2,\mathrm{loc}^C}$
- $X \in L^2((0,T), [M])$ almost surely for all $T < \infty$

but the problem is that there is no integrability in Ω.

Example 15. $X_t = e^{B_t^4}$, $M = X \cdot B$ should exist and what is M? And what about $(Y \cdot M)_t$?

Recall that $M \in \mathcal{M}_{2,\mathrm{loc}}^C \iff$ there exists a localizing sequence $\sigma_k \uparrow \infty$ such that $M^{\sigma_k} \in \mathcal{M}_2^C$

Definition 32. Let $M \in \mathcal{M}_{2,\mathrm{loc}}^C$. We say $X \in \mathcal{L}(M, P)$ if X is predictable and there exists stopping times $0 \leq \tau_1 \leq \tau_2 \leq ...$ such that

1. $P(\lim_{k \to \infty} \tau_k = \infty) = 1$
2. $M^{\tau_k} \in \mathcal{M}_2^C$ for all k
3. $\mathbf{1}_{[0,\tau_k]}X \in \mathcal{L}(M^{\tau_k}, P$ for all k.

In this case (τ_k) is called a localizing sequence for $(X \cdot M)$.
Remark: $\mathbf{1}_{[0,\tau_k]}$ is predictable, because it is adapted and left-continuous.

Now the idea is to define $(X \cdot M)$ locally:

$$Y^k = (\mathbf{1}_{[0,\tau_k]}X \cdot M^{\tau_k})$$

and let $k \to \infty$. Here k is an index.

Theorem 49 (Lemma 5.22)**.** *$M \in \mathcal{M}_{2,loc}^C$, X predictable. If τ, σ are stopping times such that $M^\sigma, M^\tau \in \mathcal{M}_2^C$ and $\mathbf{1}_{[0,\sigma]}X \in \mathcal{L}_2(M^\sigma)$, $\mathbf{1}_{[0,\tau]}X \in \mathcal{L}_2(M^\tau)$.*
Define :

$$Z_t := \int_{(0,t]} \mathbf{1}_{(0,\sigma]}X \, dM^\sigma, \qquad W_t := \int_{(0,t]} \mathbf{1}_{(0,\tau]}X \, dM^\tau$$

then $Z^{\sigma \wedge \tau} = W^{\sigma \wedge \tau}$ where we mean that the two processes are indistinguishable.

By lemma 5.22 we have that $\forall k, m \in \mathbb{N}$ almost surely and $\forall t \geq 0$

$$Y^k_{t \wedge \tau_k \wedge \tau_m} = Y^m_{t \wedge \tau_k \wedge \tau_m} \tag{3}$$

Now let $\Omega_0 = \{\omega \in \Omega : \lim_{k \to \infty} = \infty, \forall k, m \in \mathbb{N}, \forall t \geq 0 \quad (3) \text{ holds. } \}$. Then $P(\Omega_0) = 1$ by countability of $\mathbb{N} \times \mathbb{N}$.

Definition 33. Let $M \in \mathcal{M}_{2,\text{loc}}^C, X \in \mathcal{L}(M,P)$ and (τ_k) a localizing sequence for (X, M).
Now define the *stochastic integral* $\forall \omega \in \Omega_0, (X \cdot M)_t(\omega) = Y_t^k(\omega), t \leq \tau_k(\omega)$ and $X \cdot M = 0$ for $\omega \notin \Omega_0$

Remarks:

- The stochastic integral is well defined since $\tau_k(\omega) \to \infty$ and if $t \leq \tau_k(\omega) \wedge \tau_m(\omega)$, then

$$Y_t^k(\omega) = Y_{t\wedge\tau_k\wedge\tau_m}^k(\omega) = Y_{t\wedge\tau_k\wedge\tau_m}^m(\omega) = Y_t^m(\omega)$$

- $(X \cdot M)_t^{\tau_k} = (X \cdot M)_{t\wedge\tau_k} = Y_{\tau_k\wedge t}^K = (Y^k)_t^{\tau_k}$ which is in M_2^C. Thus $X \cdot M \in \mathcal{M}_{2,\text{loc}}^C$ with localizing sequence τ_k

- If we would use another localizing sequence $(\sigma_j)_{j\geq 1}$ for (X, M), this would yield the same $(X \cdot M)$ by lemma 5.22

Example 16 (Example 5.26)**.** Let B be a Brownian Motion, then

$$X \in \mathcal{L}(B,P) \Longleftrightarrow X \text{ predictable and } \forall T < \infty, \text{ a.s. } \int_0^T |X(t,\omega)|^2 \mathrm{d}t < \infty$$

Theorem 50 (Corollary 5.29)**.** *Let $M \in \mathcal{M}_{2,loc}^C$ and X continuous and adapted then $X \in \mathcal{L}(M,P)$ and hence $X \cdot M$ is well-defined*

Proof. Define $\sigma_k := \inf\{t \geq 0; |X_t| \geq k\}$ and $\tau_k := \inf\{t \geq 0 : |M_t| \geq k\}$. Now $\sigma_k \wedge \tau_k$ is a localizing sequence for $(X \cdot M)$ ☐

Standard properties of L^2-integral extend to the localized setting:

- Linearity continues to hold

- Interchanging stopping times, if $X \in \mathcal{L}(M), Y \in \mathcal{L}(N), \tau$ a stopping time. If almost surely $X_t(\omega) = Y_t(\omega)$ and $M_t(\omega) = N_t(\omega)$ for $t \leq \tau(\omega)$ then $(X \cdot M)_{t\wedge\tau} = (Y \cdot N)_{t\wedge\tau}$

Theorem 51 (Proposition 5.32)**.** *Let $M \in \mathcal{M}_{2,loc}^C$ and X be continuous and predictable. Now assume that for all $n \in \mathbb{N}$ $0 \leq \tau_0^n \leq \tau_1^n \leq \dots$ are stopping times such that almost surely $\delta_n = \sup_i \tau_{i+1}^n - \tau_i^n \to 0$ if $n \to \infty$.*
Define $R_n(t) = \sum_{i=0}^\infty X(\tau_i^n)\left(M(\tau_{i+1}^n \wedge t) - M(\tau_i^n \wedge t)\right)$, then $R_n \to X \cdot M$ uniform, in probability on compact time intervals.

8.2 Semimartingale integrators

Let Y be a continuous semimartingale, $Y_t = Y_0 + M_t + V_t$ with $M_0 = V_0 = 0$.
Technical condition: there exist stopping times σ_n such that $\forall n \in \mathbb{N} : \mathbf{1}_{(0,\sigma_n)}X$ is bounded, where X_0 is not relevant.

Definition 34. Let Y be a semimartingale and let X be a predictable process for which the technical condition is satisfied. Then we define the integral of X with respect to Y as the process

$$\int_{(0,t]} X \mathrm{d}Y = \underbrace{\int_{(0,t]} X \mathrm{d}M}_{\text{Stochastic integral in}\mathcal{M}^{C}_{2,\text{loc}}} + \underbrace{\int_{(0,t]} X \mathrm{d}\Lambda_v(\mathrm{d}s)}_{\text{Stieltjes integral for fixed } \omega}$$

Thus $X \cdot Y$ is a semimartingale again.
By the next lemma the decomposition of Y is unique, thus the stochastic integral is well defined. The well-definedness follows from the uniqueness of decomposition for continuous semimartingales $Y_t = Y_0 + M_t + V_t = Y_0 + N_t + W_t$. Thus $M_t - N_t \in \mathcal{M}^{C}_{2,\text{loc}} = W_t - V_t$. By the next result we show that $M_t = N_t$ and $W_t = V_t$.

Theorem 52 (Lemma)**.** *If $M \in \mathcal{M}^{C}_{2,loc}$ has finite variation, then $M = M_0$*

Rest of 5.3 is selfstudy Proposition 5.36 is not needed because of the above lemma. Non-continuous case is to complicated for this lecture.

9 Itô's lemma

9.1 Quadratic Covariation

The lecture starts with repeating some information about quadratic covariation. I have not reposted the old results, but here are the new results:
When the Quadratic Covariation (QCV) exists it behaves like an innerproduct

$$[\alpha X + \beta Y, Z] = \alpha[X, Z] + \beta[Y, Z]$$

Theorem 53 (Lemma 5.54). *M_n, M, N_n, N are L^2-martingales and $0 \leq T < \infty$. Furthermore suppose that $M_n(T) \to M(T)$ and $N_n(T) \to N(T)$ in L^2. Then $\mathbb{E}\left[\sup_{0 \leq t \leq T} |[M_n, N_n]_t - [M, N]_t|\right] \to 0$ as $n \to \infty$*

Theorem 54. *Let $M, N \in \mathcal{M}_{2,loc}$, $G \in \mathcal{L}(M, P)$, $H \in \mathcal{L}(N, P)$. Then $[G \cdot M, H \cdot N]_t = \int_{(0,t]} G_s H_s \, d[M, N]_s$*

9.2 Change of integrator/Substitution rule

Theorem 55 (Proposition 5.58). *Let $M \in \mathcal{M}_{2,loc}, G \in \mathcal{L}(M, P)$. We already know that $N := G \cdot M \in \mathcal{M}_{2,loc}$. Let $H \in \mathcal{L}(N, P)$. Then $HG \in \mathcal{L}(M, P)$ and $H \cdot N = (HG) \cdot M$*

Theorem 56 (Corollary 5.59). *Let Y be a cadlag semimartingale and H be predictable satisfying (5.66): there exists a sequence (σ_N) with $\sigma_n \uparrow \infty$ a.s. such that $\mathbf{1}_{(0,\sigma_n]} H$ is bounded for each n.
We know that $X = H \cdot Y$ is a cadlag semimartingale. Let G be predictable satisfying (5.66), then $\int G \, dX = \int GH \, dY$*

Theorem 57 (Theorem 5.62). *Let Y, Z be cadlag semimartingales. G, H predictable satisfying (5.66). Then $[G \cdot Y, H \cdot Z]_t = \int_{(0,t])} G_s H_s \, d[Y, Z]_t$*

Theorem 58 (Proposition 5.63). *Let Y, Z be continuous semimartingales and G an adapted, continuous process. Let $\pi = \{0 = t_0 < t_1 < t_2 < \ldots, t_i \uparrow \infty\}$ a partition of $[0, \infty)$.
Then $R_t(n) = \sum_{i=1}^{\infty} G_{t_i}(Y_{t_{i+1} \wedge t} - Y_{t_i \wedge t})(Z_{t_{i+1} \wedge t} - Z_{t_i \wedge t})$ converges to $\int_0^t G_s \, d[Y, Z]_s$ as mesh(π) $\to 0$
This is what we call convergence in probability uniformly on compact intervals.*

Theorem 59 (Theorem 5.60). *Let Y, Z be continuous semimartingales, then $[Y, Z]$ exists as continuous adapted FV process and:*

1. *$[Y, Z]_t = Y_t Z_t - Y_0 Z_0 - \int_0^t Y_s \, dZ_s - \int_0^t Z_s \, dY_s$ which is the stochastic version of integration by parts.*
2. *YZ is continuous semimartingale.*
3. *For continuous H $\int_0^t H_s \, d(YZ)_s = \int_0^t H_s Y_S \, dZ_s + \int_0^t H_s Z_s \, dY_s + \int_0^t H_s \, d[Y, Z]_s$*

9.3 Itô's lemma

Theorem 60 (Theorem 6.1.0). *Let $0 < T < \infty$ and :*

1. *$f \in C^2(\mathbb{R})$, i.e. has a continuous 2nd derivative.*
2. *Y is a continuous semimartingale with quadratic variation $[Y]$*

Then,

$$f(Y_t) = f(Y_0) + \int_0^t f'(Y_s)dY_s + \tfrac{1}{2}\int_0^t f''(Y_s)d[Y]_s \qquad \forall 0 \le t \le T$$

Both sides are continuous processes and $' =' $ means that both sides are indistinguishable on $[0,T]$, i.e., $\exists \Omega_0, P(\Omega_0) = 1$ such that $\forall \omega \in \Omega_0$ the equality holds for all $0 \le t \le T$.

Generalizations of theorem 6.1

2* Y is cadlag instead of continuous. Then the integrals become: $\int_0^t f'(Y_{s-})\mathrm{d}Y_s + \frac{1}{2}\int_0^t f''(Y_{s-})\mathrm{d}[Y]_s$. An extra term/sum involving the jumps is needed:

$$\sum_{s\in(0,t]} \left\{ f(Y_s) - f(Y_{s-}) - f'(Y_{s-})\Delta Y_s - \tfrac{1}{2} f''(Y_{s-})(\Delta Y_s)^2 \right\}$$

where the sum converges absolutely for a.e. ω. All processes are now cadlag instead of continuous.

1* $f \in C^2(D)$ where D is open in $\mathbb{R}$. We now need that $Y[0,T] \subseteq D$

3* Note that 1* and 2* combined is not enough for the theorem.

Remark 6.2: $f(Y_t)$ is a continuous semimartingale.

Theorem 61 (Corollary 6.3). **(b)** *If Y is of bounded variation on $[0,T]$ and continuous then $f(Y_t) = f(Y_0) + \int_0^t f'(Y_s)dY_s$. This is the regular, non-stochastic integration theory.*

(c) *If $Y_t = Y_0 + B_t$, where B is a standard Brownian Motion independent of Y_0 then*

$$f(B_t) = f(Y_0) + \int_0^t f'(Y_0 + B_s)dB_s + \tfrac{1}{2}\int_0^t f''(Y_0 + B_S)ds$$

9.4 Itô's formula in time and space

Theorem 62 (Theorem 6.1.1). *Let $0 < T < \infty$, $f \in C^{1,2}([0,T],\mathbb{R})$ i.e. $f(t,x)$ is continuous differentiable in 1st variable and twice continuous differentiable in the 2nd varbiable. Furthermore Y is a continuous semimartingale with quadratic variation $[Y]$. Then:*

$$f(t,Y(t)) = f(0,Y(0)) + \int_0^t f_t(s,Y(s))ds + \int_0^t f_x(s,Y(s))dY(s) + \tfrac{1}{2}\int_0^t f_{xx}(s,Y(s))d[Y]_s$$

We now generalize this theory to the d-dimension vector valued variant.

Theorem 63 (Theorem 6.5)**.** *Let $0 < T < \infty$, $f \in C^{1,2}([0,T], D)$ where D is open in $\mathbb{R}^d$. Furthermore Y is $\mathbb{R}^d$-valued and a continuous semimartingale such that $\overline{Y([0,T])} \subseteq D$ almost surely. Then:*

$$f(t, Y(t)) = f(0, Y(0)) + \int_0^t f_t(s, Y(s))\, ds + \sum_{i=1}^d \int_0^t f_{x_i}(s, Y(s))\, dY(s)$$
$$+ \frac{1}{2} \sum_{1 \le i,j \le d} \int_0^t f_{x_i x_j}(s, Y(s))\, d[Y_i, Y_j](s)$$

Short hand notation:

$$\mathrm{d}f(t, Y(t)) = f_t(t, Y(t))\mathrm{d}t + \sum_{i=1}^d f_{x_i}(t, Y(t))\mathrm{d}Y(t)$$
$$+ \frac{1}{2} \sum_{1 \le i,j \le d} f_{x_i x_j}(t, Y(t))\mathrm{d}[Y_i, Y_j](t)$$

We have the special case that $Y(t) = B(t) = (B_1(t), \dots, B_d(t))$, the $d-$dimensional Brownian Motion. Notation:

- $f \in C^{1,2}(([0,T] \times \mathbb{R}^d)$
- $\nabla_x f = (f_{x_1}, \dots, f_{x_d})$ the gradient vector
- $\Delta_x f = \nabla_x \cdot \nabla_x f = \sum_{i=1}^d f_{x_i, x_i}$, the Laplacian

Theorem 64 (Corollary 6.7)**.** *Let $B(t)$ be d-dimensional Brownian Motion, $f \in C^{1,2}([0,T] \times \mathbb{R}^d)$*
Then

$$f(t, B(t)) = f(0, B(0)) + \int_0^t \big(f_t(s, B(s) + \tfrac{1}{2}\Delta_x f(s, B(s))\big)\, ds$$
$$+ \int_0^t \nabla_x f(s, B(s))\, dB(s)$$

10 Itô's formula

The continuous semimartingale class is preserved after transformation of $f(t, Y(t)$.
This may not be the case if we work with martingales.
For $f \in C^1(\mathbb{R})$ such that $F(x) = \int_0^x f(y)\mathrm{d}y$ we have that $\int_0^t f(B_s)\mathrm{d}B_s = F(B_t) - \frac{1}{2}\int_0^t f'(B_s)\mathrm{d}s$, which is the path-wise interpretation.
The short hand notation is $\mathrm{d}f(B_t) = f'(B_t)\mathrm{d}B_t + \frac{1}{2}f''(B_t)\mathrm{d}t$. This notation has no meaning, only through the integrated version.
Application of Itô formula: Beautiful and useful results can be derived from special choices of f.

Preservation of Martingale property
Suppose that $Y(t)$ is continuous martingale and $f \in C^{1,2}([0,T] \times \mathbb{R})$.
Ito: $f(t, Y(t)) = f(0, Y(0)) + \int_0^t \left(f_t + \frac{1}{2}f_{xx}\right)(s, Y(s))\mathrm{d}[Y]_s + \int_0^t f_x(s, Y(s))\mathrm{d}Y(s)$.
If 2nd term on the right hand side is zero, then it is at least a local martingale. When is $\int_0^t f_x(s, Y(s))\mathrm{d}Y(s)$ a martingale? One sufficient condition is for example, Y is continuous L^2-martingale and $f_x(s, Y(s)) \in \mathcal{L}_2(M, P)$.

Theorem 65 (Lemma 6.9)**.** *Suppose $f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$ and $f_t + \frac{1}{2}f_{xx} = 0$. Let B_t be a one-dimensional standard Brownian Motion. Then $f(t, B_t)$ is local L^2-martingale. If further $\int_0^T \mathbb{E}\left[f_x^2(t, B_t)\right] dt < \infty$ then $f(t, B_t)$ is an L^2-martingale on $[0,T]$*

This lemma can be extended to the d-dimensional Brownian Motion.
When is a local martingale a martingale?

Exercise 3.7 X a nonnegative local martingale with $\mathbb{E}[X_0] < \infty$. X is a martingale $\iff \mathbb{E}[X_t] = \mathbb{E}[X_0]$ for all $t > 0$

Exercise 3.8 M is a right-continous local martingale and $M_t^* \in L^1(P)$ then M is a martingale

Corollary A continuous local martingale which is bounded a.s. is a martingale.

Example 17. Some applications of Lemma 6.9:

- $f(t, x) = x^2 - t \Rightarrow B_t^2 - t$ is a martingale.
- $f(t, x) = e^{\alpha x - \frac{1}{2}\alpha^2 t}$ then $f_x = \alpha f, f_{xx} = \alpha^2 f$ and $f_t = -\frac{1}{2}\alpha^2 f = -\frac{1}{2}f_{xx}$ and therefore $e^{\alpha B_t - \frac{1}{2}\alpha^2 t}$ is a martingale.

Example 18 (Exit time of Brownian Motion with drift.)**.** We have $X_t = \mu t + \sigma B_t$ with $\mu \in \mathbb{R}, \sigma \in \mathbb{R}, \sigma \neq 0$. $\tau = \inf\{t > 0 : x_t = a \text{ or } x_t = b\}$ where $a < 0, b > 0$.
What is $P(X_\tau = b)$?

Propositions 6.11 and 6.12 are about recurrent/transience properties of Brownian Motion.

- One dimensional BM is (point) recurrent.
- Two dimensional BM is not point recurrent, but neighbourhood recurrent.
- d-dimensional BM ($d \geq 3$) is transient.

Theorem 66 (Theorem 6.14)**.** *Let M be a continuous $\mathbb{R}^d$-valued local martingale and $X(t) = M(t) - M(0)$ such that $X(0) = 0$. Then X is a standard Brownian Motion relative to $\mathcal{F}_t$ iff $[x_i, X_j](t) = \delta_{i,j}t$ in particual X is independent of $\mathcal{F}_0$*

10.1 SDEs

Recall ordinary differential equations (ODE). For example it may be of the form $\dot{x} = f(t, x)$, equivalently $\mathrm{d}x(t) = f(t, x(t))\mathrm{d}t$.
SDE: The stochastic variant will involve in the simplest case a $\mathrm{d}B_t$ term. For example, $\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t$.
We have seen earlier this type of equations as short hand notation for Ito formula. But there given $X_t = f(t, B_t)$ we derived this short hand notation formula.
Now we have to do the reverse. Given this 'formula'/SDE, does there exist a process X_t which satisfy this equation? Recall that this short-hand notation must be interpreted through integral form. That is still the case.

Definition 35. Let $(\Omega, \mathcal{F}, P)$ be a complete filtered probability space, and (B_t) is a standard Brownian motion defined on it. Suppose $\mu, \sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ are measurable and η is an $\mathcal{F}_0$-measurable random variable. A stochastic process $(X_t), t \in [0, T]$ defined on $(\Omega, \mathcal{F}, P)$ is called a *strong solution* of the SDE: $\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t$ with initial condition $X_0 = \eta$ if the following assertions are true:

1. X_t is continuous and $\mathcal{F}_t$-adapted
2. $\int_0^T |\mu(t, X_t)|\mathrm{d}t + \int_0^T |\sigma(t, X_t)|^2\mathrm{d}t < \infty$ almost surely.
3. For each $t \in [0, T] : X_t = \eta + \int_0^t \mu(s, X_s)\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}B_S$ almost surely.

Note that condition 2. assures that the integrals in 3. are well defined.
So given an SDE questions are about existence of a solution, if it exists, then uniqueness of it; and not unimportant, the properties of the solutions.
In an SDE: $\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t$, μ is called drift/instantaneous growth term and σ^2 is called the diffusion coefficient/instantaneous variance.

Example 19 (7.3)**.** Consider the SDE $\mathrm{d}X_t = \mu X_t\mathrm{d}t + \sigma X_t\mathrm{d}B_t$ with $X_0 = x_0 \in \mathbb{R}$.
Let's see if $X_t = f(t, B_t)$ can be a solution to such SDE.
Applying Itô formula to $f(t, B_t)$ we have,

$$\mathrm{d}[f(t, B_t)] = \left[f_t(t, B_t) + \tfrac{1}{2}f_{xx}(t, B_t)\right]\mathrm{d}t + f_x(t, B_t)\mathrm{d}B_t$$

so if there exists f such that

$$f_t + \tfrac{1}{2}f_{xx} = \mu \cdot f \text{ and } f_x = \sigma f$$

then $X_t = f_t(t, B_t)$ will be a solution.
$f_x = \sigma f \Rightarrow f(t, x) = g(t)e^{\sigma x}$ where g is some function of t only. Plugging this into the 1st expression yields: $\frac{g'(t)}{g(t)}f + \frac{\sigma^2}{2}f = \mu f$. So if there exists a $g(t)$ such that $\frac{g'}{g} = \frac{1}{2}\sigma^2 - \mu$ then it will do.

But $\frac{g'}{g} = \mu - \frac{1}{2}\sigma^2 \Rightarrow g = ce^{(\mu-\frac{1}{2}\sigma^2)t}$ where c is the integration constant. So $f(t,x) = ce^{(\mu-\frac{1}{2}\sigma^2)t+\sigma x}$. Now consider $X_t = f(t, B_t) = ce^{(\mu-\frac{1}{2}\sigma^2)t+\sigma B_t}$. It is not difficult (using Itô) that all conditions in the definition of a solution are satisfied.
To make sure that initial condition is satisfied one needs $c = x_0$. hence the complete solution is $X_t = x_0 e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma B_t}$.
If X_0 was a random variable η (which must be $\mathcal{F}_0$-measurable and hence independent of $(B_t)_{t>0}$) then $X_t = \eta e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma B_t}$
This is one solution, are there any other solutions? That would be answered with no via a general result
Properties
$\mathbb{E}[X_t] = \mathbb{E}[\eta] e^{\mu t}$ which grows exponentially assuming that $\mathbb{E}[\eta] \neq 0$, but $X_t = \eta e^{t((\mu-\frac{1}{2}\sigma^2)+\sigma\frac{B_t}{t}}$. The strong law of large numbers says that $\frac{B_t}{t} \to 0$ a.s. thus if $(\mu - \frac{1}{2}\sigma^2) < 0$ then $X_t \to 0$ a.s. as $t \to \infty$.

Here is another example of a sequence of random variables which converges to 0 a.s. but is expectations converge to ∞.

Example 20 (7.2 (Ornstein Uhlenbeck process)**.**

$$\mathrm{d}X_t = -\alpha X_t \mathrm{d}t + \sigma \mathrm{d}B_t \qquad X_0 = x_0$$

Show that a solution of the form $X_t = f(t, B_t)$ does not exist.
So we need to use a different technique. Multiply both sides by the integrating factor $Z_t = e^{\alpha t}$. Then apply Itô formula to $(ZX)_t$ to obtain the solution:

$$X_t = x_0 e^{-\alpha t} + \int_0^t \sigma e^{-\alpha(t-s)} \mathrm{d}B_s$$

11 Applications of Itô's formula

Brownian Bridge(Example 7.4) For fixed $0 < t < 1$:

$$\mathrm{d}X_t = -\frac{X_t}{1-t}\mathrm{d}t + \mathrm{d}B_t \text{ with } X_0 = x_0$$

has the solution $X_t = x_0 + e^{-\alpha t} + \sigma(1-t)\int_0^t \frac{1}{1-s}\mathrm{d}B_s$. X_t is defined on $[0,1)$ and $X_t \to 0$ as $t \uparrow 1$. X_t is a Brownian motion conditioned at the end $(t = 1)$ to be also zero.
$X_t = B_t - tB_1$ is also a Brownian bridge

Theorem 67 (Theorem 7.8)**.** *Consider the SDE on the given space* $(\Omega, \mathcal{F}, P)$*:*

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, t \in [0,T]; X_0 = \xi \in \mathcal{F}_0$$

Suppose the coefficients b *and* σ *satisfy the* ***Lipschitz condition:***

$$|b(t,x) - b(t,y)|^2 + |\sigma(t,x) - \sigma(t,y)|^2 \le L|x-y|^2$$

for some constant $L > 0$ *and the spatial* ***Growth condition***

$$|b(t,x)|^2 + |\sigma(t,x)|^2 \le L(1 + |x|^2)$$

Then there exists a continuous, adapted process X *which is a solution of the SDE. Furthermore, the process* X *is* ***unique*** *up to indistinguishability, i.e. if* X_t *and* Y_t *are both solutions of the SDE then* $P(X_t = Y_t \text{ for all} t \in [0,T]) = 1$

Some useful results are listed below:

Theorem 68 (Gronwall's Lemma (Lemma A.20))**.** *Let* g *be an integrable Borel function on* $[a,b]$ *and* f *a non-decreasing function on* $[a,b]$*. Suppose there is a constant* c *such that*

$$g(t) \le f(t) + c\int_a^t g(s)ds \qquad \forall t \in [a,b]$$

Then $g(t) \le f(t)e^{c(t-a)}$

Theorem 69 (Doob's maximum inequality)**.** *For square integrable continuous martingale* M*, and* $0 < T < \infty$

$$\mathbb{E}\left[\sup_{0\le t\le T}|M_t|^2\right] \le 4\mathbb{E}\left[|M_T|^2\right]$$

Theorem 70 (Theorem 7.12)**.** *Suppose* ξ, η *are* $\mathcal{F}_0$*-measurable random variables. Assume* b *and* σ *satisfy the Lipschitz condition. Suppose* X *and* Y *are solutions to the same SDE with coefficients* b *and* σ *but with possibly different initial values* ξ *and* η*, respectively. Then* X *and* Y *are indistinguishable, on the event* $\{\xi = \eta\}$*, i.e.,* $P((X_t - Y_t)\mathbf{1}_{\{\xi=\eta\}} = 0, \forall t \in [0,T]) = 1$

Now a very long proof of this theorem followed, which I think is not relevant.

Theorem 71 (Theorem 7.14). *Suppose b and σ are continuous functions of (t,x) satisfying the growth and Lipschitz conditions.*
Let X be the strong solution of the SDE with coefficients b and σ (and with $\mathcal{F}_0$-measurable ξ as initial value) on the filtered probability space $(\Omega,\mathcal{F},P)$ with B a Brownian motion on it.
Let $\tilde{X}$ be the strong solution corresponding to the SDE with same coefficients b and σ but corresponding to $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P}),\tilde{B},\tilde{\xi}$.
Suppose $\xi=\tilde{\xi}$ in distribution.
Then the processes X and $\tilde{X}$ have the same probability distribution. I.e., for any measurable set A of $C_{\mathbb{R}^d}[0,T], P(X\in A)=\tilde{P}(\tilde{X}\in A)$

In the absence of the growth and Lipschitz conditions one may not always be able to find a (strong) solution defined on the given probability space $(\Omega,\mathcal{F},P)$ It is however, sometimes possible to define/construct

1. Another (filtered) probability space $(\Omega^*,\mathcal{F}^*,P^*)$
2. An SBM B_t^* on the new filtered space
3. An $\mathcal{F}_0$-measurable ξ^* with probability distribution same as that of ξ
4. A continuous adapted process X_t^* w.r.t. the new filtered space such that

$$\int_0^T |b(t,X_t^*)|\mathrm{d}t+\int_0^T |\sigma(t,X_t^*)|^2\mathrm{d}t<\infty$$

and

$$X_t^*=\xi^*+\int_0^t b(s,X_s^*)\mathrm{d}+\int_0^t \sigma(s,X_s^*)\mathrm{d}B_s^*$$

Then $(\Omega^*,\mathcal{F}^*,P^*,\xi^*,(B_t^*),(X_t^*))$ is called the *weak solution* of the SDE

$$\mathrm{d}X_t=b(t,X_t)\mathrm{d}t+\sigma(t,X_t)\mathrm{d}B_t$$

12 Girsanov's theorem

The main question of this section is: *"Can a stochastic process with drift be viewed as one without drift? Or be transformed into one?"*

$$X_t = \int_0^t \sigma_s \mathrm{d}B_s \qquad Y_t = \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s$$

Because X_t is a martingale, it is easier to analyze then Y_t!

Monte Carlo Integration
The Riemann sum is given by $\int_0^1 f(x)\mathrm{d}x \approx \frac{1}{n}\sum_{i=1}^n f(x_i)$ for $x_i = \frac{i-1}{n}$.
Monte Carlo integration is the same concept but now random variables are used to approximate the integral: $\int_0^1 f(x)\mathrm{d}x \approx \frac{1}{n}\sum_{i=1}^n f(X_i)$ for $X_i \sim \mathrm{Unif}[0,1]$.
Now the Strong Law of Large Numbers yields that if X_i's are i.i.d. with finite expectation μ, then $\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{n\to\infty} \mathbb{E}[X]$ almost surely. Therefore we can approximate $\mathbb{E}[f(x)]$ by drawing large samples $X_1, \ldots, X_n$ from the distribution of X and considering the sum $\frac{1}{n}\sum_{i=1}^n f(X_i)$

$$\int f(x)p(x)\mathrm{d}x = \mathbb{E}[f(x)] \approx \frac{1}{n}\sum_{i=1}^n f(X_i), \qquad X_i \sim p(x)$$

In theory this is a very nice idea, but in practice it doesn't work for most cases. Let's see for example the case that we are interested in $P(X > 30)$ for $X \sim N(0,1)$. Then we can approximate this probability by $P(X > 30) = \mathbb{E}[f(x)]$ for $f(x) = \mathbf{1}_{(30,\infty)}$ so that we have:

$$P(X > 30) = \mathbb{E}[f(X)] \approx \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{(30,\infty)}, \qquad X_i \sim N(0,1)$$

If we define D to be the number of draws before the first hit ($x_i > 30$), then $\mathbb{E}[D] > 10^{100}$. So in practice this approximation is quite useless.

Importance Sampling
For this problem *importance sampling* has been invented. By importance sampling we convert the problem so we can sample from more easy distributions:

$$\int f(x)p(x)\mathrm{d}x = \int f(x)\frac{p(x)}{q(x)}q(x)\mathrm{d}x = \int g(x)q(x)\mathrm{d}x$$
$$\mathbb{E}^P[f(X)] = \mathbb{E}^Q[g(X)] = \mathbb{E}^Q\left[f(X)\frac{p(X)}{q(X)}\right]$$

In order to apply importance sampling fruitfully we need the ability to draw sample from density $q(x)$, the ability to calculate $\frac{p(x)}{q(x)}$ and $q(x) > 0$ whenever $p(x) > 0$ (or equivalently $q(x) = 0 \Longleftrightarrow p(x) = 0$)
If we get back to our previous example, for $p \sim N(0,1)$; $q \sim N(\mu,1)$ such that $p(x)/q(x) = e^{-\mu x + \frac{1}{2}\mu^2}$ and $P(X > 30) \approx \frac{1}{n}\sum_{i=1}^n \left[\mathbf{1}_{\{X_i > 30\}} e^{-\mu X_i + \frac{1}{2}\mu^2}\right]$, where $X_i \sim N(\mu, 1)$. So choosing a suitable value for μ improves the approximation.

Change of Measure
So if we have the same random variable, but we want a different probability distribution? In that we case we define them on different probability measures. Consider $\Omega = \mathbb{R}$ equipped with the Borel σ-algebra $\mathcal{B}$ and a random variable $X : \Omega \to \mathbb{R}$ given by $X(\omega) = \omega$.
Consider probability measures on $(\Omega, \mathcal{B})$, given by:

- $P_1((a, b]) = (b \wedge 1) \vee 0 - (a \wedge 1) \vee 0$
- $P_2((a, b]) = \Phi(b) - \Phi(a)$

Under $P_1, X \sim U(0, 1)$ and under $P_2, X \sim N(0, 1)$

Now consider a probability space $(\Omega, \mathcal{F}, P)$ and a random variable X defined on it such that $X \sim N(0, 1)$ under P. For some $\mu \in \mathbb{R}$, let $Z = e^{\mu X - \frac{1}{2}\mu^2}$, then $Z > 0$ and $\mathbb{E}[Z] = 1$. Define a new measure Q on $(\Omega, \mathcal{F}$ by $Q(A) = \mathbb{E}[\mathbf{1}_A Z]$ for $A \in \mathcal{F}$. Now Q is a probability measure and under Q, $X \sim N(\mu, 1)$.

Theorem 72 (Girsanov Theorem)**.** *Suppose (B_t) is a d-dimensional Brownian Motion defined on the complete filtered probability space $(\Omega, \mathcal{F}, P), 0 < T < \infty$ is fixed and H is an adapted measurable $\mathbb{R}^d$-valued process such that $\int_0^T |H(t)|^2 dt < \infty$ almost surely under P.*
Let $Z_t = Z_t(H) = \exp\left\{\int_0^t H(s) dB(s) - \frac{1}{2}\int_0^t |H(s)|^2 ds\right\}$.

- *Assume that $\{Z_t, t \in [0, T]\}$ is martingale. (Equivalent assumption: $\mathbb{E}[Z_T] = \mathbb{E}^P[Z_t] = 1$.)*
- *Define the probability measure $Q = Q_T$ on $\mathcal{F}_T$ as $dQ = Z_T dP$*
- *Define the process $W(t) = B(t) - B(0) - \int_0^t H(s) ds$*

Then $\{W(t), t \in [0, T]\}$ is a d-dimensional Brownian Motion on the probability space $(\Omega, \mathcal{F}_T, Q)$ w.r.t. the filtration $\{\mathcal{F}_t\}_{t \in [0,T]}$

Remark 1: $M_t = \int_0^t H(s) \mathrm{d}B_s$ is a continuous local martingale. Then Itô formula says that $Z_t = 1 + \int_0^t Z_s \mathrm{d}M_s = 1 + \int_0^t Z_s H_s \mathrm{d}B_s$ such that Z_t is a continuous local martingale.
Remark 2: $Z_T \geq 0 \Rightarrow Q$ is a positive measure and Z_t is martingale $\Rightarrow \mathbb{E}^P[Z_T] = \mathbb{E}^P[Z_0] = 1$ hence Q is a probability measure.

A Useful Observation
For $t \in \mathbb{R}_+$, define Q_t on $(\Omega, \mathcal{F}_t)$ as $\mathrm{d}Q_t = Z_t \mathrm{d}P$. Suppose that Z_t is a martingale. Then the family of measure $\{Q_t\}$ satisfy certain consistency properties:
Let $s < t$ and $A \in \mathcal{F}_s \subset \mathcal{F}_t$

$$\begin{aligned} Q_t(A) &= \mathbb{E}^P[\mathbf{1}_A Z_t] = \mathbb{E}^P\left[\mathbb{E}^P[\nVdash_A Z_t | \mathcal{F}_s]\right] \\ &= \mathbb{E}^P\left[\mathbf{1}_A \mathbb{E}^P[Z_t | \mathcal{F}_s]\right] = \mathbb{E}^P[\mathbf{1}_A Z_s] \\ &= Q_s(A) \end{aligned}$$

Example 21 (Application 1)**.** Let B_t be a standard Brownian Motion; $\alpha < 0, \mu \in \mathbb{R}$ and σ : first time B_t hits the (space-time) line $a - \mu t$. What is the probability distribution of σ?

- Define $X_t = B_t + \mu t$. Then $\sigma = \inf\{t \geq 0 : B_t = a - \mu t\} = \inf\{t \geq 0; X_t = a\}$

- Use Girsanov's theorem with $H(s) = -\mu$ such that $Z_t = e^{-\mu B_t - \mu^2 t/2}$ and note that Z_t is indeed a martingale. Now $Q_t(A) = \mathbb{E}^P\left[\mathbf{1}_A Z_t\right]$. such that $\{X_s, 0 \leq s \leq t\}$ is a standard Brownian Motion under Q_t.

- Since $Z_t > 0$, it holds that $P(A) = \mathbb{E}^Q\left[\mathbf{1}_A Z_t^{-1}\right]$, for $A \in \mathcal{F}$ $[\mathrm{d}Q_t = Z_t \mathrm{d}P \Leftrightarrow \mathrm{d}P = Z_t^{-1}\mathrm{d}Q_t]$
$Z_t^{-1} = e^{\mu B_t + \mu^2 t/2} = e^{\mu X_t - \mu^2 t/2}$

- $$\begin{aligned}
P(\sigma > t) = P\left(\inf_{0 \leq s \leq t} X_s > a\right) &= \mathbb{E}^Q\left[\mathbf{1}_{\left\{\inf_{0 \leq s \leq t} X_s > a\right\}} Z_t^{-1}\right] \\
&= \mathbb{E}^Q\left[\mathbf{1}_{\left\{\inf_{0 \leq s \leq t} X_s > a\right\}} e^{\mu X_t - \mu^2 t/2}\right] \\
&= e^{-\mu^2 t/2}\mathbb{E}^Q\left[\mathbf{1}_{\left\{\sup_{0 \leq s \leq t}(-X_s) < -a\right\}} e^{-\mu(-X_t)}\right] \\
&= e^{-\mu^2 t/2}\mathbb{E}^P\left[\mathbf{1}_{\left\{\sup_{0 \leq s \leq t} M_t < -a\right\}} e^{-\mu B_t}\right] \text{ where } M_t = \sup_{0 \leq s \leq t} B_t
\end{aligned}$$

- The joint distribution of (B_t, M_t) is known.

Theorem 73 (Theorem 8.13)**.** *Suppose H is adapted, measurable with $\int_0^T |H(t)|^2 dt < \infty$ almost surely under P. The process $Z_t - \exp\left\{\int_0^t H(s) dB(s) - \frac{1}{2}\int_0^t |H(s)|^2 ds\right\}$ (which is a positive local martingale, and hence a supermartingale) is martingale under any of the following conditions:*

- *$H(t)$ is non-random*
- *$((H(t))$ and (B_t) are mutually independent processes.*
- *$\{H(t), t \in [0, T]\}$ is bounded*
- *$\int_0^T |H(s)|^2 ds \leq C < \infty$ almost surely*
- *Novikov conditon:* $\mathbb{E}\left[e^{\frac{1}{2}\int_0^T |H(s)|^2 ds}\right] < \infty$

Theorem 74 (Theorem 8.17)**.** *Let $0 < T < \infty$, $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ Borel measurable and B_t a d-dimensional standard Brownian Motion. Consider the SDE:*

$$dX_t = b(t, X_t)\, dt + dB_t \qquad \text{with } X_0 \sim \nu$$

If b is bounded, then the SDE has a weak solution for any initial distribution ν on $\mathbb{R}^d$