MS-E1601 BROWNIAN MOTION AND STOCHASTIC ANALYSIS

Brownian motion is one of the most important stochastic processes. Historically, it emerged in various contexts:

- Irregular motion of pollen particles in a fluid
  - botanist Robert Brown, 1827
  - physicist Albert Einstein, 1905
- Financial markets
  - mathematician Louis Bachelier, 1900
  - economists Robert C. Merton, and Paul A. Samuelson, also Fischer Black and Myron Scholes, 1973
- Quantum physics
  - physicist Richard Feynman, 1942
  - mathematician Mark Kac, 1947
- etc. etc.

Brownian motion is a continuous time stochastic process, with time t usually in the non-negative real axis

$$\mathbb{R}_+ = [0, +\infty).$$

The d-dimensional Brownian motion has state space

$$\mathbb{R}^d = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{d \text{ copies}}$$

(the d-dimensional Euclidian vector space).

We will mostly focus on $d=1$, since in general, the d components of a d-dimensional Brownian motion are simply independent one-dimensional Brownian motions.

The one-dimensional Brownian motion is therefore a collection

$$B = (B_t)_{t \in \mathbb{R}_+}$$

of $\mathbb{R}$-valued random variables B_t, which represent the values (positions) of the process at times $t \in \mathbb{R}_+ = [0, \infty)$.

Warning: We often — but not always — write the time t as a subscript index as in B_t above. Occasionally, however, it is preferable to write it as an argument, $B(t)$. Try to choose convenient practices for yourself, and do not get confused by the varying conventions.

The random variables B_t are defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$:

$$B_t : \Omega \longrightarrow \mathbb{R} \qquad \left(\text{for } t \in \mathbb{R}_+ = [0,\infty)\right)$$

The value of the process thus depend on the outcome $\omega \in \Omega$ of randomness, and we should in principle write

$$B_t(\omega) \in \mathbb{R} \qquad (t \in \mathbb{R}_+, \ \omega \in \Omega).$$

However, most of the time we do not write ω explicitly (as usually in probability theory).

Here are a few ways of looking at the Brownian motion B:

- A collection $B = (B_t)_{t \in \mathbb{R}_+}$ variables $B_t : \Omega \to \mathbb{R}$, for $t \in \mathbb{R}_+$.
- A function
$$B : \Omega \times \mathbb{R}_+ \longrightarrow \mathbb{R}$$
$$(\omega, t) \longmapsto B_t(\omega).$$
- A random function, i.e., a function-valued random variable
$$B : \Omega \longrightarrow \{ \text{functions } [0,\infty) \to \mathbb{R} \}$$
$$\omega \longmapsto \underbrace{(t \mapsto B_t(\omega))}_{\substack{\text{function of time } t, \\ \text{the path of the process} \\ \text{on the outcome } \omega \in \Omega.}}$$

The last perspective is very important. We will care a lot about the properties of the (random) paths

$$t \longmapsto B_t.$$

(continuity, non-differentiability, variation, ...).

Before the definition, let us recall properties of Gaussian random variables that will be used repeatedly.

Def: A real-valued random variable X is said to have Gaussian distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$, denoted $X \sim N(\mu, \sigma^2)$, if X has probability density

$$f_X(x) = p_{\sigma^2}(\mu, x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$

i.e. for all Borel sets $A \subset \mathbb{R}$

$$\mathbb{P}[X \in A] = \int_A p_{\sigma^2}(\mu, x)\, dx .$$

Exercise Show that if $X \sim N(\mu, \sigma^2)$, then the characteristic function of X is

$$\varphi_X(\theta) := \mathbb{E}[e^{i\theta X}] = e^{i\mu\theta - \frac{1}{2}\sigma^2\theta^2}.$$

Exercise Show that if $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$ are independent, then $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. Conclude that

$$\int_{\mathbb{R}} p_t(x,y)\, p_s(y,z)\, dy = p_{t+s}(x,z) \quad \left(\text{for } \begin{matrix} t, s > 0 \\ x, z \in \mathbb{R} \end{matrix}\right).$$

Exercise: Show that if $X \sim N(0, \sigma^2)$, then for any even positive integer $p \in \{2, 4, 6, \ldots\}$ we have

$$\mathbb{E}[X^p] = c_p \cdot (\sigma^2)^{p/2}$$

where $c_p = (p-1)(p-3)(p-5) \cdots 3 \cdot 1$.

Exercise: Show that if $X \sim N(0,1)$, then for any $x > 0$ we have

$$\frac{1}{\sqrt{2\pi}}(x^{-1} - x^{-3})\, e^{-\frac{1}{2}x^2} \leq \mathbb{P}[X > x] \leq \frac{1}{\sqrt{2\pi}}\, x^{-1}\, e^{-\frac{1}{2}x^2}.$$

Exercise For $X \sim N(\mu, \sigma^2)$ and $\lambda > 0$ show that $\lambda X \sim N(\lambda\mu, \lambda^2\sigma^2)$.

Definition A real-valued stochastic process $B = (B(t))_{t \in \mathbb{R}_+}$ is a *standard pre-Brownian motion* if it satisfies

(BM-$\mathbb{1}$) for any $0 = t_0 < t_1 < t_2 < \dots < t_n$, the increments

$$B(t_1) - B(t_0), \; B(t_2) - B(t_1), \dots, \; B(t_n - t_{n-1})$$

are independent

(BM-N) for any $0 \le s < t$ the increment is normal distributed

$$B(t) - B(s) \sim N(0, t-s).$$

(mean 0, variance $t-s$)

(BM-0) $B(0) = 0$ almost surely.

Remarks: 1) The sigma-algebra we use is that generated by the events of the form

$$\{\omega \in \Omega \mid B_{t_0}(\omega) \in A_0, \; B_{t_1}(\omega) \in A_1, \dots, B_{t_n}(\omega) \in A_n\}$$

$$= \bigcap_{j=0}^{n} \{\omega \in \Omega \mid B_{t_j}(\omega) \in A_j\}$$

where $A_0, A_1, \dots, A_n \subset \mathbb{R}$ are Borel subsets. These events are called *finite-dimensional events* (or *cylinder events*). They form a π-system (stable under finite intersections), so a probability measure on paths $[0, \infty) \to \mathbb{R}$ with this sigma-algebra is uniquely characterized by the probabilities of the finite-dimensional events, according to Dynkin's identification theorem.

For a standard (pre-) Brownian motion, we have

$$\mathbb{P}\left[B_{t_1} \in A_1 , B_{t_2} \in A_2 , \dots , B_{t_n} \in A_n\right]$$
$$= \int_{A_1} dx_1 \; p_{t_1 - t_0}(0, x_1) \int_{A_2} dx_2 \; p_{t_2 - t_1}(x_1, x_2) \cdots \int_{A_n} dx_n \; p_{t_n - t_{n-1}}(x_{n-1}, x_n)$$

by properties (BM-$\perp\!\!\!\perp$) and (BM-N).

2) The term "standard" refers to the choices $B(0) = 0$ instead of more general $B(0) = x_0 \in \mathbb{R}$ or even a random starting point and $\mathrm{Var}(B(t) - B(s)) = t - s$ instead of more general $\mathrm{Var}(B(t) - B(s)) = A \cdot (t-s)$ for some $A > 0$.

3) The distinction of a "pre-Brownian motion" and "Brownian motion" is that for the latter we moreover require the paths to be continuous, almost surely. We next note that a pre-Brownian motion can be modified slightly to become continuous.

<u>Def</u>: Two stochastic processes $X = (X_t)_t$ and $\tilde{X} = (\tilde{X}_t)_t$ on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are said to be versions (modifications) of each other, if for any t we have $\mathbb{P}\left[\{\omega \in \Omega \mid X_t(\omega) = \tilde{X}_t(\omega)\}\right] = 1$.

<u>Def</u> Two stoch. proc. $X = (X_t)_t$ and $\tilde{X} = (\tilde{X}_t)_t$ on the same proba space $(\Omega, \mathcal{F}, \mathbb{P})$ are said to be indistinguishable from each other if

$$\mathbb{P}\left[\{\omega \in \Omega \mid X_t(\omega) = \tilde{X}_t(\omega) \text{ for all } t\}\right] = 1.$$

Remark Carefully note the difference! There are uncountably many time instants t, so the former does not imply the latter!

Theorem 1.5 If B is a standard pre-Brownian motion, then there exists a version $\tilde{B}$ of B such that for any $\alpha < \frac{1}{2}$, the paths $t \mapsto \tilde{B}_t$ of $\tilde{B}$ are Hölder continuous of order α, i.e. for any $T > 0$ we have

$$\sup_{0 \le s < t \le T} \frac{|B(t) - B(s)|}{|t-s|^\alpha} < +\infty .$$

This is a consequence of the following Kolmogorov's continuity criterion.

Theorem 1.4 (Kolmogorov's continuity criterion)

Let $X = (X_t)_{t \in [0,1]}$ be a stochastic process. Suppose that there exists $p > 0$, $c > 0$, $\beta > 1$ such that for all $s,t \in [0,1]$

$$\mathbb{E}\left[|X_t - X_s|^p \right] \le c \cdot |t-s|^\beta .$$

Then there exists a version $\tilde{X}$ of X which is Hölder continuous of order α for any $\alpha < \frac{\beta - 1}{p}$.

Proof of Theorem 1.5: By scaling, it suffices to consider $T=1$.

Recall that by (BM-N) we have $B(t) - B(s) \sim N(0, t-s)$.

Then for any even integer $p \in \{2,4,6,\dots\}$ we can calculate

$$\mathbb{E}\left[(B(t) - B(s))^p \right] = c_p \cdot |t-s|^{p/2} \quad \text{where } c_p = (p-1)(p-3)\cdots 3 \cdot 1 .$$

(for example integration by parts, or differentiating characteristic fn)

Thm 1.4 applies with p and $\beta = \frac{p}{2}$, so Hölder continuity of order $\alpha < \frac{\beta-1}{p} = \frac{1}{2} - \frac{1}{p}$ follows. Finally take $p \to \infty$. $\square$

Proof of Theorem 1.4:

Let $D_n = \{ k \cdot 2^{-n} \mid k \in \{0,1,2,\dots,2^n\} \}$ be the set of dyadic numbers at level n.

Let $\alpha < \frac{\beta - 1}{p}$. Markov's inequality gives

$$\mathbb{P}\left[|X(k2^{-n}) - X((k-1)2^{-n})| > 2^{-n\alpha} \right]$$
$$\leq \frac{\mathbb{E}\left[|X(k2^{-n}) - X((k-1)2^{-n})|^p \right]}{(2^{-n\alpha})^p} \leq \frac{c \cdot (2^{-n})^\beta}{(2^{-n\alpha})^p} = c\, 2^{n(\alpha p - \beta)}.$$

Then by union bound

$$\mathbb{P}\left[\max_{k=1,2,\dots,2^n} |X(k2^{-n}) - X((k-1)2^{-n})| > 2^{-n\alpha} \right]$$
$$\leq \sum_{k=1}^{2^n} \mathbb{P}\left[|X(k2^{-n}) - X((k-1)2^{-n})| > 2^{-n\alpha} \right] \leq 2^n \cdot c\, 2^{n(\alpha p - \beta)}$$

Note that $\alpha p - \beta < -1$ so we have

$$\sum_{n=0}^{\infty} \mathbb{P}\left[\max_{k=1,\dots,2^n} |X(k2^{-n}) - X((k-1)2^{-n})| > 2^{-n\alpha} \right] \leq \sum_{n=0}^{\infty} c\, 2^{n(1+\alpha p - \beta)} < +\infty.$$

The convergence part of Borel-Cantelli lemmas thus says that we have almost surely $\max_{k=1,\dots,2^n} |X(k2^{-n}) - X((k-1)2^{-n})| \leq 2^{-n\alpha}$ except for finitely many values of n. This implies that

$$\sup_{n \in \mathbb{Z}_{\geq 0}} \max_{k=1,\dots,2^n} \frac{|X(k2^{-n}) - X((k-1)2^{-n})|}{2^{-n\alpha}} \leq \underbrace{M(\omega)}_{\text{random but finite number}} < \infty.$$

We now claim that X is then Hölder continuous of order α on the dyadic numbers $D = \bigcup_{n \in \mathbb{N}} D_n$.

Suppose that $s,t \in D = \bigcup_{n\in\mathbb{N}} D_n$, $s<t$.
Let $r \in \mathbb{N}$ be the smallest positive integer such that $t-s > 2^{-r}$. Then $2^{-r} < t-s \le 2^{1-r}$ and there exists $u = k\cdot 2^{-r} \in D_r$ and coefficients $b_1, b_2, \dots, b_m, b_1', b_2', \dots, b_m' \in \{0,1\}$ such that

$$s = k\cdot 2^{-r} - b_1'\cdot 2^{-r-1} - b_2'\cdot 2^{-r-2} - \dots - b_m'\cdot 2^{-r-m}$$
$$t = k\cdot 2^{-r} + b_1\cdot 2^{-r-1} + b_2\cdot 2^{-r-2} + \dots + b_\ell\cdot 2^{-r-m}.$$

For $\ell = 0,1,\dots,m$, let also

$$s_\ell = k2^{-r} - \sum_{j=1}^{\ell} b_j'\cdot 2^{-r-j}, \qquad t_\ell = k\cdot 2^{-r} + \sum_{j=1}^{\ell} b_j\cdot 2^{-r-j}.$$

By triangle inequality, we get

$$\begin{aligned}
|X_t - X_s| &= |X_{t_m} - X_{s_m}| \\
&\le \big| X_{t_0} + (X_{t_1} - X_{t_0}) + \dots + (X_{t_m} - X_{t_{m-1}}) \\
&\qquad - X_{s_0} - (X_{s_1} - X_{s_0}) - \dots - (X_{s_m} - X_{s_{m-1}})\big| \\
&\le \underbrace{|X_{t_0} - X_{s_0}|}_{=0} + \sum_{j=1}^{m} \underbrace{|X_{t_j} - X_{t_{j-1}}|}_{\le M\cdot(2^{-r-j})^\alpha} + \sum_{j=1}^{m} \underbrace{|X_{s_j} - X_{s_{j-1}}|}_{\le M\cdot(2^{-r-j})^\alpha} \\
&\le 2\cdot M\cdot \sum_{j=1}^{\infty} 2^{-(r+j)\alpha} = 2\cdot M\cdot 2^{-(r+1)\alpha}\cdot\frac{1}{1-2^{-\alpha}} \\
&= \frac{2M}{2^\alpha - 1}\, 2^{-r\alpha} \le \frac{2M}{2^\alpha-1}(t-s)^\alpha.
\end{aligned}$$

This shows that there exists an almost surely finite M' such that

$$\frac{|X_t - X_s|}{|t-s|^\alpha} \le M' \quad \text{for all } t,s \in D = \bigcup_{n\in\mathbb{N}} D_n,$$

i.e. that X is almost surely Hölder continuous of order α on dyadic numbers.

Since dyadic numbers are dense, $\overline{D} = [0,1]$, there is then a Hölder continuous extension $\tilde{X} : [0,1] \longrightarrow \mathbb{R}$ defined by

$$\tilde{X}_t = \lim_{n \to \infty} X_{t_n}$$

where $t_1, t_2, \ldots \in D$ is any sequence s.t. $t_n \to t$.

(On the exceptional event when X is not Hölder continuous, we can set $\tilde{X}_t = 0$ for all t, this anyway happens with probability zero.)

It remains to show that $\tilde{X}$ is a version of X.

By Fatou's lemma we estimate, for any $t \in [0,1]$,

$$\mathbb{E}\left[|X_t - \tilde{X}_t|^p \right] = \mathbb{E}\left[\lim_{n\to\infty} |X_t - X_{t_n}|^p \right]$$

$$\leq \liminf_{n\to\infty} \underbrace{\mathbb{E}\left[|X_t - X_{t_n}|^p \right]}_{\leq c \cdot |t - t_n|^\beta \xrightarrow[n\to\infty]{} 0} = 0 .$$

This shows that $|X_t - \tilde{X}_t|^p = 0$ almost surely, which implies $\mathbb{P}[X_t = \tilde{X}_t] = 1$ so that $\tilde{X}$ is indeed a version of X. $\square$

Definition: A stochastic process $B = (B_t)_{t \in \mathbb{R}_+}$ is a *standard Brownian motion* if properties (BM-$\perp\!\!\!\perp$), (BM-N), (BM-0) hold $\leftarrow$ (i.e., B is a standard pre-Brownian motion) and in addition

(BM-c): $t \mapsto B_t$ is continuous, almost surely.

Remark: We thus in particular require that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is such that

$$\{\omega \in \Omega \mid t \mapsto B_t(\omega) \text{ is continuous}\}$$

is an event in $\mathcal{F}$.

A canonical choice of probability space is the space of continuous functions

$$\mathcal{C}(\mathbb{R}_+, \mathbb{R}) := \{\omega : \mathbb{R}_+ \to \mathbb{R} \text{ continuous function}\},$$

so that an outcome ω is the path of the process. This space is equipped with the σ-algebra $\mathcal{S}$ generated by the π-system of finite-dimensional events, i.e., the smallest σ-algebra such that the evaluation at time t

$$\begin{aligned} \mathrm{ev}_t : \mathcal{C}(\mathbb{R}_+, \mathbb{R}) &\longrightarrow \mathbb{R} \\ \omega &\longmapsto \omega(t) \end{aligned}$$

is measurable for every $t \in \mathbb{R}_+$.

The choice is very natural in view of the topology of uniform convergence on compact time intervals in the space of continuous functions.

Proposition: On $\mathcal{C}([0,T],\mathbb{R}) = \{\omega : [0,T] \to \mathbb{R} \text{ continuous}\}$ equipped with the sup-norm

$$\|\omega\|_\infty = \sup_{t\in[0,T]} |\omega(t)| ,$$

the Borel σ-algebra $\mathcal{B}(\mathcal{C}([0,T],\mathbb{R}), \|\cdot\|_\infty)$ generated by open subsets of the space of functions coincides with the σ-algebra generated by the finite-dimensional events.

Proof See e.g. "Large Random Systems". $\square$

Finally, let us reassure the concerned participants of the course that the topic of the course in fact exists.

Theorem 1.2: There exists a standard Brownian motion on some probability space.

Idea of proof The details are given in the textbook and in Berestycki's lecture notes — we only outline the idea, which is similar to Thm 1.4.

1°) Construct BM on dyadic times $D = \bigcup_{n\in\mathbb{N}} D_n \subset [0,1]$.

(This step uses a clever but easy conditional independence property of refinements of the Gaussian increments from D_n to D_{n+1}.)

2°) Check a.s. uniform continuity on $D \subset [0,1]$ and extend continuously to $[0,1]$ by density. (Then check that increments still have independent Gaussian laws, by dominated convergence.)

3°) Take i.i.d. copies of the process on $[0,1]$ to be used on $[n,n+1]$ by concatenation ($B^{(n)}$, $n \in \mathbb{Z}_{\geq 0}$)

$$B_t = \sum_{n=0}^{\lfloor t \rfloor} B_1^{(n)} + B_{t-\lfloor t \rfloor}^{(\lfloor t \rfloor)} \qquad (t \in \mathbb{R}_+).$$

(This still satisfies independent Gaussian increments property.)

The constructed process is continuous as the concatenation of continuous pieces. □

As a consequence of Theorem 1.5. we get:

Corollary A standard Brownian motion is almost surely Hölder continuous of any order $\alpha < \frac{1}{2}$.

This is essentially sharp:

Theorem 1.6 For a standard Brownian motion B we have for any $\alpha > \frac{1}{2}$

$$\mathbb{P}\left[\forall t \in \mathbb{R}_+ : \limsup_{h \to 0^+} \frac{|B(t+h) - B(t)|}{h^\alpha} = +\infty \right] = 1.$$

In particular, B is almost surely not Hölder continuous of order $\alpha > \frac{1}{2}$.

Proof: See Berestycki's notes. □

GEOMETRY OF THE SPACE OF SQUARE INTEGRABLE RANDOM VARIABLES

Throughout : $(\Omega, \mathcal{F}, \mathbb{P})$ probability space.

$$m\mathcal{F} := \{ X : \Omega \to \mathbb{R} \quad \mathcal{F}\text{-measurable} \}$$

= the set of all $\mathbb{R}$-valued random variables

$$\mathcal{L}^1(\mathbb{P}) := \{ X \in m\mathcal{F} \mid \mathbb{E}[|X|] < \infty \}$$

= the set of all integrable random variables

$$\mathcal{L}^2(\mathbb{P}) := \{ X \in m\mathcal{F} \mid \mathbb{E}[X^2] < \infty \}$$

= the set of all square integrable random variables

<u>Notation</u> For $X, Y \in \mathcal{L}^2(\mathbb{P})$

$$\langle X, Y \rangle := \mathbb{E}[XY] \qquad \text{"inner product"}$$

$$\|X\| := \sqrt{\langle X, X \rangle} \qquad \text{"norm"}$$

$$X \perp Y \overset{\text{def.}}{\iff} \langle X, Y \rangle = 0 \qquad \text{"orthogonality"}$$

<u>Remark</u> We have $\|X\| = 0$ if and only if $\mathbb{E}[X^2] = 0$, which happens if and only if the non-neg. random variable X^2 is almost surely 0, in turn if and only if X is almost surely 0 : $\mathbb{P}[X = 0] = 1$.

<u>Cauchy-Schwarz inequality</u> $X, Y \in \mathcal{L}^2(\mathbb{P}) \Rightarrow X \cdot Y \in \mathcal{L}^1(\mathbb{P})$

$$\left| \mathbb{E}[XY] \right| \leq \sqrt{\mathbb{E}[X^2] \cdot \mathbb{E}[Y^2]}$$

i.e. $$\left| \langle X, Y \rangle \right| \leq \|X\| \cdot \|Y\| .$$

Properties:

Triangle inequality:

$$\|X+Y\| \leq \|X\| + \|Y\|$$

Pythagoras theorem:

if $X \perp Y$ then

$$\|X+Y\|^2 = \|X\|^2 + \|Y\|^2$$

Parallelogram law:

$$\|X+Y\|^2 + \|X-Y\|^2 = 2\cdot\|X\|^2 + 2\cdot\|Y\|^2$$

First moment bound: $\mathbb{E}[|X|] \leq \|X\|$

Sketches of proof:

$$\|X+Y\|^2 = \langle X+Y , X+Y\rangle$$
$$= \langle X,X\rangle + 2\underbrace{\langle X,Y\rangle}_{\leq \|X\|\cdot\|Y\|} + \langle Y,Y\rangle$$
$$\leq \|X\|^2 + 2\|X\|\cdot\|Y\| + \|Y\|^2$$
$$= (\|X\| + \|Y\|)^2 \rightsquigarrow \text{take square roots.} \quad \square$$

Same calculation but observe $\langle X,Y\rangle = 0$ by orthogonality. $\square$

$$\|X+Y\|^2 = \|X\|^2 + 2\langle X,Y\rangle + \|Y\|^2$$
$$\|X-Y\|^2 = \|X\|^2 - 2\langle X,Y\rangle + \|Y\|^2$$

add these up $\square$

By C-S ineq: $\mathbb{E}[|X|\cdot 1] \leq \sqrt{\mathbb{E}[X^2]\,\mathbb{E}[1^2]}$

Convergence in the space of square integrable r.v.'s

Def: Let $X_1, X_2, \ldots \in \mathcal{L}^2(\mathbb{P})$ and $X \in \mathcal{L}^2(\mathbb{P})$.
We say that X_n tends to X in $\mathcal{L}^2$ and denote $X_n \xrightarrow[n\to\infty]{\mathcal{L}^2} X$ if

$$\|X_n - X\| \xrightarrow[n\to\infty]{} 0 \quad \text{or equivalently}$$

$$\mathbb{E}\left[(X_n - X)^2\right] \xrightarrow[n\to\infty]{} 0.$$

(Almost) uniqueness of limits: If $X_n \xrightarrow{\mathcal{L}^2} X$ and $X_n \xrightarrow{\mathcal{L}^2} \tilde{X}$ then by triangle inequality

$$\|\tilde{X} - X\| = \|\tilde{X} - X_n + X_n - X\|$$
$$\leq \|\tilde{X} - X_n\| + \|X_n - X\| \longrightarrow 0 \quad \text{so } \tilde{X} = X \text{ a.s.}$$

Completeness of the space of square integrable r.v.'s

Theorem (Completeness of $\mathcal{L}^2(\mathbb{P})$)

Suppose that $X_1, X_2, \ldots \in \mathcal{L}^2(\mathbb{P})$ is Cauchy:

$$\lim_{m\to\infty} \sup_{n,n'\geq m} \| X_n - X_{n'} \| = 0.$$

Then there exists a square integrable $X \in \mathcal{L}^2(\mathbb{P})$ such that $X_n \xrightarrow{\mathcal{L}^2} X$

Proof: By Cauchy property we may choose $m_1 < m_2 < m_3 < \ldots$ such that

$$\| X_n - X_{n'} \| \leq 2^{-k} \qquad \text{whenever } n, n' \geq m_k.$$

Then also

$$\mathbb{E}\left[|X_n - X_{n'}| \right] \leq \| X_n - X_{n'} \| \leq 2^{-k} \qquad \forall n, n' \geq m_k.$$

and in particular $\mathbb{E}\left[|X_{m_{k+1}} - X_{m_k}| \right] \leq 2^{-k}$.

Therefore

$$\sum_{k=1}^{\infty} \mathbb{E}\left[|X_{m_{k+1}} - X_{m_k}| \right] < \infty.$$

This implies (recall from Probability Theory) that

$$\sum_{k=1}^{\infty} |X_{m_{k+1}} - X_{m_k}| < \infty \qquad \text{a.s.}$$

and thus we almost surely have the absolute convergence of

$$\sum_{k=1}^{\infty} \left(X_{m_{k+1}} - X_{m_k} \right)$$

Consider the sum of this series

$$S = \sum_{k=1}^{\infty} \left(X_{m_{k+1}} - X_{m_k} \right) = \lim_{\ell\to\infty} \sum_{k=1}^{\ell} \left(X_{m_{k+1}} - X_{m_k} \right)$$

$$= \lim_{\ell\to\infty} \left(X_{m_{\ell+1}} - \cancel{X_{m_\ell}} + \cancel{X_{m_\ell}} - \cdots + \cancel{X_{m_3}} - \cancel{X_{m_2}} + \cancel{X_{m_2}} - X_{m_1} \right)$$

We get that $\lim_{\ell \to \infty} X_{m_\ell} = S + X_{m_1} =: X$ a.s.

Then by Fatou's lemma, for $n \geq m_k$ we have

$$\mathbb{E}[(X_n - X)^2] = \mathbb{E}[\lim_{\ell \to \infty} (X_n - X_{m_\ell})^2]$$

$$\overset{\text{Fatou}}{\leq} \liminf_{\ell} \mathbb{E}[(X_n - X_{m_\ell})^2]$$

$$= \liminf_{\ell} \underbrace{\| X_n - X_{m_\ell} \|^2}_{\leq 4^{-k} \text{ when also } \ell \geq k, \text{ since } n \geq m_k.} \leq 4^{-k}.$$

This shows first of all that $X - X_n \in \mathcal{L}^2(\mathbb{P})$ and thus also $X = (X - X_n) + X_n \in \mathcal{L}^2(\mathbb{P})$. Moreover, it shows that

$$\lim_{n \to \infty} \mathbb{E}[(X_n - X)^2] = 0.$$

□

Orthogonal projections

Def: A vector subspace $\mathcal{V} \subset \mathcal{L}^2(\mathbb{P})$ is closed if for any sequence $X_1, X_2, \ldots \in \mathcal{V}$ in the subspace which converges $X_n \xrightarrow{\mathcal{L}^2} X$, the limit remains in the subspace: $X \in \mathcal{V}$.

Proposition (Orthogonal projection)

Let $\mathcal{V} \subset \mathcal{L}^2(\mathbb{P})$ be a closed subspace and let $X \in \mathcal{L}^2(\mathbb{P})$ be a square integrable random variable. Define

$$\Delta := \inf_{Z \in \mathcal{V}} \|X - Z\|.$$ (the distance of X to subspace $\mathcal{V}$)

Then for $Z \in \mathcal{V}$ the following conditions are equivalent:

(i): $\|X - Z\| = \Delta$ (Z minimizes the distance to X)

(ii) $X - Z \perp \mathcal{V}$. (the difference is orthogonal to the subspace)

Furthermore: there exists a random variable $Z \in \mathcal{V}$ with properties (i) and (ii) and if $\tilde{Z}$ is another such random variable, then $\tilde{Z} = Z$ a.s.

Proof:

(ii) $\Rightarrow$ (i): Assume $X - Z \perp V$ for all $V \in \mathcal{V}$.

Let $Z' \in \mathcal{V}$ be any other point in the subspace. Then $Z - Z' \in \mathcal{V}$ so $X - Z \perp Z - Z'$ so we may apply Pythagoras theorem and get

$$\|X - Z'\|^2 = \|(X - Z) + (Z - Z')\|^2$$
$$= \|X - Z\|^2 + \|Z - Z'\|^2 \geq \|X - Z\|^2$$

This shows that Z minimizes the distance to X.

(i) $\Rightarrow$ (ii): Assume Z is such that $\|X - Z\| = \Delta$. Let $V \in \mathcal{V}$. Consider, for $t \in \mathbb{R}$, the vector $Z + tV \in \mathcal{V}$. Then the distance is greater:

$$0 \leq \|X - Z - tV\|^2 - \Delta^2$$
$$= \langle X - Z - tV, X - Z - tV \rangle - \Delta^2$$
$$= \underbrace{\|X - Z\|^2}_{=\Delta^2} - 2t \cdot \langle X - Z, V \rangle + t^2 \cdot \|V\|^2 - \Delta^2$$
$$= -2t \cdot \langle X - Z, V \rangle + t^2 \cdot \|V\|^2.$$

If we would have $\langle X - Z, V \rangle \neq 0$ then the polynomial would have negative values for small positive or small negative t, which is a contradiction. Therefore $\langle X - Z, V \rangle = 0$. This proves $X - Z \perp \mathcal{V}$.

uniqueness: If Z and Z' both satisfy conditions (i) and (ii) then the Pythagoras calculation above gives

$$\Delta^2 = \|X - Z'\|^2 = \|X - Z\|^2 + \|Z - Z'\|^2 = \Delta^2 + \|Z - Z'\|^2,$$

which shows that $\|Z - Z'\| = 0$ and thus $Z = Z'$ a.s.

existence: By definition of Δ, we can find a sequence $Z_1, Z_2, \ldots \in \mathcal{V}$ such that

$$\|X - Z_n\|^2 \leq \Delta^2 + \frac{1}{n}.$$

Parallelogram rule gives, for $n, n' \in \mathbb{N}$,

$$2 \cdot \| X - Z_n \|^2 + 2 \cdot \| X - Z_{n'} \|^2$$
$$= \underbrace{\| 2X - Z_n - Z_{n'} \|^2}_{= 4 \cdot \| X - \frac{Z_n + Z_{n'}}{2} \|^2 \geq 4 \Delta^2} + \| Z_n - Z_{n'} \|^2 .$$

The choiche of Z_n then gives

$$\| Z_n - Z_{n'} \| \leq 2 \cdot \underbrace{\| X - Z_n \|^2}_{\leq \Delta^2 + \frac{1}{n}} + 2 \cdot \underbrace{\| X - Z_{n'} \|^2}_{\leq \Delta^2 + \frac{1}{n'}} - 4 \Delta^2$$
$$\leq \frac{2}{n} + \frac{2}{n'} .$$

This shows that $Z_1, Z_2, \ldots$ is Cauchy, so by completeness $Z_n \longrightarrow Z \in \mathcal{L}^2$. But since $Z_n \in \mathcal{V}$ and $\mathcal{V}$ is closed, also $Z \in \mathcal{V}$.

Finally, triangle inequality shows that

$$\| X - Z \| \leq \underbrace{\| X - Z_n \|}_{\to \Delta} + \underbrace{\| Z_n - Z \|}_{\to 0} \xrightarrow[n \to \infty]{} \Delta$$

i.e. $\| X - Z \| \leq \Delta$. This implies that Z satisfies (i). $\square$

Orthogonal projection as a best estimator

Suppose that $\mathcal{G} \subset \mathcal{F}$ is a sub-σ-algebra.

"partial information" "full information"

Then $\mathcal{V} = \mathcal{L}^2(\mathbb{P}) \cap m\mathcal{G}$

= the set of square integrable $\mathcal{G}$-measurable random variables

is a closed subspace of $\mathcal{L}^2(\mathbb{P})$.

(Why? Completeness of $\mathcal{L}^2(\mathbb{P})$ holds also in $(\Omega, \mathcal{G}, \mathbb{P})$)

Suppose that for $X \in \mathcal{L}^2(\mathbb{P})$ we want to give a "best estimate $\hat{X}$ using only information $\mathcal{G}$" in the sense of minimizing expected square error:

$$\mathbb{E}[(X-\hat{X})^2] \quad \text{minimal among} \quad \hat{X} \in \mathcal{L}^2(\mathbb{P}) \cap m\mathcal{G}.$$

Then $\hat{X}$ is just the orthogonal projection of X to the subspace $\mathcal{L}^2(\mathbb{P}) \cap m\mathcal{G} \subset \mathcal{L}^2(\mathbb{P})$.

<u>Lemma</u>: Let $X \in \mathcal{L}^2(\mathbb{P})$ and let $\hat{X}$ be the orthogonal projection of X to $\mathcal{L}^2(\mathbb{P}) \cap m\mathcal{G}$. Then for any $G \in \mathcal{G}$ we have

$$\mathbb{E}[\mathbb{1}_G \hat{X}] = \mathbb{E}[\mathbb{1}_G X].$$

<u>Proof</u>: Note that $\mathbb{1}_G \in \mathcal{L}^2(\mathbb{P}) \cap m\mathcal{G}$.

By property (ii), then, $X - \hat{X} \perp \mathbb{1}_G$, i.e.

$$0 = \langle X - \hat{X}, \mathbb{1}_G \rangle = \langle X, \mathbb{1}_G \rangle - \langle \hat{X}, \mathbb{1}_G \rangle$$
$$= \mathbb{E}[\mathbb{1}_G X] - \mathbb{E}[\mathbb{1}_G \hat{X}]. \quad \square$$

This property is taken in general as the definition of conditional expected value, given information $\mathcal{G}$.

<u>CONDITIONAL EXPECTED VALUE</u>

The notion of conditioning on information $\mathcal{G}$ is important when we ask:

e.g.
- Markov property
- martingale property

What can be said about the future of a stochastic process given the information about the past and present?

The general definition is the following.

For square integrable random variables existence comes from orthogonal projections.

Throughout: $(\Omega, \mathcal{F}, \mathbb{P})$ probability space.

<u>Definition</u>: Let $X \in \mathcal{L}^1(\mathbb{P})$ and let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra. Then a random variable $\hat{X} \in \mathcal{L}^1(\mathbb{P})$ is said to be (a version of) the conditional expected value $\mathbb{E}[X \mid \mathcal{G}]$ of X given $\mathcal{G}$ if $\hat{X} \in m\mathcal{G}$ (meaning $\hat{X}$ is $\mathcal{G}$-measurable)

for all $G \in \mathcal{G}$: $\quad \mathbb{E}[\mathbb{1}_G \hat{X}] = \mathbb{E}[\mathbb{1}_G X]$. $\circledast$

This leads to an "almost unique definition":

<u>Lemma</u> If $\hat{X}$ and $\hat{X}'$ are two versions of $\mathbb{E}[X \mid \mathcal{G}]$, then we have $\hat{X} = \hat{X}'$ almost surely.

<u>Proof</u>: Let $G_n = \{\omega \in \Omega \mid \hat{X}'(\omega) - \hat{X}(\omega) \geq \frac{1}{n}\}$. Then since $\hat{X}, \hat{X}' \in m\mathcal{G}$, we have $G_n \in \mathcal{G}$. Applying the defining equation $\circledast$, we get

$$\frac{1}{n}\mathbb{P}[G_n] \leq \mathbb{E}[\mathbb{1}_{G_n} \cdot (\hat{X}' - \hat{X})]$$

$$= \mathbb{E}[\mathbb{1}_{G_n} \hat{X}'] - \mathbb{E}[\mathbb{1}_{G_n} \hat{X}]$$

$$\overset{\circledast}{=} \mathbb{E}[\mathbb{1}_{G_n} X] - \mathbb{E}[\mathbb{1}_{G_n} X] = 0.$$

Therefore $\mathbb{P}[G_n] = 0$. By union bound

$$\mathbb{P}[X' > X] = \mathbb{P}\Big[\bigcup_{n \in \mathbb{N}} G_n\Big] \leq \sum_{n \in \mathbb{N}} \underbrace{\mathbb{P}[G_n]}_{=0} = 0,$$

so $X' \leq X$ almost surely. Similarly $X \leq X'$ a.s. $\square$

This shows that $\mathbb{E}[X \mid \mathcal{G}]$ is a well defined random variable except for modifications on an event of zero probability. We will not care about such events, so we treat $\mathbb{E}[X \mid \mathcal{G}]$ essentially as if it were unique. When we need to be careful, we speak of "a version of $\mathbb{E}[X \mid \mathcal{G}]$".

<u>Exercise</u> If $X \geq 0$ then $\mathbb{E}[X \mid \mathcal{G}] \geq 0$ (almost surely).

Construction for non-negative integrable random variables

Suppose $X \in \mathcal{L}^1(\mathbb{P}) \cap m\mathcal{F}^+$ is a non-neg. integrable random variable.

Consider the truncated random variable $X \wedge n$

$$(X \wedge n)(\omega) := \min\{X(\omega), n\}.$$

Then $X \wedge n$ is bounded and thus square integrable, $X \wedge n \in \mathcal{L}^2(\mathbb{P})$, so we can let

$$\hat{X}_n := \mathbb{E}[X \wedge n \mid \mathcal{G}] = \mathrm{proj}_{\mathcal{L}^2_{\mathcal{G}}(\mathbb{P})}(X \wedge n)$$

be its conditional expected value (projection).

Since $X \wedge n \uparrow X$ as $n \to \infty$, we have that $\hat{X}_n$ are increasing (a.s.): linearity implies

$$\hat{X}_{n+1} - \hat{X}_n = \mathbb{E}[\underbrace{X \wedge (n+1) - X \wedge n}_{\geq 0} \mid \mathcal{G}] \geq 0.$$

Therefore the limit $\lim_{n\to\infty} \hat{X}_n = \hat{X}$ exists.

For any $G \in \mathcal{G}$ by monotone convergence theorem

$$\mathbb{E}[\hat{X} \cdot \mathbb{1}_G] = \mathbb{E}[\lim_{n\to\infty} \hat{X}_n \cdot \mathbb{1}_G]$$

$$\overset{\text{MCT}}{=} \lim_{n\to\infty} \mathbb{E}[\hat{X}_n \cdot \mathbb{1}_G] = \lim_{n\to\infty} \mathbb{E}[(X \wedge n)\, \mathbb{1}_G]$$

$$\overset{\text{MCT}}{=} \mathbb{E}[\lim_{n\to\infty} (X \wedge n)\, \mathbb{1}_G] = \mathbb{E}[X \mathbb{1}_G]$$

This shows that $\hat{X} \in \mathcal{L}^1(\mathbb{P})$ (by taking $G = \Omega$) and that $\hat{X}$ is a version of $\mathbb{E}[X \mid \mathcal{G}]$.

Construction for integrable random variables

If $X \in \mathcal{L}^1(\mathbb{P})$, split to positive and negative parts: $X = X_+ - X_-$ with $X_+, X_- \in \mathcal{L}^1(\mathbb{P}) \cap m\mathcal{F}^+$.

Then $\hat{X} = \mathbb{E}[X_+ \mid \mathcal{G}] - \mathbb{E}[X_- \mid \mathcal{G}]$ satisfies the defining property $\circledast$.

Properties of conditional expected value

Theorem: Conditional expected values satisfy:

"known" (i): if $X \in m\mathcal{G}$ then $\mathbb{E}[X|\mathcal{G}] = X$ (a.s.)

"unbiased" (ii): $\mathbb{E}\big[\mathbb{E}[X|\mathcal{G}]\big] = \mathbb{E}[X]$

"linear" (iii): $X \mapsto \mathbb{E}[X|\mathcal{G}]$ is linear

"tower property" (iv): if $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ are nested sub-σ-algebras then $\mathbb{E}\big[\mathbb{E}[X|\mathcal{G}]\big|\mathcal{H}\big] = \mathbb{E}[X|\mathcal{H}]$ (a.s.)

"taking out what is known" (v): if $Z \in m\mathcal{G}$ and $ZX \in \mathcal{L}^1(\mathbb{P})$ then $\mathbb{E}[ZX|\mathcal{G}] = Z \cdot \mathbb{E}[X|\mathcal{G}]$ (a.s.)

"Jensen's inequality" (vi): If ϕ is a convex function and $\phi(X) \in \mathcal{L}^1(\mathbb{P})$ then $\phi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\phi(X)|\mathcal{G}]$

"monotone convergence theorem" (vii): if $0 \leq X_n \uparrow X \in \mathcal{L}^1(\mathbb{P})$ then $\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}]$ (a.s.)

"dominated convergence theorem" (viii): if $|X_n| \leq Z \in \mathcal{L}^1(\mathbb{P})$ and $X_n \to X$ then $\mathbb{E}[X_n|\mathcal{G}] \to \mathbb{E}[X|\mathcal{G}]$ (a.s.)

"Fatou's lemma" (ix): if $X_1, X_2, \dots \geq 0$ then $\mathbb{E}[\liminf_n X_n|\mathcal{G}] \leq \liminf_n \mathbb{E}[X_n|\mathcal{G}]$

"independent information" (x): if $X \perp\!\!\!\perp \mathcal{G}$ then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$

Sketch of proof

(i): direct from definition (*) and uniqueness

(ii): take $G = \Omega$ in (*)

(iii): the condition (*) is linear + uniqueness

(iv): for $H \in \mathcal{H} \subset \mathcal{G}$ by (*):
$$\mathbb{E}\big[\mathbb{1}_H\, \mathbb{E}[X|\mathcal{G}]\big] = \mathbb{E}[\mathbb{1}_H X]$$
and the LHS is $\mathcal{H}$-measurable + uniqueness.

(v): "standard machine": prove when
- $Z = \mathbb{1}_G$ is indicator
- $Z = \sum_{j=1}^n z_j \cdot \mathbb{1}_{G_j}$ is simple
- $Z \geq 0$ is non-negative
- Z is general $\mathcal{G}$-measurable

(vi): EXERCISE

(vii): this is the same calculation as in $\mathbb{E}[X \wedge n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}]$ above.

(viii) & (ix): Fatou's lemma and dominated convergence theorem were proved using monotone convergence theorem. The same proofs work here!

(x): "standard machine"

DISCRETE TIME MARTINGALES

Discrete time : $n \in \mathbb{Z}_{\geq 0}$

Information available at time $n \in \mathbb{Z}_{\geq 0}$ is a σ-algebra $\mathcal{F}_n \subset \mathcal{F}$, information accumulates: $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \dots \subset \mathcal{F}$.

Such an increasing collection $(\mathcal{F}_n)_{n \in \mathbb{Z}_{\geq 0}}$ of σ-algebras is called a filtration (some authors use the term history instead).

Example If $X = (X_n)_{n \in \mathbb{Z}_{\geq 0}}$ is a stochastic process, then $(\mathcal{F}^X_n)_{n \in \mathbb{Z}_{\geq 0}}$ defined by

$$\mathcal{F}^X_n = \sigma(X_0, X_1, \dots, X_n)$$

(the σ-algebra generated by the random variables $X_0, X_1, \dots, X_n$)

is a filtration. It is called the natural filtration of the process X, since $\mathcal{F}^X_n$ represents the information contained in the values of the process up to time n.

Def Let $X = (X_n)_{n \in \mathbb{Z}_{\geq 0}}$ be a real-valued stochastic process and $\mathcal{F}_\bullet = (\mathcal{F}_n)_{n \in \mathbb{Z}_{\geq 0}}$ a filtration. We say that the process X is adapted to the filtration $\mathcal{F}_\bullet$ if for all $n \in \mathbb{Z}_{\geq 0}$

$$X_n \in m\mathcal{F}_n .$$

(X_n is $\mathcal{F}_n$-measurable)

Interpretation: $X_n \in m\mathcal{F}_n$ says that the value of the process at time n can be inferred from the information available at that time!

Def Let $X = (X_n)_{n\in\mathbb{Z}_{\geq 0}}$ be a real-valued stochastic process and $\mathcal{F}_\bullet = (\mathcal{F}_n)_{n\in\mathbb{Z}_{\geq 0}}$ a filtration. We say that X is a *martingale* with respect to $\mathcal{F}_\bullet$ if it satisfies

(1°): X is adapted to $\mathcal{F}_\bullet$

(2°): $X_n \in \mathcal{L}^1(\mathbb{P})$ for all $n \in \mathbb{Z}_{\geq 0}$

(3°): $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n$ for all $n\in\mathbb{Z}_{\geq 0}$.

Similarly we say that X is a *submartingale* w.r.t. $\mathcal{F}_\bullet$ if it satisfies (1°), (2°), and

(3°_{sub}): $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \geq X_n \quad \forall n\in\mathbb{Z}_{\geq 0}$

and that X is a *supermartingale* w.r.t. $\mathcal{F}_\bullet$ if it satisfies (1°), (2°), and

(3°_{super}): $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \leq X_n \quad \forall n\in\mathbb{Z}_{\geq 0}$.

A martingale can be thought of as a "fair game", where a gamblers fortune after the next round is predicted as equal to her current fortune. A supermartingale, correspondingly, is a game where the gambler is at least predicted not to gain fortune (and a submartingale a game where the gambler is predicted not to lose — but don't expect casinos to offer you such options...).

When "betting" on a game, our stakes have to be decided before we see the results of the next round. This is captured by the following definition.

Def: A process $H = (H_n)_{n\in\mathbb{Z}_{\geq 0}}$ is *previsible* (predictable) if for all $n\in\mathbb{Z}_{\geq 0}$ we have $H_n \in m\mathcal{F}_{n-1}$.

(Usually interpret $\mathcal{F}_{-1}$ as the trivial σ-alg. $\mathcal{F}_{-1} = \{\emptyset, \Omega\}$)

MARTINGALE THEORY IN DISCRETE TIME

Filtration: $\mathcal{F}_\bullet = (\mathcal{F}_n)_{n \in \mathbb{Z}_{\geq 0}}$ collection of σ-algebras

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \ldots \subset \mathcal{F}$$

"information available at time n"

(idea: "information accumulates over time")

Def.: A stochastic process $X_\bullet = (X_n)_{n \in \mathbb{Z}_{\geq 0}}$ is adapted to $\mathcal{F}_\bullet = (\mathcal{F}_n)_{n \in \mathbb{Z}_n}$ if for each $n \in \mathbb{Z}_{\geq 0}$ the value X_n of the process at time n is an $\mathcal{F}_n$-measurable random variable, i.e., $X_n \in m\mathcal{F}_n$.

"the value of the process at any instant of time is known at that instant"

Def.: A stochastic process $H_\bullet = (H_n)_{n \in \mathbb{Z}_{\geq 0}}$ is predictable w.r.t. $\mathcal{F}_\bullet = (\mathcal{F}_n)_{n \in \mathbb{Z}_{\geq 0}}$ if for each $n \in \mathbb{Z}_{\geq 0}$ the value H_n of the process at time n is an $\mathcal{F}_{n-1}$-measurable random variable, i.e., $X_n \in m\mathcal{F}_{n-1}$.

"the value of the process at any instant of time is known before that instant"

(Idea: If H_n is the amount you bet on gambling round n, then you are not allowed to use the result of that round — but only prior rounds — to decide H_n.)

FOR DEFINITENESS:

Put: $\mathcal{F}_{-1} = \{\phi, \Omega\}$ trivial σ-algebra

i.e. H_0 is deterministic.

but never use the value H_0.

Def.: A stochastic process $X_\bullet = (X_n)_{n\in\mathbb{Z}_{\geq 0}}$ is a <u>martingale</u> w.r.t. $\mathcal{F}_\bullet = (\mathcal{F}_n)_{n\in\mathbb{Z}_{\geq 0}}$ if

(1°): $X_\bullet$ is adapted to $\mathcal{F}_\bullet$

(2°): $X_n \in \mathcal{L}^1(\mathbb{P}) \quad \forall\, n\in\mathbb{Z}_{\geq 0}$

(3°): $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n \quad \forall\, n\in\mathbb{Z}_{\geq 0}$.

"the best prediction for the next value given current information is the current value"

Def.: A stoch. proc. $X_\bullet$ is a <u>supermartingale</u> w.r.t. $\mathcal{F}_\bullet$ if it satisfies (1°), (2°), and

(3°_{super}): $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \leq X_n \quad \forall\, n\in\mathbb{Z}_{\geq 0}$.

A stoch. proc. $X_\bullet$ is a <u>submartingale</u> w.r.t. $\mathcal{F}_\bullet$ if it satisfies (1°), (2°), and

(3°_{sub}): $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \geq X_n \quad \forall\, n\in\mathbb{Z}_{\geq 0}$.

Remark: $X_\bullet$ is a supermgale if and only if its negative $-X_\bullet$ is a submgale. Also $X_\bullet$ is a mgale if and only if it is both supermgale and submgale. For these reasons we often do not give separate statements and proofs for each of the three, but leave it to the reader to translate results given for one case to the others. The following example illustrates this principle.

Lemma: If $X_\bullet$ is a submartingale, then for any $n\in\mathbb{Z}$ and $k\in\mathbb{Z}_{\geq 0}$ we have

$$\mathbb{E}[X_{n+k} \mid \mathcal{F}_n] \geq X_n.$$

Proof: Case $k=0$ is clear by property (i) of conditional expected values: since $X_n \in m\mathcal{F}_n$, we have $\mathbb{E}[X_n \mid \mathcal{F}_n] = X_n$ (a.s.).

Case $k=1$ is the defining property (3°_{sub}) of submartingales.

We then do induction on k: since $\mathcal{F}_n \subset \mathcal{F}_{n+k}$, the tower property (iv) of conditional expected values gives

$$\mathbb{E}\left[X_{n+k+1} \,\middle|\, \mathcal{F}_n\right] \overset{(iv)}{=} \mathbb{E}\Big[\underbrace{\mathbb{E}\left[X_{n+k+1} \,\middle|\, \mathcal{F}_{n+k}\right]}_{\geq X_{n+k} \text{ by } (3^\circ_{sub})} \,\Big|\, \mathcal{F}_n\Big]$$

$$\geq \mathbb{E}\left[X_{n+k} \,\middle|\, \mathcal{F}_n\right]$$

$$\geq X_n \qquad \text{by induction assumption.} \quad \square$$

(The translations to supermartingales and martingales are:

Lemma If $X_\bullet$ is a supermartingale, then
$$\mathbb{E}\left[X_{n+k} \,\middle|\, \mathcal{F}_n\right] \leq X_n \qquad \forall\, n \in \mathbb{Z}_{\geq 0},\ k \in \mathbb{Z}_{\geq 0}.$$

Lemma If $X_\bullet$ is a martingale, then
$$\mathbb{E}\left[X_{n+k} \,\middle|\, \mathcal{F}_n\right] = X_n \qquad \forall\, n \in \mathbb{Z}_{\geq 0},\ k \in \mathbb{Z}_{\geq 0}.$$
)

Proposition: If $X_\bullet = (X_n)_{n \in \mathbb{Z}_{\geq 0}}$ is a martingale and ϕ is a convex function and if $\mathbb{E}[|\phi(X_n)|] < +\infty$ for all $n \in \mathbb{Z}$, then the process $(\phi(X_n))_{n \in \mathbb{Z}_{\geq 0}}$ is a submartingale.

Proof: By conditional Jensen's inequality (vi), we get
$$\mathbb{E}\left[\phi(X_{n+1}) \,\middle|\, \mathcal{F}_n\right] \overset{(vi)}{\geq} \phi\Big(\underbrace{\mathbb{E}\left[X_{n+1} \,\middle|\, \mathcal{F}_n\right]}_{= X_n \text{ by } (3^\circ)}\Big) \overset{(3^\circ)}{=} \phi(X_n).$$

Adaptedness is clear and integrability follows from assumption. $\square$

Proposition If $X_\bullet = (X_n)_{n \in \mathbb{Z}_{\geq 0}}$ is a submartingale and ϕ is an increasing convex function and if $\mathbb{E}[|\phi(X_n)|] < +\infty$ for all $n \in \mathbb{Z}_{\geq 0}$, then the process $(\phi(X_n))_{n \in \mathbb{Z}_{\geq 0}}$ is a submartingale.

Proof: $$\mathbb{E}[\phi(X_{n+1}) \mid \mathcal{F}_n] \overset{(vi)}{\geq} \phi(\underbrace{\mathbb{E}[X_{n+1} \mid \mathcal{F}_n]}_{\geq X_n \text{ by } (3^\circ_{\text{sub}})}) \overset{\phi \text{ incr.}}{\geq} \phi(X_n). \qquad \square$$

Example $\phi(x) = (x-a)^+ = \max\{x-a, 0\}$ $(a \in \mathbb{R})$

$X_\bullet$ submartingale

Then using $|\phi(x)| = |(x-a)^+| \leq |x| + |a|$

we see $\mathbb{E}[|\phi(X_n)|] \leq \underbrace{\mathbb{E}[|X_n|]}_{<+\infty \text{ by } (1^\circ)} + |a| < +\infty$.

Thus we get that

$$\left((X_n - a)^+\right)_{n \in \mathbb{Z}_{\geq 0}} \quad \text{is a submartingale.}$$

Example Suppose that $X_\bullet$ is a supermartingale and $a \in \mathbb{R}$.

Notation: $x \wedge a = \min\{x, a\}$

Then $(X_n \wedge a)_{n \in \mathbb{Z}_{\geq 0}}$ is a supermartingale

(apply $x \mapsto \max\{x, a\} = \phi(x) + a$ to $-X_\bullet$).

(The processes in these examples are conveniently denoted
$(X_\bullet - a)^+ = \left((X_n - a)^+\right)_{n \in \mathbb{Z}_{\geq 0}}$ and $X_\bullet \wedge a = (X_n \wedge a)_{n \in \mathbb{Z}_{\geq 0}}$.)

Theorem: Suppose $X_\bullet = (X_n)_{n \in \mathbb{Z}_{\geq 0}}$ is a supermartingale and $H_\bullet = (H_n)_{n \in \mathbb{Z}_{\geq 0}}$ is a predictable process which is non-negative, $H_n \geq 0 \ \forall n$, and each value is bounded, $H_n \leq a_n < +\infty \ \forall n$ $(a_n \in \mathbb{R})$.

Then the process

$$H \cdot X = \left((H \cdot X)_n\right)_{n \in \mathbb{Z}_{\geq 0}}$$

defined by

$$(H \cdot X)_n = \sum_{k=1}^{n} H_k \cdot (X_k - X_{k-1})$$

is a supermartingale.

Proof: Exercise. $\square$

Theorem: Suppose $X_\bullet$ is martingale and $H_\bullet$ is predictable and H_n is bounded for each $n \in \mathbb{Z}_{\geq 0}$. Then $H \cdot X$ is a martingale.

Proof: Exercise. $\square$

Stock market interpretation

X_k = "stock price on day k"

H_k = "number of stocks in investors portfolio in the beginning of day k"

$X_k - X_{k-1}$ = "stock price increment during day k"

$H_k \cdot (X_k - X_{k-1})$ = "investors profit for day k"

$(H \cdot X)_n = \sum_{k=1}^{n} H_k (X_k - X_{k-1})$ = "investors cumulative profit from day 1 to day n"

Def: A random variable $\tau : \Omega \to \{0,1,2,\ldots\} \cup \{+\infty\}$ is a stopping time w.r.t. filtration $\mathcal{F}_\bullet = (\mathcal{F}_n)_{n \in \mathbb{Z}_{\geq 0}}$ if for every $k \in \mathbb{Z}_{\geq 0}$ we have

$$\{\omega \in \Omega \mid \tau(\omega) \leq k\} \in \mathcal{F}_k .$$

"Decision to stop at time k can be made with the information available at time k"

Exercise τ is a stopping time if and only if $\{\omega \in \Omega \mid \tau(\omega) = k\} \in \mathcal{F}_k$ for all $k \in \mathbb{Z}_{\geq 0}$.

Lemma: If τ is a stopping time, then the process $H_\bullet = (H_n)_{n \in \mathbb{Z}_{\geq 0}}$ defined by

$$H_n = \mathbb{1}_{\{n \leq \tau\}} = \begin{cases} 1 & \text{if } n \leq \tau \\ 0 & \text{if } n > \tau \end{cases}$$

is predictable.

Proof: Fix $n \in \mathbb{Z}_{\geq 0}$. To show that H_n is an $\mathcal{F}_{n-1}$-measurable random variable, we must show that $\{n \leq \tau\}$ is an $\mathcal{F}_{n-1}$-measurable event. This follows from

$$\{n \leq \tau\}^c = \{n > \tau\} = \bigcup_{k=0}^{n-1} \underbrace{\{\tau = k\}}_{\in \mathcal{F}_k \subset \mathcal{F}_{n-1}} \in \mathcal{F}_{n-1}.$$

□

If τ is an (almost surely) finite stopping time and $X_\bullet = (X_n)_{n \in \mathbb{Z}_{\geq 0}}$ is a process, then the value of the process at the (random) time τ is

$$X_\tau := \sum_{k=0}^{\infty} \mathbb{1}_{\{\tau = k\}} \cdot X_k$$

(The event $\{\tau = +\infty\}$ occurs with probability zero by assumption, so it does not matter what value we assign to X_τ on that event.)

Exercise If τ and σ are stopping times, then also $\tau \wedge \sigma = \min\{\tau, \sigma\}$ and $\tau \vee \sigma = \max\{\tau, \sigma\}$ are stopping times.

If τ is a stopping time then for any $n \in \mathbb{Z}_{\geq 0}$, $\tau \wedge n$ is a finite stopping time and $H_k = \mathbb{1}_{\{\tau \leq k\}}$ defines a bounded non-negative predictable process. Applying earlier results, we immediately get

Theorem (Stopped supermartingales are supermartingales) If $X_\bullet = (X_n)_{n \in \mathbb{Z}_{\geq 0}}$ is a supermartingale and τ is a stopping time then $(X_{n \wedge \tau})_{n \in \mathbb{Z}_{\geq 0}}$ is also supermartingale and $\mathbb{E}[X_{n \wedge \tau}] \leq \mathbb{E}[X_0]$ for any $n \in \mathbb{Z}_{\geq 0}$.

Proof: Let us prove (1°), (2°), and (3°_{super}).

(1°): $$X_{n\wedge\tau} = \sum_{k=0}^{n-1} \mathbb{1}_{\{\tau=k\}} X_k + \mathbb{1}_{\{\tau\geq n\}} X_n$$

$\in \mathcal{F}_k \subset \mathcal{F}_n$ $\quad$ $\in m\mathcal{F}_k \subset m\mathcal{F}_n$ $\quad$ $=\{\tau\leq n-1\}^c \in \mathcal{F}_{n-1}\subset\mathcal{F}_n$ $\quad$ $\in m\mathcal{F}_n$

This shows that $X_{n\wedge\tau}$ is $\mathcal{F}_n$-measurable, so $X_{\bullet\wedge\tau}$ is $\mathcal{F}_\bullet$-adapted.

(2°): The formula for $X_{n\wedge\tau}$ above is a finite sum of integrable terms (bounded indicator times integrable random variable) and as such also integrable, $X_{n\wedge\tau}\in\mathcal{L}^1(\mathbb{P})$.

(3°) Let $H_k = \mathbb{1}_{\{k\leq\tau\}}$. Then $H\cdot X$ is a supermartingale by earlier theorem.

But
$$(H\cdot X)_n = \sum_{k=1}^{n} H_k\cdot(X_k - X_{k-1})$$
$$= \sum_{k=1}^{n} \mathbb{1}_{\{k\leq\tau\}}(X_k - X_{k-1})$$
$$= \sum_{k=1}^{n\wedge\tau}(X_k - X_{k-1}) = X_{n\wedge\tau} - X_0 .$$

Since this is a supermartingale, also $X_{\bullet\wedge\tau}$ is, because adding X_0 to all values does not change (1°), (2°), (3°_{super}). □

Corollary: If $X_\bullet$ is a martingale and τ is a stopping time, then $X_{\bullet\wedge\tau}$ is a martingale and in particular
$$\mathbb{E}[X_{n\wedge\tau}] = \mathbb{E}[X_0].$$

Do we have $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$ for martingales?

Theorem (Doob's optional stopping theorem)

Let $X_\bullet = (X_n)_{n\in\mathbb{Z}_{\geq 0}}$ be a martingale and τ a stopping time. Then under any of the conditions (a), (b), or (c) below, we have that $X_\tau \in \mathcal{L}^1(\mathbb{P})$ and

$$\mathbb{E}[X_\tau] = \mathbb{E}[X_0].$$

Sufficient conditions:

(a): τ is a.s. bounded $\left(\exists k \text{ s.t. } \mathbb{P}[\tau\leq k]=1\right)$

(b): X is a.s. bounded $\left(\exists a \text{ s.t. } \forall n : \mathbb{P}[|X_n|\leq a]=1\right)$ and τ is a.s. finite $\left(\mathbb{P}[\tau<\infty]=1\right)$

(c): X has a.s. bounded increments $\left(\exists a \text{ st. } \forall n : \mathbb{P}[|X_n - X_{n-1}|\leq a] = 1\right)$ and τ is integrable $\left(\mathbb{E}[\tau]<\infty\right)$.

__Proof__: (a): Assume $\mathbb{P}[\tau\leq k]=1$. The previous corollary with $n=k$ gives

$$\mathbb{E}[X_{k\wedge\tau}] = \mathbb{E}[X_0].$$

But since $\tau\leq k$ a.s., we have $k\wedge\tau = \tau$ a.s., which implies $X_{k\wedge\tau} = X_\tau$ a.s. Thus the equation above proves the claim.

(b): Assume $\mathbb{P}[|X_n|\leq a]=1$ $\forall n$ and $\mathbb{P}[\tau<\infty]=1$. The previous corollary gives

$$\mathbb{E}[X_{n\wedge\tau}] = \mathbb{E}[X_0].$$

Since $\tau<\infty$ a.s., we have $n\wedge\tau\uparrow\tau$ as $n\to\infty$ a.s. Then also

$$X_{n\wedge\tau} \xrightarrow[n\to\infty]{} X_\tau \quad \text{a.s.}$$

Moreover, $X_{n\wedge\tau}$ are a.s. bounded, so the bounded convergence theorem gives

$$\mathbb{E}[X_\tau] = \mathbb{E}[\lim_{n\to\infty} X_{n\wedge\tau}]$$
$$\overset{BCT}{=} \lim_{n\to\infty} \underbrace{\mathbb{E}[X_{n\wedge\tau}]}_{=\mathbb{E}[X_0] \text{ by corollary}} = \mathbb{E}[X_0].$$

This proves the assertion.

(c): Assume $\mathbb{P}[\,|X_n - X_{n-1}| \le \alpha\,] = 1 \quad \forall n$
and $\mathbb{E}[\tau] < \infty$. (In particular $\mathbb{P}[\tau < \infty] = 1$.)

The previous corollary gives
$$\mathbb{E}[X_{n\wedge\tau}] = \mathbb{E}[X_0].$$

We will use dominated convergence theorem.

Note that a.s.
$$|X_{n\wedge\tau}| = \left| X_0 + \sum_{k=1}^{n\wedge\tau} (X_k - X_{k-1}) \right|$$
$$\le |X_0| + |\alpha| \cdot |n\wedge\tau| \qquad \text{(triangle ineq.)}$$
$$\le |X_0| + |\alpha| \cdot \tau$$

The right hand side is a dominating random variable, which is integrable, since $X_0 \in \mathcal{L}^1(\mathbb{P})$ and $\tau \in \mathcal{L}^1(\mathbb{P})$ and $\alpha \in \mathbb{R}$ is a constant.

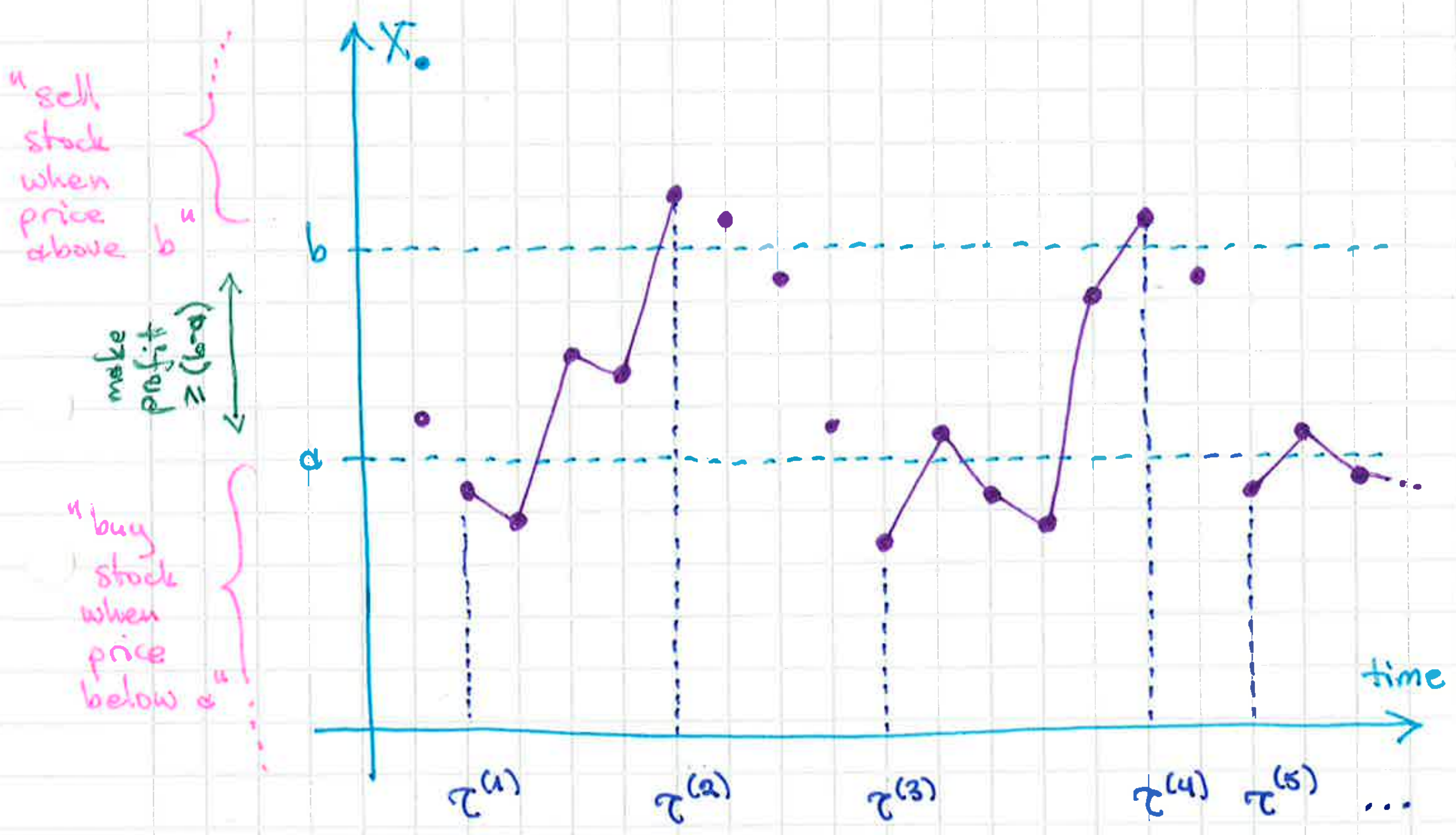
Again $X_{n\wedge\tau} \longrightarrow X_\tau$ a.s. (since $\tau < \infty$ a.s.)
so dominated convergence theorem gives
$$\mathbb{E}[X_\tau] = \mathbb{E}[\lim_{n\to\infty} X_{n\wedge\tau}]$$
$$\overset{DCT}{=} \lim_{n\to\infty} \underbrace{\mathbb{E}[X_{n\wedge\tau}]}_{=\mathbb{E}[X_0] \text{ by corollary}} = \mathbb{E}[X_0].$$

This proves the assertion. $\square$

Remark: There are versions of the optional stopping theorem for super- and submartingales. What are the statements and what changes in the proof?

Upcrossings and martingale convergence theorem

We will next prove an amazingly powerful result. The idea of the proof is an "investment strategy" of "buy low, sell high":
($a < b$)



"Buy on odd stopping times $\tau^{(1)}, \tau^{(3)}, \ldots$
Sell on even stopping times $\tau^{(2)}, \tau^{(4)}, \ldots$
In between make profit $\geq b-a$."

Set $\tau^{(0)} = -1$.

Define stopping times $\tau^{(1)} < \tau^{(2)} < \tau^{(3)} < \dots$ recursively

$$\tau^{(2j-1)} = \inf \{ n > \tau^{(2j-2)} \mid X_n \leq a \}$$ "time to buy!"

$$\tau^{(2j)} = \inf \{ n > \tau^{(2j-1)} \mid X_n \geq b \}$$ "time to sell!"

Now our "portfolio" is

$$H_k = \begin{cases} 1 & \text{if } \tau^{(2j-1)} < k \leq \tau^{(2j)} \text{ for some } j \\ 0 & \text{otherwise} \end{cases}$$

(see picture above!).

Let $U_n := \sup \{ j \in \mathbb{N} \mid \tau^{(2j)} \leq n \}$ be the number of upcrossings completed up to time n.

<u>Lemma</u> (Doob's upcrossing lemma)

If $X_\bullet$ is a submartingale, then

$$(b-a) \cdot \mathbb{E}[U_n] \leq \mathbb{E}\left[(X_n - a)^+\right] - \mathbb{E}\left[(X_0 - a)^+\right].$$

<u>Proof</u>: Let H be the "portfolio" above. It is predictable, bounded, and non-negative. The process $Y_\bullet = (Y_n)_{n \in \mathbb{Z}_{\geq 0}}$ defined by $Y_n = a + (X_n - a)^+$ is a submartingale by an earlier example. Thus also $H \cdot Y$ is a submartingale. Now observe that

$$(H \cdot Y)_n \geq (b-a) \cdot U_n$$

since every upcrossing gives at least profit $b-a$ and any final incomplete upcrossing attempt gives non-negative profit (because Y was truncated to never go below level a).

Consider also $K_k = 1 - H_k$. Then

$$Y_n - Y_0 = (H \cdot Y)_n + (K \cdot Y)_n .$$

Since $K \cdot Y$ is also a submartingale, we have $\mathbb{E}[(K\cdot Y)_n] \geq \mathbb{E}[(K\cdot Y)_0] = 0$.

Take expected values to get:

$$\mathbb{E}[Y_n - Y_0] \geq \mathbb{E}[(H\cdot Y)_n] \geq (b-a)\,\mathbb{E}[U_n].$$

$\square$

<u>Theorem</u> (Martingale convergence theorem)

If $X_\bullet = (X_n)_{n \in \mathbb{Z}_{\geq 0}}$ is a submartingale and $\sup_{n \in \mathbb{Z}_{\geq 0}} \mathbb{E}[X_n^+] < +\infty$, then as $n \longrightarrow \infty$, X_n converges almost surely to a limit random variable X with $\mathbb{E}[|X|] < +\infty$.

<u>Proof</u>: Note that if a sequence $x_0, x_1, x_2, \ldots \in \mathbb{R}$ of numbers does not converge, then

$$\liminf_n x_n < \limsup_n x_n ,$$

and then there exists rational numbers $a, b \in \mathbb{Q}$ with $\liminf_n x_n < a < b < \limsup_n x_n$.

Thus the sequence $(x_n)_{n \in \mathbb{Z}_{\geq 0}}$ upcrosses the interval $[a,b]$ infinitely many times.

For the asserted almost sure convergence it is therefore sufficient to show that almost surely $(X_n)_{n \in \mathbb{Z}_{\geq 0}}$ does not upcross any such interval infinitely many times.

For fixed a, b, the total number U of upcrossings is the increasing limit $U_n \uparrow U$ of the numbers of upcrossings by time n, as $n \to \infty$.

Doob's upcrossing lemma gives

$$\mathbb{E}[U_n] \leq \frac{\mathbb{E}[(X_n-a)^+]}{b-a} \leq \frac{\mathbb{E}[X_n^+]+|a|}{b-a}.$$

The right hand side is bounded, by assumption, so by Fatou's lemma

$$\mathbb{E}[U] = \mathbb{E}[\lim_{n\to\infty} U_n]$$
$$\leq \liminf_n \mathbb{E}[U_n] \leq \frac{1}{b-a}\left(\sup_n \mathbb{E}[X_n^+]+|a|\right)$$
$$< +\infty.$$

In particular U is almost surely finite, i.e. $\mathbb{P}[(X_n) \text{ upcrosses } [a,b] \text{ infinitely often}] = 0$.

Union bound (countable unions over rationals) gives

$$\mathbb{P}[\text{there exists } a,b\in\mathbb{Q} \text{ s.t. } (X_n) \text{ upcrosses } [a,b] \text{ infinitely often}]$$

$$\leq \sum_{a,b\in\mathbb{Q}} \mathbb{P}[(X_n) \text{ upcrosses } [a,b] \text{ inf. oft.}] = \sum_{a,b\in\mathbb{Q}} 0 = 0.$$

Therefore $\mathbb{P}[\lim_{n\to\infty} X_n \text{ exists}] = 1$.

Denote $X := \lim_{n\to\infty} X_n$. Fatou's lemma gives

$$\mathbb{E}[X^+] = \mathbb{E}[\lim_{n\to\infty} X_n^+] \leq \liminf_n \mathbb{E}[X_n^+]$$
$$\leq \sup_n \mathbb{E}[X_n^+] < \infty.$$

On the other hand, since $X_\bullet$ is a submgale,

$$\mathbb{E}[X_n^-] = \mathbb{E}[X_n^+] - \mathbb{E}[X_n] \leq \mathbb{E}[X_n^+] - \mathbb{E}[X_0].$$

Again Fatou's lemma gives

$$\mathbb{E}[X^-] = \mathbb{E}[\lim_{n\to\infty} X_n^-] \leq \liminf_n \mathbb{E}[X_n^-]$$
$$\leq \sup_n \mathbb{E}[X_n^+] - \mathbb{E}[X_0] < \infty.$$

This shows that $\mathbb{E}[|X|] < \infty$. $\square$

MARKOV PROPERTIES OF BROWNIAN MOTION

We will now start discussing fundamental properties of the Brownian motion as a continuous time stochastic process. In particular, our goal is to prove Markov properties, which intuitively state that:

"For predicting what happens in the future, any past information besides the current state is irrelevant."

Time will now be indexed by

$$t \in \mathbb{R}_+ = [0, \infty),$$

so let us begin by discussing how some basic concepts generalize to continuous time.

Def: A <u>filtration</u> is a colection $\mathcal{F}_\bullet = (\mathcal{F}_t)_{t \in \mathbb{R}_+}$ of σ-algebras $\mathcal{F}_t \subset \mathcal{F}$ such that

$$\mathcal{F}_s \subset \mathcal{F}_t \quad \text{whenever} \quad s < t.$$

Def: A random variable $\tau : \Omega \to [0, +\infty]$ is a <u>stopping time</u> w.r.t. $\mathcal{F}_\bullet$ if

$$\{\omega \in \Omega \mid \tau(\omega) \leq t\} \in \mathcal{F}_t \qquad \forall t \in \mathbb{R}_+$$

Remark: Unlike in discrete time, it is not particularly important whether we use $\{\tau \leq t\}$ or $\{\tau < t\}$ in the definition above — at least if the filtration $\mathcal{F}_\bullet$ is <u>right continuous</u> so that $\mathcal{F}_t = \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}$.

Indeed, if $\{\tau \leq t\} \in \mathcal{F}_t \quad \forall t \in \mathbb{R}_+$, then

$$\{\tau < t\} = \bigcup_{n=1}^{\infty} \underbrace{\{\tau \leq t - \tfrac{1}{n}\}}_{\in \mathcal{F}_{t-1/n} \subset \mathcal{F}_t} \in \mathcal{F}_t .$$

Conversely, if $\{\tau < t\} \in \mathcal{F}_t \quad \forall t \in \mathbb{R}_+$, then

$$\{\tau \leq t\} = \bigcap_{n=n_0}^{\infty} \underbrace{\{\tau < t + \tfrac{1}{n}\}}_{\in \mathcal{F}_{t+1/n} \subset \mathcal{F}_{t+1/n_0}} \in \mathcal{F}_{t+1/n_0} .$$

for any $n_0 \in \mathbb{N}$.

Therefore, if $\mathcal{F}_\bullet$ is right continuous, then

$$\{\tau \leq t\} \in \bigcap_{n_0 \in \mathbb{N}} \mathcal{F}_{t+1/n_0} = \mathcal{F}_t .$$

Def: A stochastic process $X_\bullet = (X_t)_{t \in \mathbb{R}_+}$ is adapted to a filtration $\mathcal{F}_\bullet = (\mathcal{F}_t)_{t \in \mathbb{R}_+}$ if for all $t \in \mathbb{R}_+$ the random variable X_t is $\mathcal{F}_t$-measurable: $X_t \in m\mathcal{F}_t$.

Def: A stochastic process $X_\bullet = (X_t)_{t \in \mathbb{R}_+}$ is a martingale w.r.t. filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ if we have

(1°) $X_\bullet$ is adapted to $\mathcal{F}_\bullet$

(2°) $X_t \in \mathcal{L}^1(\mathbb{P})$ for each $t \in \mathbb{R}_+$

(3°) for all $0 \leq s < t$ we have

$$\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s .$$

Supermartingales and submartingales are defined with "$\leq$" and "$\geq$" in place of "$=$" in property (3°).

WIENER SPACE AND WIENER MEASURE

Recall that a stochastic process $B = (B_t)_{t \in \mathbb{R}_+}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a standard Brownian motion if

(BM-$\perp\!\!\!\perp$): for $0 = t_0 < t_1 < t_2 < \cdots < t_n$, the increments $B_{t_1} - B_{t_0}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent

(BM-N): for $0 \leq s < t$, the increment has Gaussian distribution $B_t - B_s \sim N(0, t-s)$.

(BM-0): $\mathbb{P}[B_0 = 0] = 1$

(BM-c): $\mathbb{P}[t \mapsto B_t \text{ is continuous}] = 1$.

Then the law of B is a probability measure on the space of continuous functions

$$C(\mathbb{R}_+, \mathbb{R}) = \{ w : \mathbb{R}_+ \to \mathbb{R} \mid w \text{ continuous} \}$$

equipped with the σ-algebra

$$\mathcal{W} = \sigma(ev_t : t \in \mathbb{R}_+)$$

generated by the "projections" (evaluations at a single time instant)

$$\begin{aligned} ev_t : C(\mathbb{R}_+, \mathbb{R}) &\longrightarrow \mathbb{R} \\ w &\longmapsto w_t \end{aligned}$$

The π-system of "finite dimensional events" of the form

$$A = \{ w \in C(\mathbb{R}_+, \mathbb{R}) \mid w_{t_1} \in A_1, \ldots, w_{t_n} \in A_n \}$$

for $n \in \mathbb{N}$, $0 \leq t_1 < t_2 < \cdots < t_n$
$A_1, A_2, \ldots, A_n \subset \mathbb{R}$ Borel

generates the σ-algebra $\mathcal{W}$.

The law of a standard Brownian motion is denoted by $\mathbb{W}_0$ and called "the Wiener measure" — thus

$$\left(C(\mathbb{R}_+, \mathbb{R}) \,,\, \mathcal{W} \,,\, \mathbb{W}_0 \right)$$

is a probability space.

(On this space, the "canonical process" $W_\bullet = (W_t)_{t \in \mathbb{R}_+}$ given by $W_t(w) = w_t$ is a standard Brownian motion...)

We also equip the space with other probability measures. Most importantly, for $x \in \mathbb{R}$, let $\mathbb{W}_x$ denote the law of the process

$$(B_t + x)_{t \in \mathbb{R}_+} \quad \text{where} \quad (B_t)_{t \in \mathbb{R}} \text{ is a std BM.}$$

This is the "Brownian motion started from x".

It is also meaningful to start from a random point X, but X should then be independent of how the process continues! The corresponding measure is

$$A \longmapsto \int_{\mathbb{R}} dP_X(x) \; \mathbb{W}_x[A] \qquad \text{for} \quad A \in \mathcal{W}$$

law of X

One more piece of notation will be extensively used later. If

$$F : C(\mathbb{R}_+, \mathbb{R}) \longrightarrow \mathbb{R}$$

is a $\mathcal{W}$-measurable function which is integrable w.r.t. $\mathbb{W}_x$, i.e., $F \in \mathcal{L}^1(\mathbb{W}_x)$ (for example just bounded) then we set

$$\mathbb{E}_x\left[F(B) \right] := \int_{C(\mathbb{R}_+, \mathbb{R})} F(w) \, d\mathbb{W}_x(w) \, .$$

This is the expected value of the functional F of the Brownian path B, which is started from $x \in \mathbb{R}$.

The notation on the LHS may be slightly confusing, since we use this also when we already have a Brownian motion B started from 0 (or some other point). If you are ever confused by it, then use the defining formula on the RHS which manifestly does not involve any new B.

This will be used with random starting point.

(Note that we thus average over the randomness in the Brownian path in $\int \ldots d\mathbb{W}_x$ but we keep the randomness of the starting point.)

For $X: \Omega \to \mathbb{R}$ such random starting point, the notation

$$\mathbb{E}_X[F(B)]$$

is the random variable $\Omega \to \mathbb{R}$ given by

$$\omega \longmapsto \mathbb{E}_{X(\omega)}[F(B)] = \int_{C(\mathbb{R}_+,\mathbb{R})} F(w)\, d\mathbb{W}_{X(\omega)}(w).$$

this is omega

omega

this is double-U

this is omega

this is double-U

The space $C(\mathbb{R}_+,\mathbb{R})$ also has time-shift operators θ_s, for $s \geq 0$, defined by

$$\theta_s : C(\mathbb{R}_+,\mathbb{R}) \longrightarrow C(\mathbb{R}_+,\mathbb{R})$$

$$\overset{\omega}{(w_t)_{t\in\mathbb{R}_+}} \overset{\theta_s}{\longmapsto} (w_{s+t})_{t\in\mathbb{R}_+}.$$

These are used in one formulation of the Markov property.

Natural filtrations of Brownian motion

Finally, before starting, let us define the filtrations used.

Let $B = (B_t)_{t\in\mathbb{R}_+}$ be a standard Brownian motion.

The "naive" definition of information available at time $t\in\mathbb{R}_+$ is

$$\mathcal{F}_t^0 := \sigma(B_s : s\leq t),$$

the sigma-algebra generated by past and present values.

A minor issue is that not all subsets of zero probability events are measurable, so we complete it by adding these:

$$\mathcal{F}_t^B := \sigma(\mathcal{F}_t^0 \cup \mathcal{N})$$

where $\mathcal{N} = \{A\subset\Omega \mid A\subset N\in\mathcal{F} \text{ for some } N \text{ s.t. } \mathbb{P}[N]=0\}$.

Another issue is that the filtration is not yet right continuous. This is solved by "allowing an infinitesimal peek into the future":

$$\mathcal{F}_{t^+}^B = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}^B .$$

We mainly use this filtration

$$(\mathcal{F}_{t^+}^B)_{t\in\mathbb{R}_+} .$$

Also denote $\mathcal{F}_\infty^B = \sigma(\mathcal{F}_t^B : t\in\mathbb{R}_+)$ the σ-algebra generated by the entire process (completed by zero proba events)

THE SIMPLE MARKOV PROPERTY

(ABBREVIATE: std BM = standard Brownian motion)

Theorem 1.10 (Simple Markov property)

Let $B = (B_t)_{t \in \mathbb{R}_+}$ be a std BM and $s \geq 0$.
Then $\tilde{B} = (\tilde{B}_t)_{t \in \mathbb{R}_+}$ defined by $\tilde{B} = \theta_s(B) - B_s$

i.e. $\tilde{B}_t = B_{s+t} - B_s \quad \forall\, t \in \mathbb{R}_+$

is a standard Brownian motion independent of $\mathcal{F}^B_{s+}$.

We first check that $\tilde{B}$ has the right law.

Proof that $\tilde{B}$ is a std BM:

Let $0 = t_0 < t_1 < t_2 < \cdots < t_n$. Now

$$\tilde{B}_{t_1} - \tilde{B}_{t_0} = (B_{s+t_1} - B_s) - (B_{s+t_0} - B_s) = B_{s+t_1} - B_{s+t_0}$$
$$\vdots$$
$$\tilde{B}_{t_n} - \tilde{B}_{t_{n-1}} = (B_{s+t_n} - B_s) - (B_{s+t_{n-1}} - B_s) = B_{s+t_n} - B_{s+t_{n-1}}.$$

Since B is a std BM, these are independent Gaussian random variables with mean 0 and variances $(s+t_1)-(s+t_0) = t_1 - t_0$, ..., $(s+t_n)-(s+t_{n-1}) = t_n - t_{n-1}$.
Such finite dimensional distributions are enough to specify the law of $\tilde{B}$, so indeed $\tilde{B}$ is a std BM. $\square$

It is also easy to prove that $\tilde{B}$ is independent at least of $\mathcal{F}^B_s$ — we use the π-system of finite dimensional events before and after time s. A little extra observation will be needed to show independence of $\mathcal{F}^B_{s+}$.

Proof that $\tilde{B} \perp\!\!\!\perp \mathcal{F}_s^B$.

Let $0 = s_0 < s_1 < \dots < s_m = s$ and

$0 = t_0 < t_1 < \dots < t_n$.

Then since B is a std BM, we have that

$$B_{s_1} - B_{s_0}, \dots, B_{s_{m-1}} - B_{s_{m-2}}, B_s - B_{s_{m-1}},$$
$$\underbrace{B_{s+t_1} - B_s}_{= \tilde{B}_{t_1} - \tilde{B}_{t_0}}, \dots, \underbrace{B_{s+t_n} - B_{s+t_{n-1}}}_{= \tilde{B}_{t_n} - \tilde{B}_{t_{n-1}}}$$

are independent. Therefore for any $A_1, \dots, A_m \subset \mathbb{R}$ and $\tilde{A}_1, \dots, \tilde{A}_n \subset \mathbb{R}$ Borel, the events

$$A = \{B_{s_1} \in A_1, \dots, B_{s_m} \in A_m\} \quad \text{and}$$
$$\tilde{A} = \{\tilde{B}_{t_1} \in \tilde{A}_1, \dots, \tilde{B}_{t_n} \in \tilde{A}_n\}$$

are independent (the former can be expressed in terms of $B_{s_1} - B_{s_0}, \dots, B_s - B_{s_{m-1}}$ and the latter in terms of $B_{s+t_1} - B_s, \dots, B_{s+t_n} - B_{s+t_{n-1}}$).

Events of these forms constitute π-systems, which generate σ-algebras $\mathcal{F}_s^0$ and $\sigma(\tilde{B}_t : t \in \mathbb{R}_+)$ respectively, so the σ-algebras are independent.

Zero probability events do not affect independence, so indeed $\mathcal{F}_s^B \perp\!\!\!\perp \sigma(\tilde{B}_t : t \in \mathbb{R}_+)$. $\square$

Then let us handle the infinitesimal peek into the future.

We will use the same finite dimensional marginal on times $0 < t_1 < t_2 < \dots < t_n$.

Moreover, to characterize the measure on $\mathbb{R}^n$, we use the observation:

Lemma: Suppose that μ and ν are two finite Borel measures on $\mathbb{R}^n$ such that for every bounded continuous $g : \mathbb{R}^n \to \mathbb{R}$ we have
$$\int_{\mathbb{R}^n} g \, d\mu = \int_{\mathbb{R}^n} g \, d\nu .$$
Then $\mu = \nu$.

(Sketch: For $U \subset \mathbb{R}^n$ open can approximate $0 \leq g_n \uparrow \mathbb{1}_U$, so MCT gives $\mu[U] = \lim_{n\to\infty} \int g_n d\mu = \lim_{n\to\infty} \int g_n d\nu = \nu[U]$. Complements U^c form π-system. $\square$)

So let us finish the proof of Theorem 1.10.

For details, see e.g. MS-E1602 "Large Random Systems"

Proof that $\tilde{B} \perp\!\!\!\perp \mathcal{F}^B_{s+}$:

Let $A \in \mathcal{F}^B_{s+} := \bigcap_{\varepsilon>0} \mathcal{F}^B_{s+\varepsilon}$. We wish to show that for any $0 = t_0 < t_1 < \cdots < t_n$ and any bounded continuous $g : \mathbb{R}^n \to \mathbb{R}$ we have
$$\mathbb{E}\big[\mathbb{1}_A \cdot g(\tilde{B}_{t_1} - \tilde{B}_{t_0}, \ldots, \tilde{B}_{t_n} - \tilde{B}_{t_{n-1}})\big] \qquad (\star)$$
$$= \mathbb{P}[A] \cdot \mathbb{E}\big[g(\tilde{B}_{t_1} - \tilde{B}_0, \ldots, \tilde{B}_{t_n} - \tilde{B}_{t_{n-1}})\big]$$

Each side of $(\star)$ is an integral of g against a certain measure, so by the above lemma this will guarantee $\mathbb{P}[A \cap \tilde{A}] = \mathbb{P}[A] \cdot \mathbb{P}[\tilde{A}]$ for any $\tilde{A}$ finite dimensional event for $\tilde{B}$. This implies indedendence of $\mathcal{F}^B_{s+}$ of the π-system of finite-dim. events for $\tilde{B}$, which is sufficient to conclude the proof.

It remains to prove $(\star)$. First take $\varepsilon > 0$. We have already shown that $(B_{s+\varepsilon+t_1} - B_{s+\varepsilon}, \ldots, B_{s+\varepsilon+t_n} - B_{s+\varepsilon+t_{n-1}})$ is independent of $\mathcal{F}^B_{s+\varepsilon}$. Since also $A \in \mathcal{F}^B_{s+} \subset \mathcal{F}^B_{s+\varepsilon}$, we get

$$\mathbb{E}\left[\mathbb{1}_A \cdot g\left(B_{s+\varepsilon+t_1}-B_{s+\varepsilon}, \dots, B_{s+\varepsilon+t_n}-B_{s+\varepsilon+t_{n-1}}\right)\right]$$
$$= \mathbb{P}[A] \cdot \mathbb{E}\left[g\left(B_{s+\varepsilon+t_1}-B_{s+\varepsilon}, \dots, B_{s+\varepsilon+t_n}-B_{s+\varepsilon+t_{n-1}}\right)\right].$$

Now we want to let $\varepsilon \downarrow 0$.

Note first that $t \mapsto B_t$ is continuous, so

$$B_{s+\varepsilon+t_j} \xrightarrow[\varepsilon \downarrow 0]{} B_{s+t_j} \quad \text{for } j=0,1,\dots,n.$$

Therefore by continuity of g we have

$$g\left(B_{s+\varepsilon+t_1}-B_{s+\varepsilon}, \dots, B_{s+\varepsilon+t_n}-B_{s+\varepsilon+t_{n-1}}\right)$$
$$\xrightarrow[\varepsilon \downarrow 0]{} g\left(\tilde{B}_{t_1}-\tilde{B}_{t_0}, \dots, \tilde{B}_{t_n}-\tilde{B}_{t_{n-1}}\right).$$

Moreover, g is bounded, so by bounded convergence theorem in the limit $\varepsilon \downarrow 0$ we obtain (★). This finishes the proof. □

This may have seemed like a lot of work to gain just an infinitesimal peek into the future. There are remarkable consequences, however.

<u>Theorem 1.12</u> (Blumenthal's 0-1 law)

Let B be a standard Brownian motion. Then the σ-algebra $\mathcal{F}^B_{0+} = \bigcap_{\varepsilon>0} \mathcal{F}^B_\varepsilon$ is trivial in the following sense: if $A \in \mathcal{F}^B_{0+}$ then $\mathbb{P}[A]=0$ or $\mathbb{P}[A]=1$.

<u>Proof</u>: By Thm 1.10 with $s=0$, B itself is independent of $\mathcal{F}^B_{0+}$. This shows $\mathcal{F}^B_\infty \perp\!\!\!\perp \mathcal{F}^B_{0+}$. But of course $\mathcal{F}^B_{0+} \subset \mathcal{F}^B_\infty$, so a fortiori we have $\mathcal{F}^B_{0+} \perp\!\!\!\perp \mathcal{F}^B_{0+}$. Therefore any $A \in \mathcal{F}^B_{0+}$ satisfies

$$\mathbb{P}[A] = \mathbb{P}[A \cap A] = \mathbb{P}[A] \cdot \mathbb{P}[A].$$

The only solutions to $p = p^2$ are $p=0, p=1$. □

Let us give some examples of surprising consequences.

__Corollary__ Let $S_t = \sup_{0 \le s \le t} B_s$ and $I_t = \inf_{0 \le s \le t} B_s$. Then for any $\varepsilon > 0$ we have almost surely $S_\varepsilon > 0$ and $I_\varepsilon < 0$.

__Proof__: For any $t > 0$ we have $\mathbb{P}[B_t > 0] = \frac{1}{2}$ (centered Gaussian B_t). Take a sequence $(t_n)_{n \in \mathbb{N}}$ of positive times tending to zero, $t_n \to 0$. Then reverse Fatou's lemma gives

$$\mathbb{P}\Big[\limsup_n \{B_{t_n} > 0\}\Big] \geq \limsup_n \mathbb{P}[B_{t_n} > 0] \geq \frac{1}{2}.$$

But $\limsup_n \{B_{t_n} > 0\} \in \mathcal{F}^B_\varepsilon$ for any $\varepsilon > 0$ and therefore $\limsup_n \{B_{t_n} > 0\} \in \mathcal{F}^B_{0^+}$.
By Blumenthal's 0-1 law, then, we must have

$$\mathbb{P}\Big[\limsup_n \{B_{t_n} > 0\}\Big] = 1.$$

This proves that $S_\varepsilon > 0$ almost surely, for any $\varepsilon > 0$. The assertion $I_\varepsilon < 0$ a.s. is proven similarly. $\square$

__Corollary__ For any $\varepsilon > 0$, almost surely there exists a time instant $0 < t < \varepsilon$ such that $B_t = 0$.

__Proof__ Since $I_\varepsilon < 0$ and $S_\varepsilon > 0$ and $t \mapsto B_t$ is continuous, this follows from mean value theorem $\square$

__Remark__: There are in fact infinitely many zeroes on $(0, \varepsilon)$. Indeed, if there were only finitely many, then some smaller interval $(0, \delta)$ would not have zeroes, which contradicts the above.

We do not prove it here, but there are in fact uncountably many zeroes on $(0, \varepsilon)$, the Hausdorff dimension of the set of zeroes is $\frac{1}{2}$.

Corollary: Almost surely $\sup_{t\in\mathbb{R}_+} B_t = +\infty$ and $\inf_{t\in\mathbb{R}_+} B_t = -\infty$.

Proof: By scaling property $B_{\lambda t} \overset{\text{law}}{=} \sqrt{\lambda}\cdot B_t$ we get for $S_\infty = \sup_{t\in\mathbb{R}_+} B_t$ that $\sqrt{\lambda}\, S_\infty \overset{\text{law}}{=} S_\infty$. This is only possible if S_∞ takes values in $\{0,+\infty\}$ (think about the c.d.f.). But by earlier corollary $S_\infty \geq S_\varepsilon > 0$, so we must have $\mathbb{P}[S_\infty = +\infty] = 1$. The other claim is similar (or obtained by reflection). $\square$

Let us then turn to another formulation of the Markov property.

Markov property as a conditional expectation

Theorem (Conditional formulation of Markov property)
Let $B = (B_t)_{t\in\mathbb{R}_+}$ be a standard Brownian motion. Then for any $s\geq 0$ and any bounded $\mathcal{W}$-measurable function

$$F : C(\mathbb{R}_+,\mathbb{R}) \longrightarrow \mathbb{R}$$

we have

$$\mathbb{E}\left[F(\theta_s(B)) \,\middle|\, \mathcal{F}^B_{s^+}\right] = \mathbb{E}_{B_s}\left[F(B)\right]$$

conditional expected value of the real random variable $\omega \mapsto F(\theta_s(B(\omega)))$ given information $\mathcal{F}^B_{s^+}$ at time s.

random variable obtained as the function $x \mapsto \mathbb{E}_x[F] = \int F \, d\mathcal{W}_x$ applied to the random variable B_s, i.e. $\omega \mapsto \mathbb{E}_{B_s(\omega)}[F]$.

Recall notation: $\theta_s : C(\mathbb{R}_+,\mathbb{R}) \to C(\mathbb{R}_+,\mathbb{R})$ shift by s time units $(\omega_t)_{t\in\mathbb{R}_+} \mapsto (\omega_{s+t})_{t\in\mathbb{R}_+}$

Sketch of proof: The RHS is a function of B_s, and as such it is clearly measurable w.r.t. $\mathcal{F}_s^B$ and a fortiori w.r.t. $\mathcal{F}_{s+}^B$.

Also since F is bounded, both $F(\theta_s(B))$ and $\mathbb{E}_{B_s}[F]$ are bounded random variables, and in particular integrable.

It therefore remains to show that for all $A \in \mathcal{F}_{s+}^B$ we have

$$\mathbb{E}\left[\mathbb{1}_A \cdot F(\theta_s(B))\right] = \mathbb{E}\left[\mathbb{1}_A \cdot \mathbb{E}_{B_s}[F]\right]. \qquad \text{(MP)}$$

The idea is to show this first for simple enough A and simple enough F, and then argue that it must hold more generally.

If A is a finite dimensional event and F is of the form

$$F(w) = \prod_{j=1}^{n} g_j(w_{t_j})$$

for $0 < t_1 < t_2 < \dots < t_n$ and bounded continuous functions $g_1, g_2, \dots, g_n : \mathbb{R} \to \mathbb{R}$,

then (MP) is proven by calculations similar to what we did in the proof of Thm 1.10.

From the π-system of finite-dim. events, one can extend to all $A \in \mathcal{F}_{s+}^B$ as before.

It is straightforward to show that the collection of F for which (MP) holds is a monotone class of functions. The special case checked first implies that the class contains indicators of events in a generating π-syst. so Monotone Class Theorem concludes $\square$

The simplest application of this conditional formulation of Markov property is:

<u>Proposition</u>: Let $0 \leq s < t$ and let $f : \mathbb{R} \to \mathbb{R}$ be a bounded Borel function. Then we have

$$\mathbb{E}\left[f(B_t) \mid \mathcal{F}^B_{s+} \right] = \mathbb{E}_{B_s}\left[f(B_{t-s}) \right]$$

<u>Proof</u>: Use $F : C(\mathbb{R}_+, \mathbb{R}) \to \mathbb{R}$ given by

$F(w) = f(w_{t-s})$. Then we have

$F(\theta_s(w)) = f(w_t)$, so $F(\theta_s(B)) = f(B_t)$.

The statement therefore follows directly from the Markov property (conditional formul.) $\square$

The martingale property of Brownian motion is one (relatively) direct consequence. (It could be proven with the simple Markov property, too...)

<u>Theorem</u>: Standard Brownian motion $(B_t)_{t \in \mathbb{R}_+}$ is a martingale w.r.t. filtration $(\mathcal{F}^B_{t+})_{t \in \mathbb{R}_+}$.

<u>Proof</u> Define $f_n : \mathbb{R} \to \mathbb{R}$ by $f_n(x) = \begin{cases} +n & \text{if } x \geq n \\ x & \text{if } -n < x < n \\ -n & \text{if } x \leq -n \end{cases}$.

Then f_n is bounded and continuous (thus Borel).

Also $f_n(x) \longrightarrow x$ as $n \to \infty$, for any $x \in \mathbb{R}$.

The previous proposition gives

$$\mathbb{E}\left[f_n(B_t) \mid \mathcal{F}^B_{s+} \right] = \mathbb{E}_{B_s}\left[f_n(B_{t-s}) \right].$$

Since B_t is Gaussian, it is integrable. We have $|f_n(B_t)| \leq |B_t|$, so dominated convergence gives in the limit $n \to \infty$

$$\mathbb{E}\left[B_t \mid \mathcal{F}^B_{s+} \right] = \mathbb{E}_{B_s}\left[B_{t-s} \right].$$

Of course $\mathbb{E}_x[B_u] = x$ for any $x \in \mathbb{R}$ and $u \geq 0$ (Brownian motion started from x has mean x at time u), so the RHS is B_s and we have shown

$$\mathbb{E}[B_t \mid \mathcal{F}^B_{s+}] = B_s .$$

B is also adapted to $(\mathcal{F}^B_{t+})_{t \in \mathbb{R}_+}$, so it is a martingale. $\square$

Let us give another example application.

<u>Claim:</u> Let $\tau_0 = \inf\{t > 0 \mid B_t = 0\}$ be the first hitting time of origin, and define

$$h(x;t) = \mathbb{E}_x\left[\mathbb{1}_{\{\tau_0 > t\}}\right] = \mathbb{P}_x[\tau_0 > t].$$

Let $R^{(1)} = \inf\{t > 1 \mid B_t = 0\}$ be the first "return" time to origin after time 1. Then

$$\mathbb{P}[R^{(1)} > 1+u] = \int_{\mathbb{R}} p_1(0,x)\, h(x;u)\, dx$$

<u>Proof:</u> On $C(\mathbb{R}_+, \mathbb{R}) = \{w : \mathbb{R}_+ \to \mathbb{R} \text{ continuous function}\}$ we define correspondingly

$$\tau_0(w) = \inf\{t > 0 \mid w_t = 0\} \qquad \text{and}$$

$$R^{(1)}(w) = \inf\{t > 1 \mid w_t = 0\}.$$

and furthermore, for given $u \geq 0$,

$$F(w) = \mathbb{1}_{\{\tau_0(w) > u\}} = \begin{cases} 1 & \text{if } \tau_0(w) > u \\ 0 & \text{otherwise.} \end{cases}$$

If $\theta_1 : C(\mathbb{R}_+, \mathbb{R}) \to C(\mathbb{R}_+, \mathbb{R})$ is the unit time shift operator

$$(w_t)_{t \in \mathbb{R}_+} \xmapsto{\theta_1} (w_{1+t})_{t \in \mathbb{R}_+}$$

then $F(\theta_1(w)) = \mathbb{1}_{\{R^{(1)} > 1+u\}}$ since

$$\tau_0(\theta_1(w)) = \inf\{t \geq 0 \mid w_{1+t} = 0\} = \inf\{t' \geq 1 \mid w_{t'} = 0\} - 1$$

$$= R^{(1)}(w) - 1.$$

Note also that by definition

$$\mathbb{E}_x[F(B)] =: h(x;u).$$

Let us then apply Markov property with $s=1$ and this F:

$$\mathbb{E}[F(\theta_1(B)) \mid \mathcal{F}^B_{1^+}] = \mathbb{E}_{B_1}[F(B)]$$
$$= h(B_1;u).$$

Now take expected values of this equality (and use property (ii) of conditional expected value):

$$\mathbb{E}[F(\theta_1(B))] = \mathbb{E}[h(B_1;u)].$$

The LHS and RHS above can be evaluated as follows:

$$\text{LHS} = \mathbb{E}[F(\theta_1(B))] = \mathbb{E}[\mathbb{1}_{\{R^{(1)} > 1+u\}}]$$
$$= \mathbb{P}[R^{(1)} > 1+u]$$

$$\text{RHS} = \mathbb{E}[h(B_1;u)] \qquad (B_1 \sim N(0,1))$$
$$= \int_{\mathbb{R}} p_1(0,x)\, h(x,u)\, dx.$$

This proves the claim. $\square$

$(\Omega, \mathcal{F}, \mathbb{P})$ probability space

STOPPING TIMES FOR CONTINUOUS TIME PROCESSES

Assume throughout that $\mathcal{F}_\bullet = (\mathcal{F}_t)_{t \in \mathbb{R}_+}$ is a right continuous filtration, i.e.

- $\forall t \in \mathbb{R}_+ :$ $\mathcal{F}_t \subset \mathcal{F}$ is a sub-σ-algebra
- $\forall\, 0 \leq s < t :$ $\mathcal{F}_s \subset \mathcal{F}_t$ "increasing information"
- $\forall t \in \mathbb{R}_+ :$ $\mathcal{F}_t = \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}$ "right-continuity".

Denote also $\mathcal{F}_\infty = \sigma\Big(\bigcup_{t \in \mathbb{R}_+} \mathcal{F}_t \Big)$ the smallest σ-algebra which contains each $\mathcal{F}_t$, $t \in \mathbb{R}_+$.

Recall that a random variable $\tau : \Omega \to \mathbb{R}_+ \cup \{+\infty\}$ is said to be a stopping time if for every $t \in \mathbb{R}_+$

$$\{\tau \leq t\} = \{\omega \in \Omega \mid \tau(\omega) \leq t\} \in \mathcal{F}_t .$$

Remark: For a right-continuous filtration equivalently

$$\{\tau < t\} = \{\omega \in \Omega \mid \tau(\omega) < t\} \in \mathcal{F}_t \qquad \forall t \in \mathbb{R}_+ .$$

Let us give some examples: the first hitting times of (topologically nice) subsets.

Let $X = (X_t)_{t \in \mathbb{R}_+}$ be a stochastic process with values in a metric space $(\mathfrak{X}, \varrho)$, and such that the paths $t \longmapsto X_t(\omega)$ are continuous $\mathbb{R}_+ \to \mathfrak{X}$ for every $\omega \in \Omega$. Assume also, as always, that X is adapted to the filtration $\mathcal{F}_\bullet$ in the sense that

$$X_t : \Omega \to \mathfrak{X} \quad \text{is} \quad \mathcal{F}_t \,/\, \mathcal{B}(\mathfrak{X})\text{-measurable}$$

Borel σ-algebra on $\mathfrak{X}$.

Then:

Lemma (i) For any open subset $U \subset \mathcal{X}$ the first hitting time $\tau_U = \inf\{t \geq 0 \mid X_t \in U\}$ is a stopping time.

(ii) For any closed subset $F \subset \mathcal{X}$ the first hitting time $\tau_F = \inf\{t \geq 0 \mid X_t \in F\}$ is a stopping time.

Proof (i): For any $\omega \in \Omega$, the set
$$\{t \in \mathbb{R}_+ \mid X_t(\omega) \in U\}$$
is an open subset of $\mathbb{R}_+$ by continuity of the path $t \mapsto X_t(\omega)$.

Therefore
$$\{\omega \in \Omega \mid \tau_U(\omega) < t\}$$
$$= \underbrace{\bigcup_{q \in [0,t) \cap \mathbb{Q}}}_{\text{countable union}} \underbrace{\{\omega \in \Omega \mid \tau_U(\omega) \leq q\}}_{\substack{\in \mathcal{F}_q \subset \mathcal{F}_t \\ \text{since } \tau \text{ is} \\ \text{a stopping time}}} \in \mathcal{F}_t .$$

(Since $\mathcal{F}_t$ is a σ-alg.)

Since the filtration is right continuous, this implies that τ is a stopping time.

(ii): Define $\operatorname{dist}(x, F) = \inf_{y \in F} \rho(x, y)$.

$x \mapsto \operatorname{dist}(x, F)$ is a continuous function $\mathcal{X} \to [0, \infty)$ which vanishes only for $x \in F$. Now

$$\{\omega \in \Omega \mid \tau_F(\omega) \leq t\} = \{\omega \in \Omega \mid \inf_{s \in [0,t]} \operatorname{dist}(X_s(\omega), F) = 0\}$$
$$= \{\omega \in \Omega \mid \underbrace{\inf_{q \in [0,t] \cap \mathbb{Q}}}_{\text{countable infimum}} \underbrace{\operatorname{dist}(X_q(\omega), F)}_{\substack{\mathcal{F}_q\text{-measurable} \\ \Downarrow \\ \text{also } \mathcal{F}_t\text{-measurable}}} = 0\} \in \mathcal{F}_t .$$

$\square$

From given stopping times, we can also construct new ones.

<u>Exercise</u> If σ and τ are stopping times, then also $\sigma \wedge \tau := \min(\sigma, \tau)$ and $\sigma \vee \tau := \max(\sigma, \tau)$ are stopping times.

<u>Lemma</u> Let $\tau_1, \tau_2, \ldots$ be stopping times.

(i): If $\tau_n \uparrow \tau$ as $n \to \infty$, then also τ is a stopping time.

(ii): If $\tau_n \downarrow \tau$ as $n \to \infty$, then also τ is a stopping time.

<u>Proof</u> (i): $\{\tau \leq t\} = \bigcap_{n=1}^{\infty} \underbrace{\{\tau_n \leq t\}}_{\substack{\in \mathcal{F}_t \text{ since} \\ \tau_n \text{ is stopping time}}} \in \mathcal{F}_t.$

(ii): $\{\tau \leq t\} = \bigcap_{m=m_0}^{\infty} \bigcup_{n=1}^{\infty} \underbrace{\{\tau_n \leq t + \tfrac{1}{m}\}}_{\in \mathcal{F}_{t+1/m} \subset \mathcal{F}_{t+1/m_0}} \in \mathcal{F}_{t+1/m_0}.$

This holds for any m_0, so

$\{\tau \leq t\} \in \bigcap_{m_0=1}^{\infty} \mathcal{F}_{t+1/m_0} = \mathcal{F}_t$ by right-continuity. $\square$

Information available at stopping times

Def: Let τ be a stopping time w.r.t. filtration $\mathcal{F}_\bullet = (\mathcal{F}_t)_{t\in\mathbb{R}_+}$. We define

$$\mathcal{F}_\tau = \left\{ A \in \mathcal{F}_\infty \;\middle|\; \forall t \in \mathbb{R}_+ : \; A \cap \{\tau \le t\} \in \mathcal{F}_t \right\}$$

"at time t, if the stopping time τ has already occurred, then we can decide about the occurrence of the event A"

Exercise If τ is a deterministic stopping time $\tau(\omega) = t \quad \forall \omega \in \Omega$ then $\mathcal{F}_\tau = \mathcal{F}_t$.

(This is a sanity check, which shows that the notion is a reasonable generalization of "the information available at a given time".)

Lemma $\mathcal{F}_\tau$ is a σ-algebra.

Proof Let us verify the three defining properties of σ-algebras.

1°) $A = \Omega$?

We have $\Omega \cap \{\tau \le t\} = \{\tau \le t\} \in \mathcal{F}_t$ since τ is a stopping time.

This shows $\Omega \in \mathcal{F}_\tau$.

2°) Assume $A \in \mathcal{F}_\tau$. Check $A^c = \Omega \setminus A$?

By De Morgan's laws, $A^c \cup \{\tau > t\} = \big(\underbrace{A \cap \{\tau \le t\}}_{\in \mathcal{F}_t \text{ since } A \in \mathcal{F}_\tau}\big)^c \in \mathcal{F}_t$.

Then observe

$$A^c \cap \{\tau \le t\} = \underbrace{\left(A^c \cup \{\tau > t\}\right)}_{\in \mathcal{F}_t \text{ by the above}} \cap \underbrace{\{\tau \le t\}}_{\in \mathcal{F}_t \text{ since } \tau \text{ is stopping time}} \in \mathcal{F}_t.$$

This shows $A^c \in \mathcal{F}_\tau$.

3°) Assume $A_1, A_2, \ldots \in \mathcal{F}_\tau$. Check $\bigcup_{n=1}^{\infty} A_n$?

We have

$$\left(\bigcup_{n=1}^{\infty} A_n\right) \cap \{\tau \le t\} = \bigcup_{n=1}^{\infty} \underbrace{(A_n \cap \{\tau \le t\})}_{\in \mathcal{F}_t \text{ since } A_n \in \mathcal{F}_\tau} \in \mathcal{F}_t .$$

This shows $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}_\tau$. $\square$

<u>Lemma</u> Assume that σ and τ are stopping times such that $\sigma \le \tau$ $(\sigma(\omega) \le \tau(\omega) \ \forall \omega \in \Omega)$. Then we have $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$.

<u>Proof</u> Suppose $A \in \mathcal{F}_\sigma$.

Observe that $\{\tau \le t\} \subset \{\sigma \le t\}$ since $\sigma \le t$.

Therefore we can write

$$A \cap \{\tau \le t\} = \underbrace{A \cap \{\sigma \le t\}}_{\in \mathcal{F}_t \text{ since } A \in \mathcal{F}_\sigma} \cap \underbrace{\{\tau \le t\}}_{\in \mathcal{F}_t \text{ since } \tau \text{ is stopping time}} \in \mathcal{F}_t .$$

This shows $A \in \mathcal{F}_\tau$. $\square$

<u>Lemma</u> Suppose that $\tau_1, \tau_2, \ldots$ are stopping times such that $\tau_n \downarrow \tau$ as $n \to \infty$. Then

$$\mathcal{F}_\tau = \bigcap_{n=1}^{\infty} \mathcal{F}_{\tau_n} .$$

<u>Proof</u>: By previous lemma $\mathcal{F}_\tau \subset \bigcap_{n=1}^{\infty} \mathcal{F}_{\tau_n}$ since $\tau \le \tau_n \ \forall n$.

For converse inclusion assume $A \in \bigcap_{n=1}^{\infty} \mathcal{F}_{\tau_n}$. Then

$$A \cap \{\tau < t\} = A \cap \left(\bigcup_{n=1}^{\infty} \{\tau_n < t\}\right) = \bigcup_{n=1}^{\infty} \underbrace{(A \cap \{\tau_n < t\})}_{\in \mathcal{F}_t} \in \mathcal{F}_t .$$

For a right continuous filtration this shows $A \in \mathcal{F}_\tau$. $\square$

Proposition: Let τ be a stopping time and $T: \Omega \to \mathbb{R}_+ \cup \{+\infty\}$ a random variable which is $\mathcal{F}_\tau$-measurable and $T \geq \tau$. Then also T is a stopping time.

Proof: We have $\{T \leq t\} \subset \{\tau \leq t\}$ so we can write $\{T \leq t\} = \{T \leq t\} \cap \{\tau \leq t\}$. Since T is $\mathcal{F}_\tau$-measurable, the RHS $\in \mathcal{F}_t$, so $\{T \leq t\} \in \mathcal{F}_t$. $\square$

Notation: In the rest of this lecture, for $n \in \mathbb{N}$ and $t \geq 0$ denote $[t]_n = \inf\left(2^{-n} \cdot \mathbb{Z} \cap [t, \infty)\right)$ the smallest number of the form $k \cdot 2^{-n}$, $k \in \mathbb{Z}_{\geq 0}$, such that $t \leq k \cdot 2^{-n}$.

Note that for any $t \in \mathbb{R}_+$ we have $[t]_n \downarrow t$ as $n \to \infty$.

We also interpret $[+\infty]_n = +\infty$ for any $n \in \mathbb{N}$.

Corollary If τ is a stopping time, then for any $n \in \mathbb{N}$, also $[\tau]_n$ is a stopping time, with values in $2^{-n} \cdot \mathbb{Z}_{\geq 0}$. We have $[\tau_n] \downarrow \tau$.

Proof: Exercise. $\square$

If $X_\bullet = (X_t)_{t \in \mathbb{R}_+}$ is a $\mathbb{R}$-valued stochastic process, and τ is a finite stopping time (i.e. $\tau(\omega) < +\infty$ $\forall \omega \in \Omega$) then the value of the process at time τ is the random variable ⊛ X_τ defined by

$$\omega \longmapsto X_{\tau(\omega)}(\omega).$$

More generally, if τ is any stopping time, then we define $\mathbb{1}_{\{\tau < +\infty\}} X_\tau$ as

$$\omega \longmapsto \begin{cases} 0 & \text{if } \tau(\omega) = +\infty \\ X_{\tau(\omega)}(\omega) & \text{if } \tau(\omega) < \infty. \end{cases}$$

⊛ we have not yet verified measurability, though...

Lemma If $X_\bullet = (X_t)_{t\in\mathbb{R}_+}$ is a continuous process (for every ω, the path $t \mapsto X_t(\omega)$ is continuous) adapted to a right-continuous filtration $\mathcal{F}_\bullet = (\mathcal{F}_t)_{t\in\mathbb{R}_+}$ and τ is a stopping time w.r.t. $\mathcal{F}_\bullet$, then the random variable $\mathbb{1}_{\{\tau<\infty\}} X_\tau$ is $\mathcal{F}_\tau$-mble.

Proof: Use the approximation $[\tau]_n \downarrow \tau$ above.
By continuity of paths, we have

$$\mathbb{1}_{\{\tau<\infty\}} \cdot X_\tau = \lim_{n\to\infty} \mathbb{1}_{\{[\tau]_n<\infty\}} \cdot X_{[\tau]_n}$$

$$= \lim_{n\to\infty} \sum_{k=0}^{\infty} \mathbb{1}_{\{[\tau]_n = k2^{-n}\}} \cdot X_{k2^{-n}} .$$

Suppose then that $A \subset \mathbb{R}$ is Borel and $0 \notin A$.
Then for a stopping time σ and $s \in \mathbb{R}_+$ we have

$$\{\mathbb{1}_{\{\sigma=s\}} X_s \in A\} \cap \{\sigma \leq t\}$$

$$= \begin{cases} \emptyset & \text{if } t<s \\ \underbrace{\{X_s\in A\}}_{\in \mathcal{F}_s \subset \mathcal{F}_t} \cap \underbrace{\{\sigma=s\}}_{\in \mathcal{F}_s\subset\mathcal{F}_t} \cap \underbrace{\{\sigma\leq t\}}_{\in \mathcal{F}_t} & \text{if } t\geq s \end{cases}$$

We thus have $\{\mathbb{1}_{\{\sigma=s\}} X_s \in A\} \in \mathcal{F}_\sigma$.
If $0 \in A$ then the same is true, since

$$\{\mathbb{1}_{\{\sigma=s\}} X_s \in A\} = \{\mathbb{1}_{\{\sigma=s\}} X_s \in A^c\}^c \in \mathcal{F}_\sigma .$$

Therefore $\mathbb{1}_{\{\sigma=s\}} X_s$ is $\mathcal{F}_\sigma$-measurable.
Applying this to $\sigma = [\tau]_n$ and $s = k2^{-n}$ we see that terms $\mathbb{1}_{\{[\tau]_n = k2^{-n}\}} X_{k2^{-n}}$ are $\mathcal{F}_{[\tau]_n}$-mble, so also the sum

$$\sum_{k=0}^{\infty} \mathbb{1}_{\{[\tau]_n = k2^{-n}\}} X_{k2^{-n}} \quad \text{is } \mathcal{F}_{[\tau]_n}\text{-mble.}$$

For $n' \geq n$ we have $[\tau]_{n'} \leq [\tau]_n$ and $\mathcal{F}_{[\tau]_{n'}} \subseteq \mathcal{F}_{[\tau]_n}$
so $\mathbb{1}_{\{\tau<\infty\}} X_\tau$ is $\mathcal{F}_{[\tau]_n}$-mble for any $n \in \mathbb{N}$.
Since $\bigcap_n \mathcal{F}_{[\tau]_n} = \mathcal{F}_\tau$, this shows $\mathcal{F}_\tau$-measurability. $\square$

THE STRONG MARKOV PROPERTY OF BROWNIAN MOTION

Markov property states that whatever you know of the past of the Brownian motion at a given time does not give you better predictions about the future than the current value does. The strong Markov property extends this from a deterministically given time to a stopping time: only the value of the process at a stopping time is relevant for the continuation.

The precise statement is the following.

Theorem 1.14 (Strong Markov property of Brownian motion)
Let $B = (B_t)_{t \in \mathbb{R}_+}$ be a standard Brownian motion and let τ be an a.s. finite stopping time w.r.t. filtration $(\mathcal{F}_{t^+}^B)_{t \in \mathbb{R}_+}$. Define, for $t \geq 0$

$$\tilde{B}_t = (B_{\tau + t} - B_\tau) \mathbb{1}_{\{\tau < \infty\}}.$$

Then the process $\tilde{B} = (\tilde{B}_t)_{t \in \mathbb{R}_+}$ is a std BM independent of $\mathcal{F}_{\tau^+}^B$.

Proof Disregarding a zero probability event, suppose $\tau < \infty$. Let $A \in \mathcal{F}_{\tau^+}^B$ and consider $0 = t_0 < t_1 < t_2 < \dots < t_m$. Our goal is to show that for every bounded continuous $g : \mathbb{R}^m \to \mathbb{R}$ we have

$$\circledast : \quad \mathbb{E}\big[\mathbb{1}_A \, g(\tilde{B}_{t_1}, \dots, \tilde{B}_{t_m})\big] = \mathbb{P}[A] \, \mathbb{E}\big[g(B_{t_1}, \dots, B_{t_m})\big].$$

By taking $A = \Omega$ in $\circledast$ we see that $\tilde{B}$ has the same finite-dimensional distributions as B, and thus $\tilde{B}$ is a std BM.

Letting A vary over $\mathcal{F}_{\tau^+}^B$ in $\circledast$ shows that $\mathcal{F}_{\tau^+}^B$ is independent of $\tilde{B}$ (first the π-system of finite-dim. marginals of $\tilde{B}$ and consequently the entire σ-algebra generated by $\tilde{B}$).

To prove ⊛, note that by continuity of $t \mapsto B_t$ and of $g: \mathbb{R}^m \to \mathbb{R}$ we have

$$g(\tilde{B}_{t_1}, \dots, \tilde{B}_{t_k}) = \lim_{n\to\infty} g(B_{[\tau]_n+t_1} - B_{[\tau_n]-t_0}, \dots).$$

Since g is also bounded, we get

$$\mathbb{E}\left[\mathbb{1}_A\, g(\tilde{B}_{t_1}, \dots, \tilde{B}_{t_m})\right] = \lim_{n\to\infty} \mathbb{E}\left[\mathbb{1}_A \cdot g(B_{[\tau]_n+t_1} - B_{[\tau]_n}, \dots)\right]$$

$$= \lim_{n\to\infty} \sum_{k=0}^{\infty} \mathbb{E}\left[\mathbb{1}_{A\cap\{[\tau]_n = k2^{-n}\}}\, g(B_{k2^{-n}+t_1} - B_{k2^{-n}}, \dots)\right]$$

Observe that

$$A \cap \{[\tau]_n = k2^{-n}\} = A \cap \{(k-1)2^{-n} < \tau \le k2^{-n}\}$$

$$= \underbrace{A\cap\{\tau \le k2^{-n}\}}_{\in \mathcal{F}^B_{k2^{-n}}} \setminus \underbrace{(A\cap\{\tau \le (k-1)2^{-n}\})}_{\in \mathcal{F}^B_{(k-1)2^{-n}} \subset \mathcal{F}^B_{k2^{-n}}} \in \mathcal{F}^B_{k2^{-n}}.$$

Therefore by the simple Markov property we have

$$\mathbb{E}\left[\mathbb{1}_{A\cap\{[\tau]_n = k2^{-n}\}}\, g(B_{k2^{-n}+t_1} - B_{k2^{-n}}, \dots)\right]$$

$$= \mathbb{P}\left[A \cap \{[\tau]_n = k2^{-n}\}\right] \cdot \mathbb{E}\left[g(B_{t_1}, \dots, B_{t_m})\right].$$

Summing over k gives

$$\mathbb{E}\left[\mathbb{1}_A\, g(\tilde{B}_{t_1}, \dots, \tilde{B}_{t_m})\right]$$

$$= \lim_{n\to\infty} \underbrace{\left(\sum_{k=0}^{\infty} \mathbb{P}\left[A\cap\{[\tau]_n = k2^{-n}\}\right]\right)}_{= \mathbb{P}[A]} \cdot \mathbb{E}\left[g(B_{t_1}, \dots, B_{t_m})\right]$$

which establishes ⊛. This finishes the proof. □

As an example application of the strong Markov property, we determine the distribution of the maximum

$$S_t = \sup_{s \in [0,t]} B_s$$

of the Brownian motion up to time t. In fact, we derive the joint distribution of S_t and B_t.

Theorem (Reflection principle)
Let $0 < a$ and $b \leq a$ ant $t > 0$. Then for a standard Brownian motion we have

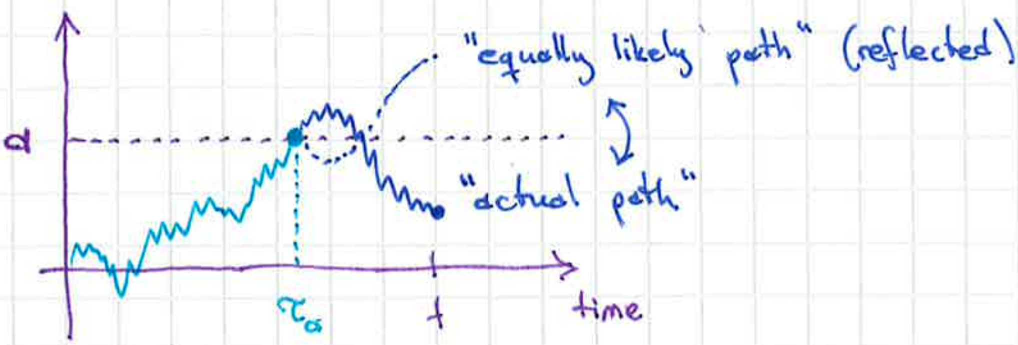
$$\mathbb{P}[S_t \geq a, B_t \leq b] = \frac{1}{\sqrt{2\pi t}} \int_{(2a-b)/\sqrt{t}}^{\infty} e^{-\frac{1}{2}u^2} du .$$

Remark The RHS is the complement of a cdf of a Gaussian, and equals

$$\mathbb{P}[B_t \geq 2a - b].$$

We prove the statement in this form.

The idea of the proof is to stop at the time τ_a when the BM first hits level $a > 0$ and then notice that going up or down after it happens with equal probability (by strong Markov prop.). Pictorially:



Proof Let $\tau_a = \inf \{ t \geq 0 \mid B_t = a \}$.
Then τ_a is a stopping time.
We have $S_t \geq a \iff \tau_a \leq t$ so we get

$$\mathbb{P}[S_t \geq a, B_t \leq b] = \mathbb{P}[\tau_a \leq t, B_t \leq b]$$

$$= \mathbb{P}[\tau_a \leq t, \tilde{B}_{t-\tau_a} \leq b-a]$$

where $\tilde{B}_s = B_{\tau_a + s} - B_{\tau_a}$, since on the event $\{\tau_a \leq t\}$ we have $B_{\tau_a} = a$.
By the strong Markov property, $\tilde{B}$ is a std BM independent of $\mathcal{F}^B_{\tau_a +}$. In particular the joint law of τ_a and $\tilde{B}$ is the same as the joint law of τ_a and $-\tilde{B}$ (both equal the product measure $P_{\tau_a} \otimes W_0$ of the law of τ_0 and the law of std BM).
Therefore we can write alternatively

$$\mathbb{P}[\tau_a \leq t, \tilde{B}_{t-\tau_a} \leq b-a]$$

$$= \mathbb{P}[\tau_a \leq t, \tilde{B}_{t-\tau_a} \geq a-b]$$

$$= \mathbb{P}[\tau_a \leq t, B_t \geq 2a-b]$$

$$= \mathbb{P}[B_t \geq 2a-b]$$

since $\{B_t \geq 2a-b\} \subset \{\tau_a \leq t\}$.
This finishes the proof. $\square$

Corollary: For any $a > 0$, we have

$$\mathbb{P}[S_t \geq a] = 2 \cdot \mathbb{P}[B_t \geq a] = \sqrt{\frac{2}{\pi t}} \cdot \int_{a/\sqrt{t}}^{\infty} e^{-\frac{1}{2}u^2} \, du$$

Proof: $$\mathbb{P}[S_t \geq a] = \underbrace{\mathbb{P}[S_t \geq a, B_t \leq a]}_{= \mathbb{P}[B_t \geq 2a - a]} + \underbrace{\mathbb{P}[S_t \geq a, B_t \geq a]}_{= \mathbb{P}[B_t \geq a]}$$

$$= 2 \cdot \mathbb{P}[B_t \geq a]. \qquad \square$$

Exercise: Show that both τ_a and $\left(\frac{a}{B_1}\right)^2$ have the same distribution, with c.d.f

$$F(t) = \mathbb{P}[\tau_a \leq t] = \mathbb{P}\left[\left(\frac{a}{B_1}\right)^2 \leq t\right]$$

given by

$$F(t) = \sqrt{\frac{2}{\pi}} \cdot \int_{a/\sqrt{t}}^{\infty} e^{-\frac{1}{2}u^2} \, du$$

and probability density given by

$$f(t) = \frac{a}{\sqrt{2\pi}} \cdot t^{-3/2} \cdot \exp\left(-\frac{1}{2t} a^2\right).$$

We now start to work towards

Stochastic integration

We want to make sense of

$$\int_0^t H_s \, dX_s$$

for suitable stochastic processes $H = (H_t)_{t \in \mathbb{R}_+}$ and $X = (X_t)_{t \in \mathbb{R}_+}$, in analogy with the discrete time "stochastic integral"

$$(H \cdot X)_n = \sum_{k=1}^{n} H_k \cdot (X_k - X_{k-1}) .$$

Recall that such an integral has for instance the interpretation of an investors cumulative profit up to time n, if H_k is the number of stocks in her portfolio at time k, and X_k is the unit stock price at time k. The continuous stochastic integral has a similar interpretation.

The theory of stochastic integration also offers powerful tools for calculations with martingales, especially the celebrated Itô's formula. In combination with optional stopping theorems, these tools yield concrete results that would be hard if not impossible to derive otherwise.

The classes of stochastic processes used:

integrand $H = (H_t)_{t \in \mathbb{R}_+}$	integrator $X = (X_t)_{t \in \mathbb{R}_+}$
▶ predictable process	▶ finite variation process OR ▶ continuous local martingale

FINITE VARIATION PROCESSES AND INTEGRALS

Let us first look at an instructive case where integrals are relatively easy to define.

Def: A function $\alpha : \mathbb{R}_+ \to \mathbb{R}$ (denoted $t \mapsto \alpha_t$) is said to be of finite variation if

- $\alpha_0 = 0$
- $t \mapsto \alpha_t$ is continuous $\mathbb{R}_+ \to \mathbb{R}$
- $v(t) := \lim_{n\to\infty} \sum_{k=1}^{\lfloor 2^n t \rfloor} |\alpha_{k2^{-n}} - \alpha_{(k-1)2^{-n}}| < \infty$

for all $t \in \mathbb{R}_+$.

$v(t)$ is called the total variation of α on $[0,t]$.

Lemma 2.1 For any continuous function $\alpha : \mathbb{R}_+ \to \mathbb{R}$ set
$$v^n(t) = \sum_{k=1}^{\lfloor 2^n t \rfloor} |\alpha_{k2^{-n}} - \alpha_{(k-1)2^{-n}}| .$$
Then the limit $v(t) = \lim_{n\to\infty} v^n(t)$ exists and is an increasing function of $t \in \mathbb{R}_+$.

Proof: It is clear from the definition of v^n that for $0 \le s < t$ we have $v^n(s) \le v^n(t)$. Thus if the limit $v = \lim_{n\to\infty} v^n$ exists, it is also increasing, $v(s) \le v(t)$ for $s<t$.

To prove that the limit exists, note that for fixed $t \in \mathbb{R}_+$ the sequence $(v^n(t))_{n\in\mathbb{N}}$ is increasing. Indeed, the dyadic intervals are split to two in going from n to $n+1$ and triangle inequality gives
$$|\alpha_{k2^{-n}} - \alpha_{(k-1)2^{-n}}| \le |\alpha_{(2k-1)2^{-n-1}} - \alpha_{(2k-2)2^{-n-1}}| + |\alpha_{2k2^{-n-1}} - \alpha_{(2k-1)2^{-n-1}}| .$$

There may also be one non-negative term in $v^{n+1}(t)$ in addition to the above. $\square$

Lemma For any finite variation function $\alpha: \mathbb{R}_+ \to \mathbb{R}$ and any $0 \le s < t$, we have $|\alpha_t - \alpha_s| \le v(t) - v(s)$.

Proof: Denote by $s_n^- = 2^{-n} \cdot \lfloor 2^n s \rfloor$ and $t_n^+ = 2^{-n} \cdot \lceil 2^n t \rceil$. Then by triangle inequality

$$|\alpha_{t_n^+} - \alpha_{s_n^-}| \le \sum_{k=2^n s_n^- + 1}^{2^n t_n^+} |\alpha_{k2^{-n}} - \alpha_{(k-1)2^{-n}}|$$

$$= v^n(t_n^+) - v^n(s_n^-) \le v^n(t) - v^n(s).$$

As $n \to \infty$ the LHS tends to $|\alpha_t - \alpha_s|$ by continuity and RHS tends to $v(t) - v(s)$ $\square$

Corollary For any $0 \le s = t_0 < t_1 < \dots < t_m = t$, we have

$$\sum_{j=1}^{m} |\alpha_{t_j} - \alpha_{t_{j-1}}| \le v(t) - v(s).$$

Here is a characterization of finite variation functions which we will use to see that integration with respect to them is easy.

Proposition 2.2: A continuous function $\alpha: \mathbb{R}_+ \to \mathbb{R}$ with $\alpha_0 = 0$ is of finite variation if and only if it is the difference $\alpha = \tilde{\alpha}^+ - \tilde{\alpha}^-$ of two continuous increasing functions $\tilde{\alpha}^+, \tilde{\alpha}^- : \mathbb{R}_+ \to \mathbb{R}$ with $\tilde{\alpha}_0^+ = 0$, $\tilde{\alpha}_0^- = 0$.

Proof: "if": Suppose $\alpha = \tilde{\alpha}^+ - \tilde{\alpha}^-$ as above.

Then by triangle inequality

$$v^n(t) = \sum_{k=1}^{\lfloor 2^n t \rfloor} \underbrace{|\alpha_{k2^{-n}} - \alpha_{(k-1)2^{-n}}|}_{\le |\tilde{\alpha}^+_{k2^{-n}} - \tilde{\alpha}^+_{(k-1)2^{-n}}| + |\tilde{\alpha}^-_{k2^{-n}} - \tilde{\alpha}^-_{(k-1)2^{-n}}|}$$

$$\le \sum_{k=1}^{\lfloor 2^n t \rfloor} (\tilde{\alpha}^+_{k2^{-n}} - \tilde{\alpha}^+_{(k-1)2^{-n}}) + \sum_{k=1}^{\lfloor 2^n t \rfloor} (\tilde{\alpha}^-_{k2^{-n}} - \tilde{\alpha}^-_{(k-1)2^{-n}})$$

$$= \tilde{\alpha}^+_{2^{-n}\cdot\lfloor 2^n t\rfloor} - \tilde{\alpha}^+_0 + \tilde{\alpha}^-_{2^{-n}\cdot\lfloor 2^n t\rfloor} - \tilde{\alpha}^-_0$$

since $\tilde{\alpha}^+$ and $\tilde{\alpha}^-$ are increasing.
Letting $n \to \infty$ we get

$$v(t) \leq \tilde{\alpha}^+_t + \tilde{\alpha}^-_t < \infty .$$

This shows that α is of finite variation.

"only if": Suppose α is of finite variation and let $v(t)$ denote its total variation on $[0,t]$. Define, for $t \in \mathbb{R}_+$,

$$\alpha^+_t = \frac{1}{2}\left(v(t) + \alpha_t\right)$$

$$\alpha^-_t = \frac{1}{2}\left(v(t) - \alpha_t\right).$$

Clearly $\alpha_t = \alpha^+_t - \alpha^-_t$ for any $t \in \mathbb{R}_+$.
Also for $s \leq t$ we have

$$\alpha^+_t - \alpha^+_s = \frac{1}{2}\left(v(t) + \alpha_t\right) - \frac{1}{2}\left(v(s) + \alpha_s\right)$$

$$= \frac{1}{2}\left(v(t) - v(s)\right) - \frac{1}{2}\left(\alpha_t - \alpha_s\right) \geq 0$$

since $|\alpha_t - \alpha_s| \leq v(t) - v(s)$. This shows that α^+ is increasing. Similar calculation shows that α^- is increasing.

Clearly $\alpha^+_0 = 0$ and $\alpha^-_0 = 0$. In order to finish the proof, it suffices to show that $v : \mathbb{R}_+ \to \mathbb{R}$ is continuous — the continuity of α^+ and α^- then follows.

The proof of continuity of v is left as an exercise.

$\square$

The decomposition $\alpha = \alpha^+ - \alpha^-$ given in the "only if" part is minimal in the sense that for any other decomposition $\alpha = \tilde{\alpha}^+ - \tilde{\alpha}^-$ we have:

<u>Exercise</u>: $\alpha^+_t \leq \tilde{\alpha}^+_t$ and $\alpha^-_t \leq \tilde{\alpha}^-_t$ $\forall t \in \mathbb{R}_+$.

Note that increasing functions are essentially cumulative distribution functions of measures. More specifically, if $\alpha : \mathbb{R}_+ \to \mathbb{R}$ is of finite variation and $\alpha^+, \alpha^- : \mathbb{R}_+ \to \mathbb{R}$ are the increasing functions above, then there exist two Borel measures μ^+, μ^- on $(0,\infty)$ determined by the conditions

$$\left.\begin{aligned} \mu^+[(s,t]] &= \alpha^+(t) - \alpha^+(s) \\ \mu^-[(s,t]] &= \alpha^-(t) - \alpha^-(s) \end{aligned}\right\} \quad \forall\ 0 \leq s < t.$$

Therefore the integration w.r.t. α is defined naturally as

$$\int_0^\infty f(s) \cdot d\alpha_s := \int_{(0,\infty)} f(s)\, d\mu^+(s) - \int_{(0,\infty)} f(s)\, d\mu^-(s)$$

whenever $f \in \mathcal{L}^1(\mu^+) \cap \mathcal{L}^1(\mu^-)$.

Integration up to time t is defined as usual by inserting the indicator function

$$\int_0^t f(s)\, d\alpha_s := \int_0^\infty \mathbb{1}_{(0,t]}(s) \cdot f(s)\, d\alpha_s$$

$$= \int_{(0,\infty)} \mathbb{1}_{(0,t]}(s) f(s)\, d\mu^+(s) - \int_{(0,\infty)} \mathbb{1}_{(0,t]}(s) f(s)\, d\mu^-(s).$$

We will also use the notation

$$\int_0^\infty f(s)\cdot|da_s| := \int_{(0,\infty)} f(s)\, d\mu^+(s) + \int_{(0,\infty)} f(s)\, d\mu^-(s)$$

and $\int_0^t f(s)\cdot|da_s|$ defined similarly. The following lemma is an easy consequence of these definitions.

__Lemma__: We have $\left|\int_0^t f(s)\, da_s\right| \leq \int_0^t |f(s)|\cdot|da_s|$.

Another useful observation is:

__Lemma__: The total variation $v(t)$ of a on $[0,t]$ is given by $v(t) = \int_0^t |da_s|$.

__Proof__:

$$\int_0^t |da_s| = \int_{(0,\infty)} \mathbb{1}_{(0,t]}(s)\, d\mu^+(s) + \int_{(0,\infty)} \mathbb{1}_{(0,t]}(s)\, d\mu^-(s)$$

$$= \mu^+[(0,t]] + \mu^-[(0,t]]$$

$$= (a_t^+ - a_0^+) + (a_t^- - a_0^-) = a_t^+ + a_t^-$$

$$= \tfrac{1}{2}(v(t) + a(t)) + \tfrac{1}{2}(v(t) - a(t))$$

$$= v(t). \qquad \square$$

For stochastic processes, we make the following definition. An underlying filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is considered fixed.

__Def__: $A = (A_t)_{t\in\mathbb{R}_+}$ is a finite variation process if A is adapted and for every $\omega\in\Omega$ the function $t \mapsto A_t(\omega)$ is a finite variation function. If moreover $t\mapsto A_t(\omega)$ is increasing for all $\omega\in\Omega$ then A is said to be an increasing process.

By our conventions in particular $A_0 = 0$ and $t \mapsto A_t$ is continuous in the above cases.

CONTINUOUS LOCAL MARTINGALES

As we saw, defining integrals with respect to finite variation processes is basically straightforward. The complementary class of stochastic processes with respect to which we will develop stochastic integration theory is continuous local martingales.

These are defined via stopped processes, so let us start with related basic results.

Throughout we consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ fixed. Also a filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is considered fixed. We assume:

- $\mathbb{P}$-completeness: If $N \in \mathcal{F}$ is such that

  $\mathbb{P}[N] = 0$ then each subset $E \subset N$ is measurable already at time zero

  $$E \in \mathcal{F}_0 \subset \mathcal{F}_t \subset \underbrace{\mathcal{F}_\infty}_{=\sigma\left(\bigcup_{t\in\mathbb{R}_+} \mathcal{F}_t\right)} \subset \mathcal{F} \qquad (\text{for } t \in \mathbb{R}_+).$$

- right continuity of filtration:

  $$\mathcal{F}_t = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} \qquad \forall\, t \in \mathbb{R}_+$$

These requirements are known as the *usual conditions*.

A fundamentally important use of stopping times is via optional stopping theorems. We have not yet proven continuous time versions of them, but they are derived from discrete time versions.

Def: If $X = (X_t)_{t \in \mathbb{R}_+}$ is a stochastic process and $\tau : \Omega \to \mathbb{R}_+$ is a finite stopping time, then we define the *stopped process* $X^\tau = (X^\tau_t)_{t \in \mathbb{R}_+}$ by

$$X^\tau_t := X_{t \wedge \tau} \qquad \left(X^\tau_t(\omega) := X_{t \wedge \tau(\omega)}(\omega) \right)$$

From the results of last lecture we get:

Lemma If X is continuous and adapted, then also X^τ is continuous and adapted.

Proof: Continuity is clear. We saw that $X_{t\wedge\tau}$ is $\mathcal{F}_{t\wedge\tau}$-measurable and since $t \wedge \tau \leq t$ we have $\mathcal{F}_{t \wedge \tau} \subset \mathcal{F}_t$. This shows that $X^\tau_t = X_{t \wedge \tau}$ is $\mathcal{F}_t$-measurable. □

We give two formulations of optional stopping theorems in continuous time. The first one is a characterization of continuous martingales.

Theorem 2.7 (Characterization of continuous martingales by optional stopping)

Let $X = (X_t)_{t \in \mathbb{R}_+}$ be a continuous, adapted, integrable ($X_t \in \mathcal{L}^1(P)\ \forall t$) process. Then the following are equivalent:

(i): X is a martingale

(ii): X^τ is a martingale for all bounded stopping times τ.

(iii): For all bounded stopping times σ and τ, we have $E[X_\tau \mid \mathcal{F}_\sigma] = X_{\sigma \wedge \tau}$.

(iv): $E[X_\tau] = E[X_0]$ for all bounded stopping times τ.

The main drawback in this formulation is its restriction to only bounded stopping times. Some conditions have to be imposed in general, anyway. The following notion is useful:

Def: A collection $(X_j)_{j\in J}$ of $\mathbb{R}$-valued random variables $X_j : \Omega \to \mathbb{R}$ is *uniformly integrable* if
$$\lim_{c\to\infty} \sup_{j\in J} \mathbb{E}\left[|X_j| \cdot \mathbb{1}_{\{|X_j|\geq c\}}\right] = 0.$$

Exercise: If $|X_j| \leq Z \in \mathcal{L}^1(\mathbb{P})$ for all $j \in J$, then the collection $(X_j)_{j\in J}$ is uniformly integrable.

Def: A stochastic process $X = (X_t)_{t\in\mathbb{R}_+}$ is called *uniformly integrable* if the collection $(X_t)_{t\in\mathbb{R}_+}$ of its values is a unif. int. collection.

Theorem (Optional stopping for uniformly integrable martingales)

Suppose that $X = (X_t)_{t\in\mathbb{R}_+}$ is a continuous, uniformly integrable martingale. Then there exists an integrable random variable $X_\infty \in \mathcal{L}^1(\mathbb{P})$ such that $X_t \xrightarrow[t\to\infty]{\text{a.s.}} X_\infty$ and $X_t \xrightarrow[t\to\infty]{\mathcal{L}^1} X_\infty$.

Moreover, if σ and τ are two stopping times s.t. $\sigma \leq \tau$, then we have
$$X_\sigma = \mathbb{E}[X_\tau \mid \mathcal{F}_\sigma].$$
(and both X_σ and X_τ are in $\mathcal{L}^1(\mathbb{P})$)

In particular
$$X_\sigma = \mathbb{E}[X_\infty \mid \mathcal{F}_\sigma]$$
and
$$\mathbb{E}[X_\tau] = \mathbb{E}[X_0].$$

Let us also record one basic trick used repeatedly in manipulations with square integrable martingales.

<u>Lemma 2.14</u>: Let $M=(M_t)_{t\in\mathbb{R}_+}$ be a martingale and $0\leq s<t$ such that $\mathbb{E}[M_s^2]<\infty$ and $\mathbb{E}[M_t^2]<\infty$. Then we have

$$\mathbb{E}\big[M_t^2-M_s^2 \,\big|\, \mathcal{F}_s\big] = \mathbb{E}\big[(M_t-M_s)^2 \,\big|\, \mathcal{F}_s\big].$$

<u>Proof</u>: Expand the square on the right hand side:

$$\mathbb{E}\big[(M_t-M_s)^2 \,\big|\, \mathcal{F}_s\big] = \mathbb{E}\big[M_t^2 \,\big|\, \mathcal{F}_s\big] - 2\cdot\mathbb{E}\big[M_sM_t \,\big|\, \mathcal{F}_s\big] + \mathbb{E}\big[M_s^2 \,\big|\, \mathcal{F}_s\big].$$

But since M_s is $\mathcal{F}_s$-measurable, properties (v) and (i) of conditional expected values give

$$\cdots = \mathbb{E}\big[M_t^2 \,\big|\, \mathcal{F}_s\big] - 2\cdot M_s\cdot \underbrace{\mathbb{E}\big[M_t \,\big|\, \mathcal{F}_s\big]}_{=M_s \text{ by martingale prop.}} + M_s^2$$

$$= \mathbb{E}\big[M_t^2 \,\big|\, \mathcal{F}_s\big] - 2\cdot M_s^2 + M_s^2$$

$$= \mathbb{E}\big[M_t^2 \,\big|\, \mathcal{F}_s\big] - M_s^2 \;=\; \mathbb{E}\big[M_t^2-M_s^2 \,\big|\, \mathcal{F}_s\big]. \qquad \square$$

<u>Corollary</u> Suppose that $(M_t)_{t\in\mathbb{R}_+}$ is a square integrable martingale, $M_t\in\mathcal{L}^2(\mathbb{P})$ $\forall t\in\mathbb{R}_+$. Let $0\leq s=s_0<s_1<\cdots<s_m$. Then we have

$$\mathbb{E}\big[M_{s_m}^2-M_{s_0}^2 \,\big|\, \mathcal{F}_{s_0}\big] = \mathbb{E}\Big[\sum_{j=1}^m (M_{s_j}-M_{s_{j-1}})^2 \,\Big|\, \mathcal{F}_{s_0}\Big]$$

<u>Proof</u> By the tower property (iv) of conditional exp., the RHS is

$$\sum_{j=1}^m \mathbb{E}\big[(M_{s_j}-M_{s_{j-1}})^2 \,\big|\, \mathcal{F}_{s_0}\big] = \sum_{j=1}^m \mathbb{E}\Big[\mathbb{E}\big[(M_{s_j}-M_{s_{j-1}})^2 \,\big|\, \mathcal{F}_{s_{j-1}}\big] \,\Big|\, \mathcal{F}_{s_0}\Big]$$

$$\overset{\text{Lemma 2.14}}{=} \sum_{j=1}^m \mathbb{E}\Big[\mathbb{E}\big[M_{s_j}^2-M_{s_{j-1}}^2 \,\big|\, \mathcal{F}_{s_{j-1}}\big] \,\Big|\, \mathcal{F}_{s_0}\Big]$$

$$\overset{\text{(iv)}}{=} \sum_{j=1}^m \mathbb{E}\big[M_{s_j}^2-M_{s_{j-1}}^2 \,\big|\, \mathcal{F}_{s_0}\big] \overset{\text{telescopic cancellations}}{=} \mathbb{E}\big[M_{s_m}^2-M_{s_0}^2 \,\big|\, \mathcal{F}_{s_0}\big].$$

$\square$

Local martingales

Def. Let $N = (N_t)_{t\in\mathbb{R}_+}$ be a continuous adapted process. Set $M_t = N_t - N_0$ for all $t\in\mathbb{R}_+$. We say that N is a continuous local martingale if there exists a sequence $\tau_1, \tau_2, \dots$ of stopping times such that $\tau_n \uparrow +\infty$ as $n\to\infty$ and for each $n\in\mathbb{N}$ the process M^{τ_n} is a martingale.

Remarks:

- N_0 does not have to be integrable!
- Often it is sufficient to prove properties only for the case $N_0 = 0$ and then observe that adding the same $\mathcal{F}_0$-measurable random variable to all values does not change validity.
- If N is a local mgale and τ is a stopping time, then N^τ is also a local mgale.
- If N and N' are local mgales, then also $c_1 N + c_2 N'$ is a local mgale for any $c_1, c_2 \in \mathbb{R}$. (As a sequence of stopping times take $\tau_n \wedge \tau'_n$.)

Terminology: We say that a sequence $\tau_1, \tau_2, \dots$ of stopping times reduces a local mgale $(M_t)_{t\in\mathbb{R}_+}$ if $\tau_n \uparrow +\infty$ and M^{τ_n} is a martingale for every $n\in\mathbb{N}$.

Proposition:

(i): If $N = (N_t)_{t\in\mathbb{R}_+}$ is a continuous local martingale s.t. $N_t \geq 0$ $\forall t\in\mathbb{R}_+$ and $N_0 \in \mathcal{L}^1(\mathbb{P})$, then N is a supermartingale.

(ii) If $N = (N_t)_{t\in\mathbb{R}_+}$ is a cont. loc. mgale s.t. for some $Z \in \mathcal{L}^1(\mathbb{P})$ we have $|N_t| \leq Z$ $\forall t\in\mathbb{R}_+$, then N is a uniformly integrable martingale.

Proof (i): Exercise.

(ii): Write $N_t = N_0 + M_t$. By definition, there exists a sequence $\tau_1, \tau_2, \dots$ of stopping times which reduces M. Then for any $0 \leq s < t$ apply optional stopping to the mgale M^{τ_n}:

$$M_{s\wedge\tau_n} = \mathbb{E}[M_{t\wedge\tau_n} \mid \mathcal{F}_s].$$

Add to both sides N_0, which by assumption is in $\mathcal{L}^1(\mathbb{P})$ (since $|N_0| \leq Z$) and is $\mathcal{F}_0$-measurable — we get

$$N_{s\wedge\tau_n} = \mathbb{E}[N_{t\wedge\tau_n} \mid \mathcal{F}_s].$$

As $n\to\infty$ we have $N_{s\wedge\tau_n} \to N_s$ and $N_{t\wedge\tau_n} \to N_t$ (since $\tau_n \uparrow +\infty$). All terms are dominated by $Z \in \mathcal{L}^1(\mathbb{P})$ so by DCT

$$N_s = \lim_{n\to\infty} N_{s\wedge\tau_n} = \lim_{n\to\infty} \mathbb{E}[N_{t\wedge\tau_n} \mid \mathcal{F}_s] \overset{DCT}{=} \mathbb{E}[N_t \mid \mathcal{F}_s].$$

Also $N_t \in \mathcal{L}^1(\mathbb{P})$ $\forall t$ and $(N_t)_{t\in\mathbb{R}_+}$ is uniformly integrable by the domination.

□

Let us finish the current discussion of local martingales by observing that this class of integrator processes is totally complementary to the other class we discussed, namely finite variation processes — the intersection of the classes is trivial.

Theorem 2.13 (Continuous local mgales of finite variation are zero)

Suppose that $M = (M_t)_{t \in \mathbb{R}_+}$ is a continuous local martingale and also a process of finite variation (in particular $M_0 = 0$). Then M is indistinguishable from the zero process,

$$\mathbb{P}\left[M_t = 0 \quad \forall t \in \mathbb{R}_+ \right] = 1.$$

Proof Define, for $n \in \mathbb{N}$,

the total variation of M up to time t

$$\tau_n = \inf \left\{ t \geq 0 \;\middle|\; \int_0^t |dM_s| \geq n \right\}.$$

This is a stopping time — it is the hitting time of the closed set $[n, \infty)$ by the continuous process $t \mapsto \int_0^t |dM_s|$.

Also $\tau_n \uparrow +\infty$ as $n \to \infty$ by the finite variation assumption $\int_0^t |dM_s| < \infty \quad \forall t \in \mathbb{R}_+$.

Since $M_0 = 0$, for any $t \in \mathbb{R}_+$ we get

$$|M_t^{\tau_n}| = |M_{t \wedge \tau_n}| \leq \int_0^{t \wedge \tau_n} |dM_s| \leq n.$$

This shows that M^{τ_n} is a bounded process. A bounded local martingale is a martingale, so M^{τ_n} is a martingale. Boundedness also implies square integrability, so we can use the basic trick for square integrable martingales as follows.

Fix $t \in \mathbb{R}_+$. For any subdivision $0 = t_0 < t_1 < \dots < t_m = t$ the basic trick gives

$$\mathbb{E}\left[(M_t^{\tau_n})^2\right] = \mathbb{E}\left[(M_t^{\tau_n})^2 - (M_0^{\tau_n})^2\right]$$

$$= \mathbb{E}\left[\sum_{j=1}^m (M_{t_j}^{\tau_n} - M_{t_{j-1}}^{\tau_n})^2\right]$$

$$\leq \mathbb{E}\left[\sup_{j=1,\dots,m} |M_{t_j}^{\tau_n} - M_{t_{j-1}}^{\tau_n}| \cdot \underbrace{\sum_{j=1}^m |M_{t_j}^{\tau_n} - M_{t_{j-1}}^{\tau_n}|}_{\leq \int_0^{t\wedge\tau_n} |dM_s| \leq n}\right]$$

$$\leq \mathbb{E}\left[\sup_{j=1,\dots,m} |M_{t_j}^{\tau_n} - M_{t_{j-1}}^{\tau_n}|\right].$$

By choosing a sequence of subdivisions $0 = t_0^{(\ell)} < t_1^{(\ell)} < \dots < t_{m_\ell}^{(\ell)} = t$, for $\ell = 1, 2, \dots$, with mesh tending to zero, we have by continuity of $t \mapsto M_t^{\tau_n}$ that

$$\sup_{j=1,\dots,m_\ell} |M_{t_j^{(\ell)}}^{\tau_n} - M_{t_{j-1}^{(\ell)}}^{\tau_n}| \xrightarrow[\ell\to\infty]{} 0.$$

Moreover, this sup is bounded by $2 \cdot n$ (since $|M_t^{\tau_n}| \leq n \ \forall t$) so bounded convergence theorem gives

$$\mathbb{E}\left[(M_t^{\tau_n})^2\right] \leq \mathbb{E}\left[\sup_{j=1,\dots,m_\ell} |M_{t_j^{(\ell)}}^{\tau_n} - M_{t_{j-1}^{(\ell)}}^{\tau_n}|\right] \xrightarrow[\ell\to\infty]{} 0.$$

This shows that $M_t^{\tau_n} = 0$ a.s. for any $t \in \mathbb{R}_+$ and any $n \in \mathbb{N}$. Taking the limit $n \to \infty$ we get $M_t = 0$ a.s. for every $t \in \mathbb{R}_+$. Then also (by union bound)

$$\mathbb{P}\left[M_q = 0 \ \forall q \in \mathbb{Q}_+\right] \geq 1 - \sum_{q \in \mathbb{Q}_+} \mathbb{P}\left[M_q \neq 0\right] = 1 - 0 = 1.$$

Finally by continuity this shows

$$\mathbb{P}\left[M_t = 0 \ \forall t \in \mathbb{R}_+\right] = 1.$$

$\square$

THE SPACES OF CONTINUOUS MARTINGALES

In what follows, the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is considered fixed, as well as the filtration $\mathcal{F}_\bullet = (\mathcal{F}_t)_{t\in\mathbb{R}_+}$.

We assume that these satisfy the usual conditions:

- $\mathcal{F}_0$ is $\mathbb{P}$-complete:
  for any $N \subset A \subset \Omega$ s.t. $A \in \mathcal{F}$ and $\mathbb{P}[A]=0$ we have also $N \in \mathcal{F}_0$
- $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is right-continuous:
  for any $t\in\mathbb{R}_+$ we have $\mathcal{F}_t = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}$.

Let us denote the space of continuous martingales by

$$\mathcal{M}_c := \left\{ (M_t)_{t\in\mathbb{R}_+} \text{ continuous martingale w.r.t } \mathcal{F}_\bullet \right\}.$$

To be explicit, $M \in \mathcal{M}_c$ means

- M is adapted to $\mathcal{F}_\bullet$:
  $\forall t\in\mathbb{R}_+ \; : \; M_t$ is $\mathcal{F}_t$-measurable
- M is integrable:
  $\forall t\in\mathbb{R}_+ \; : \; M_t \in \mathcal{L}^1(\mathbb{P})$, i.e., $\mathbb{E}\left[|M_t|\right] < \infty$.
- M has the martingale property:
  $\forall\, 0 \leq s < t \; : \; \mathbb{E}\left[M_t \mid \mathcal{F}_s\right] = M_s$
- M has (a.s.) continuous paths:
  the path $t \mapsto M_t(\omega)$ is continuous $\mathbb{R}_+ \to \mathbb{R}$ for $\mathbb{P}$-almost all $\omega\in\Omega$

We will use a particularly well-behaved subspace — martingales which are bounded in $\mathcal{L}^2(\mathbb{P})$:

$$\mathcal{M}_c^2 := \left\{ M \in \mathcal{M}_c \;\middle|\; \sup_{t \in \mathbb{R}_+} \mathbb{E}[M_t^2] < \infty \right\}.$$

This space has a very useful norm, and it is complete w.r.t. that norm (in fact $\mathcal{M}_c^2$ is a Hilbert space). We just need a few tools...

Facts from martingale theory

We use the following two fundamental facts of martingale theory, whose proofs are postponed for the moment.

Theorem 3.1 (Martingale convergence theorem for $\mathcal{M}_c^2$).

Suppose that $M \in \mathcal{M}_c^2$. Then there exists a square integrable random variable M_∞ s.t.

$$M_t \xrightarrow[t \to \infty]{} M_\infty \quad \text{almost surely.}$$

Moreover, $M_t \xrightarrow[t\to\infty]{\mathcal{L}^2} M_\infty$ (i.e. $\mathbb{E}[(M_t - M_\infty)^2] \to 0$)

and

$$M_t = \mathbb{E}[M_\infty \mid \mathcal{F}_t] \qquad \forall t \in \mathbb{R}_+.$$

Theorem 3.2 (Doob's L^2-inequality)

For any $M \in \mathcal{M}_c^2$ we have

$$\mathbb{E}\left[\left(\sup_{t\in\mathbb{R}_+} |M_t|\right)^2\right] \leq 4 \cdot \mathbb{E}[M_\infty^2].$$

Norms in the space of martingales bounded in $\mathcal{L}^2$

Observe first that both $\mathcal{M}_c$ and $\mathcal{M}_c^2$ are vector spaces:

Lemma If $M^{(1)}, M^{(2)} \in \mathcal{M}_c$ (resp. $M^{(1)}, M^{(2)} \in \mathcal{M}_c^2$) and $c_1, c_2 \in \mathbb{R}$ then also $c_1 M^{(1)} + c_2 M^{(2)} \in \mathcal{M}_c$ (resp. $c_1 M^{(1)} + c_2 M^{(2)} \in \mathcal{M}_c^2$).

Proof Easy exercise!

We then equip $\mathcal{M}_c^2$ with two different norms: for $M \in \mathcal{M}_c^2$ denote

$$\|M\|_{\mathcal{M}_c^2} := \sqrt{\mathbb{E}[M_\infty^2]} = \|M_\infty\|_{\mathcal{L}^2(\mathbb{P})}$$

and

$$|||M||| := \sqrt{\mathbb{E}\Big[\big(\sup_{t\geq 0} |M_t|\big)^2\Big]} = \Big\| \sup_{t\geq 0} |M_t| \Big\|_{\mathcal{L}^2(\mathbb{P})}.$$

Lemma For $M \in \mathcal{M}_c^2$ we have

$$\|M\|_{\mathcal{M}_c^2} \leq |||M||| \leq 2 \cdot \|M\|_{\mathcal{M}_c^2}$$

Proof: The inequality $|||M||| \leq 2 \cdot \|M\|_{\mathcal{M}_c^2}$ is just Doob's L^2-inequality after taking square roots.

To prove the first inequality, note that $\mathcal{L}^2$-convergence $M_t \xrightarrow[t\to\infty]{\mathcal{L}^2} M_\infty$ implies

$$\|M_\infty\|_{\mathcal{L}^2(\mathbb{P})} \leq \sup_{t\geq 0} \|M_t\|_{\mathcal{L}^2(\mathbb{P})} = \sup_{t\geq 0} \sqrt{\mathbb{E}[M_t^2]}$$

$$\leq \sqrt{\mathbb{E}\big[\sup_{t\geq 0} M_t^2\big]} = \sqrt{\mathbb{E}\big[(\sup_{t\geq 0} |M_t|)^2\big]} = |||M|||.$$

$\square$

Observe then that if $\|M\|_{\mathcal{M}_c^2} = 0$ then $\sup_{t\geq 0} |M_t| = 0$ almost surely, so M is indistinguishable from the zero process. We will work with processes "up to indistinguishability", i.e. we consider two indistinguishable processes equal.

(To elaborate, this means that if X and Y are two processes such that

$$\mathbb{P}[\, X_t = Y_t \quad \forall t \in \mathbb{R}_+ \,] = 1$$

then we write $X = Y$.)

<u>Proposition 3.4</u> (Completeness of $\mathcal{M}_c^2$)

The space $\mathcal{M}_c^2$ equipped with the norm $\|\cdot\|_{\mathcal{M}_c^2}$ is a Hilbert space.

<u>Proof</u>: Clearly the norm comes from an inner product, since it is just the $\mathcal{L}^2(\mathbb{P})$-norm of M_∞.

It only remains to show that the space $\mathcal{M}_c^2$ is complete.

Suppose $M^{(1)}, M^{(2)}, \ldots \in \mathcal{M}_c^2$ is a Cauchy sequence, i.e.

$$\lim_{r\to\infty} \sup_{n,m\geq r} \| M^{(n)} - M^{(m)} \|_{\mathcal{M}_c^2} = 0. \qquad \text{(CAUCHY)}$$

This literally means that $(M_\infty^{(n)})_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{L}^2(\mathbb{P})$ and thus

$$M_\infty^{(n)} \xrightarrow[n\to\infty]{\mathcal{L}^2} X \in \mathcal{L}^2(\mathbb{P})$$

The remaining task is to understand the corresponding process, i.e. the random function of time. For $t \in \mathbb{R}_+$, the value is

$$X_t = \mathbb{E}[X \mid \mathcal{F}_t].$$

Note that at least

$$M_t^{(n)} = \mathbb{E}\left[M_\infty^{(n)} \,|\, \mathcal{F}_t \right] \xrightarrow[n\to\infty]{\mathcal{L}^2} \mathbb{E}[X \,|\, \mathcal{F}_t] = X_t$$

since $M_\infty^{(n)} \to X$ in $\mathcal{L}^2$ and the conditional expected value $\mathbb{E}[\cdot \,|\, \mathcal{F}_t]$ is a projection to a subspace and therefore 1-Lipschitz and in particular continuous $\mathcal{L}^2(\mathbb{P}) \to \mathcal{L}^2(\mathbb{P})$.

This shows also that $(X_t)_{t\in\mathbb{R}_+}$ is bounded in $\mathcal{L}^2(\mathbb{P})$:

$$\| X_t \|_{\mathcal{L}^2(\mathbb{P})} = \| \mathbb{E}[X|\mathcal{F}_t] \|_{\mathcal{L}^2(\mathbb{P})} \leq \| X \|_{\mathcal{L}^2(\mathbb{P})} < \infty .$$

The martingale property is automatic also: for $0 \leq s < t$ by the tower property (iv) of conditional expectation we have

$$\begin{aligned} \mathbb{E}[X_t \,|\, \mathcal{F}_s] &= \mathbb{E}\big[\mathbb{E}[X|\mathcal{F}_t] \,\big|\, \mathcal{F}_s \big] \\ &= \mathbb{E}[X \,|\, \mathcal{F}_s] \qquad \text{since } \mathcal{F}_s \subset \mathcal{F}_t \\ &= X_s . \end{aligned}$$

It only remains to show that the limit process has continuous paths.

By (CAUCHY), choose a subsequence with indices $n_1 < n_2 < \ldots$ such that

$$\| M^{(n)} - M^{(m)} \|_{\mathcal{M}_c^2} < 2^{-j} \quad \text{whenever } n, m \geq n_j .$$

In particular

$$\| M^{(n_{j+1})} - M^{(n_j)} \|_{\mathcal{M}_c^2} < 2^{-j}$$

and so by the previous lemma also

$$|\!|\!| M^{(n_{j+1})} - M^{(n_j)} |\!|\!| < 2 \cdot 2^{-j} .$$

Therefore the following series converges

$$\sum_{j=1}^{\infty} \mathbb{E}\left[\sup_{t\geq 0} | M_t^{(n_{j+1})} - M_t^{(n_j)} | \right]$$

(The $\mathcal{L}^1$-norm is controlled by $\mathcal{L}^2$-norm)

$$\leq \sum_{j=1}^{\infty} \mathbb{E}\left[\left(\sup_{t\geq 0} | M_t^{(n_{j+1})} - M_t^{(n_j)} | \right)^2 \right]^{1/2} < 2 \cdot \sum_{j=1}^{\infty} 2^{-j} < \infty .$$

This shows by the usual reasoning (based on monotone conv.) that

$$\sum_{j=1}^{\infty} \sup_{t\geq 0} |M_t^{(n_{j+1})} - M_t^{(n_j)}| < \infty \quad \text{almost surely.}$$

As a consequence, almost surely the sequence of functions $t \longmapsto M_t^{(n_j)}$ indexed by $j=1,2,\dots$ converges uniformly: for any $\varepsilon > 0$ we can choose j_ε such that

$$\sum_{j=j_\varepsilon}^{\infty} \sup_{t\geq 0} |M_t^{(n_{j+1})} - M_t^{(n_j)}| < \varepsilon,$$

which implies for any $k > l \geq j_\varepsilon$

$$\sup_{t\geq 0} |M_t^{(n_k)} - M_t^{(n_l)}| \leq \sum_{j=l}^{k-1} \sup_{t\geq 0} |M_t^{(n_{j+1})} - M_t^{(n_j)}|$$

$$\leq \sum_{j=j_\varepsilon}^{\infty} \sup_{t\geq 0} |M_t^{(n_{j+1})} - M_t^{(n_j)}| < \varepsilon$$

so the sequence is Cauchy w.r.t. sup-norm.

The limit function $t \longmapsto M_t = \lim_{j\to\infty} M_t^{(n_j)}$ is therefore continuous, as the uniform limit of continuous functions.

There may be an exceptional event E of zero probability, on which we define for example $M_t(\omega) = 0$ (if $\omega \in E$) to have continuity anyway.

We anyway have $M_t^{(n_j)} \xrightarrow[j\to\infty]{} M_t$ for all $t \in \mathbb{R}_+$ almost surely. Of course this a.s. limit of subseq. coincides with the $\mathcal{L}^2(\mathbb{P})$ limit of the entire sequence $M_t^{(n)} \xrightarrow[n\to\infty]{} X_t$, i.e.

$$X_t = M_t \quad (\text{a.s.}), \quad \text{for any } t \in \mathbb{R}_+.$$

Therefore in particular $M = (M_t)_{t\in\mathbb{R}_+}$ is a continuous martingale bounded in $\mathcal{L}^2(\mathbb{P})$.

The value M_∞ is X (why?) so from $M^{(n)} \xrightarrow[n\to\infty]{\mathcal{L}^2} X$ we deduce $\| M^n - M \|_{\mathcal{M}_c^2} \longrightarrow 0$ as $n \to \infty$. $\square$

QUADRATIC VARIATION

Theorem 3.5 (Quadratic variation)

For every $M \in \mathcal{M}_{c,loc}$ there exists a unique (up to indistinguishability) increasing process $\langle M,M \rangle = (\langle M,M\rangle_t)_{t\in\mathbb{R}_+}$ such that $M^2 - \langle M,M\rangle \in \mathcal{M}_{c,loc}$. Moreover, for any $t \in \mathbb{R}_+$ we have

$$\sum_{k=1}^{\lfloor 2^n t \rfloor} \left(M_{k2^{-n}} - M_{(k-1)2^{-n}}\right)^2 \xrightarrow[n\to\infty]{\mathbb{P}} \langle M,M\rangle_t .$$

Terminology: $\langle M,M\rangle$ is called the quadratic variation process of M.

Example (Quadratic variation of Brownian motion)

Let $B = (B_t)_{t\in\mathbb{R}_+}$ be a standard Brownian motion. Then (exercise) for any $t \geq 0$

$$\sum_{k=1}^{\lfloor 2^n t \rfloor} \left(B_{k2^{-n}} - B_{(k-1)2^{-n}}\right)^2 \xrightarrow[n\to\infty]{a.s.} t .$$

Also (exercise) the process $(B_t^2 - t)_{t\in\mathbb{R}_+}$ is a martingale.

Either one of the above two facts (exercises) shows that the quadratic variation process of a standard Brownian motion is given by

$$\langle B,B\rangle_t = t \qquad \forall\, t\in\mathbb{R}_+ .$$

We will begin the proof of Thm 3.5 now, and finish it in the next lecture.

Proof of uniqueness in Thm 3.5:

Suppose that $A = (A_t)_{t\in\mathbb{R}_+}$ and $A' = (A'_t)_{t\in\mathbb{R}_+}$ are increasing processes such that $M^2 - A$ and $M^2 - A'$ are cont. loc. mgales.

Then $A - A'$ is a finite variation process.

But also

$$A - A' = \underbrace{(M^2 - A')}_{\in \mathcal{M}_{c,loc}} - \underbrace{(M^2 - A)}_{\in \mathcal{M}_{c,loc}} \in \mathcal{M}_{c,loc} .$$

Therefore by Thm 2.13 we have $A_t - A'_t = 0 \ \forall t$, a.s., i.e., $A = A'$ (up to indistinguishability). $\square$

Proof of existence in Thm 3.5 assuming M bounded:

Assume $|M_t| \leq C < \infty$ and $M_0 = 0$. Fix also $T > 0$.

For $n \in \mathbb{N}$ define the process $Y^{(n)} = (Y^{(n)}_t)_{t\in\mathbb{R}_+}$ by

$$Y^{(n)}_t = \sum_{k=1}^{\lfloor 2^n T \rfloor} M_{(k-1)2^{-n}} \cdot \left(M_{k2^{-n}\wedge t} - M_{(k-1)2^{-n}\wedge t} \right).$$

Each term is a bounded continuous mgale, so $Y^{(n)} \in \mathcal{M}_c$ and $Y^{(n)}$ is bounded, so $Y^{(n)} \in \mathcal{M}_c^2$.

In two lemmas below we will show that the sequence $Y^{(1)}, Y^{(2)}, \ldots$ is Cauchy in $\mathcal{M}_c^2$:

$$\lim_{r\to\infty} \sup_{n,m\geq r} \| Y^{(n)} - Y^{(m)} \|_{\mathcal{M}_c^2} = 0 .$$

By completeness there exists a $Y \in \mathcal{M}_c^2$ such that $Y^{(n)} \xrightarrow[n\to\infty]{\mathcal{M}_c^2} Y$.

We next address the process $(M_t^2 - Y_t)_{t\in\mathbb{R}_+}$.

Let us therefore look at $M_t^2 - Y_t^{(n)}$ for given $n \in \mathbb{N}$ (and in the end take $n \to \infty$).

For any $k = 0, 1, \dots, \lfloor 2^n T \rfloor$ we calculate

$$\left(M_{k2^{-n}}\right)^2 - 2 \cdot Y^{(n)}_{k \cdot 2^{-n}}$$

$$= \sum_{j=1}^{k} \left((M_{j2^{-n}})^2 - (M_{(j-1)2^{-n}})^2 \right) - 2 \cdot \sum_{j=1}^{k} M_{(j-1)2^{-n}} \left(M_{j \cdot 2^{-n}} - M_{(j-1)2^{-n}} \right)$$

$$= \sum_{j=1}^{k} \left(M_{j \cdot 2^{-n}} - M_{(j-1)2^{-n}} \right)^2$$

This expression shows that along the sequence of times of the form $k \cdot 2^{-n}$, the process $M_t^2 - Y_t^{(n)}$ is increasing. By taking $n \to \infty$ and using continuity, we get that the process $t \longmapsto M_t^2 - Y_t = A_t$ is increasing.

Thus this process A is increasing and $M^2 - A = Y \in \mathcal{M}_c^2$ is a continuous martingale.

This finishes⊛ the existence proof for bounded M. $\square$

⊛ Actually this was ok only on $t \in [0,T]$. But T is arbitrary and we can let $T \to \infty$. Uniqueness part shows that A becomes consistently defined.

We left out of the proof the check that $(Y^{(n)})_{n \in \mathbb{N}}$ is a Cauchy sequence. This is proven in the lemmas below.

<u>Lemma 3.6</u> Suppose that $M \in \mathcal{M}_c$ is bounded, $|M_t| \leq C$ for all $t \in \mathbb{R}_+$. Then for any $0 = t_0 < t_1 < \dots < t_m < \infty$ we have

$$\mathbb{E}\left[\left(\sum_{k=1}^{m} (M_{t_k} - M_{t_{k-1}})^2 \right)^2 \right] \leq 16 \cdot C^4 < \infty.$$

<u>Proof</u>: Note first that

$$\mathbb{E}\left[\left(\sum_{k=1}^{m} (M_{t_k} - M_{t_{k-1}})^2 \right)^2 \right]$$

$$= \sum_{k=1}^{m} \mathbb{E}\left[(M_{t_k} - M_{t_{k-1}})^4 \right] + 2 \cdot \sum_{k=1}^{m} \mathbb{E}\left[(M_{t_k} - M_{t_{k-1}})^2 \cdot \sum_{l=k+1}^{m} (M_{t_l} - M_{t_{l-1}})^2 \right].$$

Let us now apply the Corollary to Lemma 2.14 and $\underbrace{\text{properties of cond. exp.}}_{\text{(ii) and (v)}}$ in the second term:

$$\mathbb{E}\Big[(M_{t_k}-M_{t_{k-1}})^2 \cdot \sum_{\ell=k+1}^{m} (M_{t_\ell}-M_{t_{\ell-1}})^2\Big]$$

$$\overset{\text{(ii) and (v)}}{=} \mathbb{E}\Big[(M_{t_k}-M_{t_{k-1}})^2 \cdot \mathbb{E}\Big[\sum_{\ell=k+1}^{m} (M_{t_\ell}-M_{t_{\ell-1}})^2 \,\Big|\, \mathcal{F}_{t_k}\Big]\Big]$$

$$\overset{\text{Corollary}}{=} \mathbb{E}\Big[(M_{t_k}-M_{t_{k-1}})^2 \cdot \mathbb{E}\big[M_{t_m}^2 - M_{t_k}^2 \,\big|\, \mathcal{F}_{t_k}\big]\Big]$$

$$\overset{\text{(ii) and (v)}}{=} \mathbb{E}\Big[(M_{t_k}-M_{t_{k-1}})^2 \cdot (M_{t_m}^2 - M_{t_k}^2)\Big].$$

Inserting this into the first formula of the proof, we estimate

$$\mathbb{E}\Big[\Big(\sum_{k=1}^{m} (M_{t_k}-M_{t_{k-1}})^2\Big)^2\Big]$$

$$\leq \mathbb{E}\Big[\Big(\underbrace{\sup_j |M_{t_j}-M_{t_{j-1}}|^2}_{\leq 4C^2} + 2\cdot \sup_j \underbrace{|M_{t_m}^2 - M_{t_j}^2|}_{\leq 2C^2}\Big) \sum_{k=1}^{m} (M_{t_k}-M_{t_{k-1}})^2\Big]$$

$$\leq 8\cdot C^2 \cdot \mathbb{E}\Big[\sum_{k=1}^{m} (M_{t_k}-M_{t_{k-1}})^2\Big]$$

$$\overset{\text{Corollary}}{=} 8\cdot C^2 \cdot \mathbb{E}\big[\underbrace{M_{t_m}^2 - M_{t_0}^2}_{\leq 2\cdot C^2}\big] \leq 16\cdot C^4. \qquad \square$$

Finally, the Cauchy-property is verified in the following.

Lemma Suppose that $M \in \mathcal{M}_c$ is bounded, $|M_t| \leq C$ for all $t \in \mathbb{R}_+$. Define

$$Y_t^{(n)} = \sum_{k=1}^{\lfloor 2^n T \rfloor} M_{(k-1)2^{-n}} \left(M_{k2^{-n}\wedge t} - M_{(k-1)2^{-n}\wedge t} \right)$$

as in the proof of Thm 3.5. Then $(Y^{(n)})_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{M}_c^2$.

Proof: By construction we have $Y^{(n)}_\infty = Y^{(n)}_T$.
Therefore to understand $\|Y^{(n)} - Y^{(m)}\|_{\mathcal{M}^2_c} = \|Y^{(n)}_\infty - Y^{(m)}_\infty\|_{\mathcal{L}^2(\mathbb{P})}$
we need to calculate

$$\mathbb{E}\left[\left(Y^{(n)}_T - Y^{(m)}_T\right)^2\right] = \mathbb{E}\left[(Y^{(n)}_T)^2\right] - 2\,\mathbb{E}\left[Y^{(n)}_T Y^{(m)}_T\right] + \mathbb{E}\left[(Y^{(m)}_T)^2\right].$$

Let us fix $n \le m$ and compute the cross term:

$$\mathbb{E}\left[Y^{(n)}_T\, Y^{(m)}_T\right]$$

$$= \sum_{k=1}^{\lfloor 2^n T \rfloor} \sum_{l=1}^{\lfloor 2^m T \rfloor} \mathbb{E}\left[M_{(k-1)2^{-n}}\left(M_{k2^{-n}} - M_{(k-1)2^{-n}}\right) \cdot M_{l2^{-m}}\left(M_{l2^{-m}} - M_{(l-1)2^{-m}}\right)\right].$$

We claim that only very particular terms in this sum are non-zero.

Indeed, the dyadic intervals are either disjoint or one contained in the other – specifically

either $[(l-1)2^{-m}, l\cdot 2^{-m}) \subset [(k-1)2^{-n}, k\cdot 2^{-n})$

or $[(l-1)2^{-m}, l\cdot 2^{-m}) \cap [(k-1)2^{-n}, k\cdot 2^{-n}) = \emptyset$.

In the second case we use the following calculation for $0 \le t_1 < t_2 \le t_3 < t_4$:

$$\mathbb{E}\left[M_{t_1}(M_{t_2} - M_{t_1})\, M_{t_3}(M_{t_4} - M_{t_3})\right]$$

(ii) and (v)

$$= \mathbb{E}\Big[M_{t_1}(M_{t_2} - M_{t_1})\, M_{t_3} \cdot \underbrace{\mathbb{E}\left[M_{t_4} - M_{t_3} \mid \mathcal{F}_{t_3}\right]}_{=0 \text{ by martingale property}}\Big] = 0$$

to see that the corresponding term vanishes.

In the first case we use the following calculation for $0 \le t_1 \le s_1 < s_2 \le t_2$: by the same argument as above

$$\mathbb{E}\Big[M_{t_1}\underbrace{(M_{t_2} - M_{t_1})}_{= (M_{t_2} - M_{s_2}) + (M_{s_2} - M_{s_1}) + (M_{s_1} - M_{t_1})} \cdot M_{s_1} \cdot (M_{s_2} - M_{s_1})\Big]$$

$$= 0 + \mathbb{E}\left[M_{t_1} M_{s_1} (M_{s_2} - M_{s_1})^2\right] + 0$$

$$= \mathbb{E}\left[M_{t_1} M_{s_1} (M_{s_2} - M_{s_1})^2\right].$$

For a given ℓ, there is exactly one value of k such that the dyadic intervals are nested. If this value is denoted by $k = k_{n,m}(\ell)$, then the above considerations show

$$\mathbb{E}\left[Y_T^{(n)} Y_T^{(m)}\right]$$

$$= \sum_{\ell=1}^{\lfloor 2^m T \rfloor} \mathbb{E}\left[M_{(k_{n,m}(\ell)-1)2^{-n}} \cdot M_{(\ell-1)2^{-m}} \cdot \left(M_{\ell 2^{-m}} - M_{(\ell-1)2^{-m}}\right)^2\right]$$

As a special case of this, for $n = m$, we get

$$\mathbb{E}\left[\left(Y_T^{(m)}\right)^2\right] = \sum_{\ell=1}^{\lfloor 2^m T \rfloor} \mathbb{E}\left[\left(M_{(\ell-1)2^{-m}}\right)^2 \cdot \left(M_{\ell 2^{-m}} - M_{(\ell-1)2^{-m}}\right)^2\right].$$

As the third and last such case, we have

$$\mathbb{E}\left[\left(Y_T^{(n)}\right)^2\right] = \sum_{k=1}^{\lfloor 2^n T \rfloor} \mathbb{E}\left[\left(M_{(k-1)2^{-n}}\right)^2 \left(M_{k2^{-n}} - M_{(k-1)2^{-n}}\right)^2\right]$$

$$= \sum_{k=1}^{\lfloor 2^n T \rfloor} \mathbb{E}\left[\left(M_{(k-1)2^{-n}}\right)^2 \underbrace{\mathbb{E}\left[\left(M_{k2^{-n}} - M_{(k-1)2^{-n}}\right)^2 \mid \mathcal{F}_{k2^{-n}}\right]}_{= \sum_{\ell : k_{n,m}(\ell) = k} \mathbb{E}\left[\left(M_{\ell 2^{-m}} - M_{(\ell-1)2^{-m}}\right)^2 \mid \mathcal{F}_{k2^{-n}}\right]}\right]$$

(by Lemma 2.14 + Corollary)

$$= \sum_{\ell=1}^{\lfloor 2^m T \rfloor} \mathbb{E}\left[\left(M_{(k_{n,m}(\ell)-1)2^{-n}}\right)^2 \cdot \left(M_{\ell 2^{-m}} - M_{(\ell-1)2^{-m}}\right)^2\right].$$

Combining the three, we get

$$\mathbb{E}\left[\left(Y_T^{(n)} - Y_T^{(m)}\right)^2\right]$$

$$= \mathbb{E}\left[\sum_{\ell=1}^{\lfloor 2^m T \rfloor} \left(M_{(k_{n,m}(\ell)-1)2^{-n}} - M_{(\ell-1)2^{-m}}\right)^2 \cdot \left(M_{\ell 2^{-m}} - M_{(\ell-1)2^{-m}}\right)^2\right]$$

$$\leq \mathbb{E}\left[\sup_{\ell} \left(M_{(k_{n,m}(\ell)-1)2^{-n}} - M_{(\ell-1)2^{-m}}\right)^2 \cdot \sum_{\ell=1}^{\lfloor 2^m T \rfloor} \left(M_{\ell 2^{-m}} - M_{(\ell-1)2^{-m}}\right)^2\right]$$

Cauchy-Schwarz inequality

$$\leq \mathbb{E}\left[\sup_{\ell}\left(M_{(k_{n,m}(\ell)-1)2^{-n}} - M_{(\ell-1)2^{-m}}\right)^4\right]^{1/2} \cdot \underbrace{\mathbb{E}\left[\left(\sum_{\ell=1}^{\lfloor 2^m T\rfloor}\left(M_{\ell 2^{-m}} - M_{(\ell-1)2^{-m}}\right)^2\right)^2\right]^{1/2}}_{\leq \sqrt{16 \cdot C^4} \text{ by Lemma 3.6}} .$$

The second factor is bounded by Lemma 3.6. By continuity, the expression inside expectation in the first factor tends to zero as $\min(n,m) \longrightarrow \infty$. Bounded convergence theorem therefore implies that the first factor tends to zero. We conclude

$$\lim_{r\to\infty} \sup_{r\leq n\leq m} \mathbb{E}\left[\left(Y_T^{(n)} - Y_T^{(m)}\right)^2\right] = 0.$$

Recalling that $\| Y^{(n)} - Y^{(m)} \|_{\mathcal{M}_c^2} = \sqrt{\mathbb{E}\left[\left(Y_T^{(n)} - Y_T^{(m)}\right)^2\right]}$, this proves that the sequence $(Y^{(n)})_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{M}_c^2$.

$\square$

NOTES ON DISCRETE TIME MARTINGALES

These notes cover some basic properties of discrete time martingales for a course on stochastic calculus. We assume that the basic concept of a (sub/super)martingale is familiar, along with some results related to optional stopping and the martingale convergence theorem are familiar. We prove Doob's martingale inequality, Doob's maximum L^2 inequality and discuss uniform integrability in the setting of martingales. The discussion covers some of the material from [1, Chapter 5.4 and Chapter 5.5]. We first review some basic definitions and results we shall make use of, then discuss Doob's inequalities, and finally uniform integrability.

1. Background

Let us consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Recall that a *filtration* on $(\Omega, \mathcal{F}, \mathbb{P})$ is an increasing sequence of sub σ-algebras $\mathcal{F}_k$: $\mathcal{F}_k \subset \mathcal{F}_{k+1} \subset \mathcal{F}$ for all $k \in \mathbb{N} = \{0, 1, 2, ...\}$. A sequence of random variables $(X_k)_{k\in\mathbb{N}}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is *adapted* to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$ if X_k is measurable with respect to $\mathcal{F}_k$ for all $k \in \mathbb{N}$. We say that a sequence of real-valued random variables $(X_k)_{k\in\mathbb{N}}$ is a *martingale* (with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$) if

(M1) $\mathbb{E}\,|X_k| < \infty$ for all $k \in \mathbb{N}$,
(M2) $(X_k)_{k\in\mathbb{N}}$ is adapted to $(\mathcal{F}_k)_{k\in\mathbb{N}}$,
(M3) $\mathbb{E}\,(X_{k+1}|\mathcal{F}_k) = X_k$ for all $k \in \mathbb{N}$ (which implies that $\mathbb{E}\,X_k = \mathbb{E}\,X_0$ for all k).

Similarly, we say that a sequence of real-valued random variables $(X_k)_{k\in\mathbb{N}}$ is a *submartingale* (with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$) if

(SUBM1) $\mathbb{E}\,|X_k| < \infty$ for all $k \in \mathbb{N}$,
(SUBM2) $(X_k)_{k\in\mathbb{N}}$ is adapted to $(\mathcal{F}_k)_{k\in\mathbb{N}}$,
(SUBM3) $\mathbb{E}\,(X_{k+1}|\mathcal{F}_k) \geqslant X_k$ for all $k \in \mathbb{N}$ (which implies that $\mathbb{E}\,X_k \leqslant \mathbb{E}\,X_{k+1}$ for all k).

and a *supermartingale* (with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$) if

(SUPM1) $\mathbb{E}\,|X_k| < \infty$ for all $k \in \mathbb{N}$,
(SUPM2) $(X_k)_{k\in\mathbb{N}}$ is adapted to $(\mathcal{F}_k)_{k\in\mathbb{N}}$,
(SUPM3) $\mathbb{E}\,(X_{k+1}|\mathcal{F}_k) \leqslant X_k$ for all $k \in \mathbb{N}$ (which implies that $\mathbb{E}\,X_{k+1} \leqslant \mathbb{E}\,X_k$ for all k).

Note that if $(X_k)_{k\in\mathbb{N}}$ is a supermartingale, then $(-X_k)_{k\in\mathbb{N}}$ is a submartingale, so statements about one automatically imply a converse statement for the other. Moreover, martingales are both submartingales and supermartingales, so results for sub/supermartingales automatically extend to martingales.

A basic that is occasionally useful is that convex functions of martingales are submartingales (if they are in L^1) and increasing convex functions of submartingales are submartingales.

Proposition 1.1. *Let $(X_k)_{k\in\mathbb{N}}$ be a martingale with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$ and $f : \mathbb{R} \to \mathbb{R}$ a convex function such that $\mathbb{E}\,|f(X_k)| < \infty$ for all $k \in \mathbb{N}$. Then $(f(X_k))_{k\in\mathbb{N}}$ is a submartingale with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$.*

Moreover, if $(X_k)_{k\in\mathbb{N}}$ is a submartingale with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$ and $f : \mathbb{R} \to \mathbb{R}$ an increasing convex function such that $\mathbb{E}\,|f(X_k)| < \infty$ for all $k \in \mathbb{N}$, then $(f(X_k))_{k\in\mathbb{N}}$ is a submartingale.

Remark 1.2. *Note that if $(X_k)_{k\in\mathbb{N}}$ is a supermartingale and f is increasing and concave, then an immediate corollary of this proposition is that $(f(X_k))_{k\in\mathbb{N}}$ is a supermartingale. Typical applications of this proposition (and its corollary for supermartingales) are the following:*

- *If $(X_k)_{k\in\mathbb{N}}$ is a martingale, then $(|X_k|^p)_{k\in\mathbb{N}}$ is a submartingale for $p \geqslant 1$ ($x \mapsto |x|^p$ is convex).*

- *If $(X_k)_{k\in\mathbb{N}}$ is a submartingale, then for each $a \in \mathbb{R}$, $(\max(X_k - a, 0))_{k\in\mathbb{N}}$ is a submartingale ($x \mapsto \max(0, x - a)$ is increasing and convex).*
- *If $(X_k)_{k\in\mathbb{N}}$ is a supermartingale, then for each $a \in \mathbb{R}$, $(\min(X_k, a))_{k\in\mathbb{N}}$ is a supermartingale ($x \mapsto \min(x, a)$ is increasing and concave).*

We will also need the notion predictability: we say that a sequence of random variables $(H_k)_{k\in\mathbb{Z}_+}$ (where $\mathbb{Z}_+ = \{1, 2, ...\}$) is *predictable* with respect to a filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$ if for each $k \in \mathbb{Z}_+$, H_k is measurable with respect to the σ-algebra $\mathcal{F}_{k-1}$. We also introduce the notion of the *martingale transform* or *discrete stochastic integral* of $(H_k)_{k\in\mathbb{Z}_+}$ with respect to $(X_k)_{k\in\mathbb{N}}$ by

$$(H \cdot X)_k := \begin{cases} \sum_{j=1}^k H_j(X_j - X_{j-1}), & k \geqslant 1 \\ 0, & k = 0 \end{cases}.$$

The basic use of the discrete stochastic integral in the setting of discrete time martingale theory is the following result.

Theorem 1.3. *Let $(H_k)_{k\in\mathbb{Z}_+}$ be predictable with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$, non-negative and bounded in that there exists some non-random finite $C > 0$ such that almost surely $0 \leqslant H_k \leqslant C$ for all $k \in \mathbb{Z}_+$, and let $(X_k)_{k\in\mathbb{N}}$ be a supermartingale with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$. Then $((H \cdot X)_k)_{k\in\mathbb{N}}$ is a supermartingale with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$.*

Next we recall the notion of a *stopping time*. A random variable N taking values in $\mathbb{N} \cup \{\infty\}$ is called a stopping time with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$ if for each $k \in \mathbb{N}$, the event $\{N \leqslant k\}$ is an element of $\mathcal{F}_k$. One of the main points of discussing stopping times in the setting of martingale theory is the following result.

Theorem 1.4. *If N is a stopping time (with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$) and $(X_k)_{k\in\mathbb{N}}$ is a supermartingale (with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$), then $(X_{\min(N,k)})_{k\in\mathbb{N}}$ is a supermartingale (with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$).*

We will also make use of the following result, known as the martingale convergence theorem, which says that martingales behave very nicely under limiting procedures.

Theorem 1.5. *If $(X_k)_{k\in\mathbb{N}}$ is a submartingale (with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$) with $\sup_{k\in\mathbb{N}} \mathbb{E}\max(X_k, 0) < \infty$, then as $k \to \infty$, X_k converges almost surely to some limit X which satisfies $\mathbb{E}|X| < \infty$.*

This has a particularly nice implication for supermartingales.

Corollary 1.6. *If $(X_k)_{k\in\mathbb{N}}$ is a non-negative supermartingale (with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$), i.e., almost surely $X_k \geqslant 0$ for all $k \in \mathbb{N}$, then as $k \to \infty$, X_k converges almost surely to some limit X which satisfies $X \geqslant 0$ almost surely and $\mathbb{E}X \leqslant \mathbb{E}X_0 < \infty$.*

Proof. Now $(-X_k)_{k\in\mathbb{N}}$ is a non-positive submartingale. In particular, due to non-positivity, $\max(-X_k, 0) = 0$ for all k so we can apply Theorem 1.5 to deduce that as $k \to \infty$, $-X_k$ converges almost surely to some limit, let's say $-X$, which is of course almost surely non-positive by almost sure convergence, and satisfies $\mathbb{E}|-X| < \infty$. This in turn implies that X_k converges to X, and since being a supermartingale implies that $\mathbb{E}X_k \leqslant \mathbb{E}X_0$, we find the final claim by Fatou's lemma. □

We now turn to Doob's martingale inequality.

2. Doob's martingale inequality and maximum L^2 inequality

The basic idea of Doob's inequalities is that if $(X_k)_{k\in\mathbb{N}}$ is a non-negative submartingale, then it's unlikely that $\max_{0\leqslant k\leqslant n} X_k$ is big if it's unlikely for X_n to be big. The form this is typically used in is moment bounds, which is the content of the L^2-inequality, but the L^2-inequality is proven using the martingale inequality. We will formulate the statements more precisely shortly. We begin with a simple result we'll make use of in the proof of Doob's martingale inequality.

Lemma 2.1. *Let $(X_k)_{k\in\mathbb{N}}$ be a submartingale with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$. Let $K \in \mathbb{Z}_+$ be fixed (i.e. non-random), and let N be a stopping time with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$ such that $N \leqslant K$ almost surely. Then*

$$\mathbb{E}\, X_0 \leqslant \mathbb{E}\, X_N \leqslant \mathbb{E}\, X_K.$$

Note that if $(X_k)_{k\in\mathbb{N}}$ is martingale, then $\mathbb{E}\, X_0 = \mathbb{E}\, X_k$ and we have identities here. Also the boundedness assumption here is important: take $(X_k)_{k\in\mathbb{N}}$ to be a simple random walk started at 1 and $N = \min\{k \in \mathbb{Z}_+ : X_k = 0\}$.

Proof. By Theorem 1.4, $(X_{\min(k,N)})_{k\in\mathbb{N}}$ is a submartingale. Thus

$$\mathbb{E}\, X_0 = \mathbb{E}\, X_{\min(N,0)} \leqslant \mathbb{E}\, X_{\min(N,K)} = \mathbb{E}\, X_N,$$

which proves the left inequality in the statement of the lemma. For the right inequality, let us define the sequence of random variables $(H_k)_{k\in\mathbb{Z}_+}$, where $H_k = \mathbf{1}_{\{N\leqslant k-1\}}$ for each $k \in \mathbb{Z}_+$. Since N is a stopping time, we see that H_k is measurable with respect to $\mathcal{F}_{k-1}$ – $(H_k)_{k\in\mathbb{Z}_+}$ is predictable. Moreover, it is of course non-negative and bounded so Theorem 1.3 implies (note that we use it here for $(-X_k)_{k\in\mathbb{N}}$ and then translate it into a statement for submartingales) that $((H \cdot X)_k)_{k\in\mathbb{N}}$ is a submartingale. In particular, we have

$$0 = \mathbb{E}\,(H \cdot X)_0 \leqslant \mathbb{E}\,(H \cdot X)_K. \tag{2.1}$$

But now, note that

$$\begin{aligned}
(H \cdot X)_k &= \sum_{j=1}^{k} H_j (X_j - X_{j-1}) \\
&= \sum_{j=1}^{k} \mathbf{1}_{\{N\leqslant j-1\}} (X_j - X_{j-1}) \\
&= \mathbf{1}_{\{N+1\leqslant k\}} \sum_{j=N+1}^{k} (X_j - X_{j-1}) \\
&= \mathbf{1}_{\{N+1\leqslant k\}} (X_k - X_N) \\
&= X_k - X_{\min(k,N)}.
\end{aligned}$$

Since $\min(K, N) = N$ almost surely, we see that (2.1) becomes

$$0 \leqslant \mathbb{E}\,(X_K - X_N)$$

which is the right inequality in the statement of the lemma. $\square$

This allows us to prove Doob's (sub)martingale inequality.

Theorem 2.2. *Let $(X_k)_{k\in\mathbb{N}}$ be a non-negative submartingale with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$. For each $\lambda > 0$ and $k \in \mathbb{N}$, let us define the event*

$$A_k(\lambda) = \left\{ \max_{0\leqslant j\leqslant k} X_j \geqslant \lambda \right\}.$$

Then

$$\mathbb{P}\,(A_k(\lambda)) \leqslant \frac{\mathbb{E}\, \mathbf{1}_{A_k(\lambda)} X_k}{\lambda} \leqslant \frac{\mathbb{E}\, X_k}{\lambda}.$$

Proof. Let us define the stopping time $N = \min(k, \inf\{m : X_m \geqslant \lambda\})$ (minimum of two stopping times is a stopping time). Since on the event $A_k(\lambda)$, $X_N \geqslant \lambda$, we see that

$$\lambda \mathbb{P}\left(A_k(\lambda)\right) \leqslant \mathbb{E}\, X_N \mathbf{1}_{A_k(\lambda)}.$$

Now applying Lemma 2.1 (N is bounded by k) and noting that on $A_k(\lambda)^{\mathsf{c}}$, $X_N = X_k$, we see that

$$\begin{aligned}\mathbb{E}\, X_N \mathbf{1}_{A_k(\lambda)} &= \mathbb{E}\, X_N - \mathbb{E}\, X_N \mathbf{1}_{A_k(\lambda)^{\mathsf{c}}} \\ &\leqslant \mathbb{E}\, X_k - \mathbb{E}\, X_k \mathbf{1}_{A_k(\lambda)^{\mathsf{c}}} \\ &= \mathbb{E}\, X_k \mathbf{1}_{A_k(\lambda)}\end{aligned}$$

so combining our two inequalities yields

$$\mathbb{P}\left(A_k(\lambda)\right) \leqslant \frac{\mathbb{E}\, X_k \mathbf{1}_{A_k(\lambda)}}{\lambda},$$

which is just the left inequality in the statement of the theorem. For the right inequality, we note that $\mathbb{E}\, X_k \mathbf{1}_{A_k(\lambda)} \leqslant \mathbb{E}\, X_k$ by non-negativity. □

Remark 2.3. *Note that in view of Remark 1.2, if one drops the assumption of non-negativity of the submartingale from Theorem 2.2, one still can have results of the same flavour: e.g. by replacing $(X_k)_{k\in\mathbb{N}}$ with $(\max(X_k, 0))_{k\in\mathbb{N}}$ which is a non-negative submartingale (by Remark 1.2).*

As an application of this, we prove Doob's L^2 maximal inequality which one perhaps uses more often than Theorem 2.2.

Theorem 2.4. *Let $(X_k)_{k\in\mathbb{N}}$ be a non-negative submartingale (with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$). Then*

$$\mathbb{E}\left(\sup_{0\leqslant j\leqslant k} X_j^2\right) \leqslant 4\mathbb{E}\, X_k^2.$$

In particular, if $(X_k)_{k\in\mathbb{N}}$ is a martingale (not necessarily non-negative), then

$$\mathbb{E}\left[\sup_{0\leqslant j\leqslant k} |X_j|^2\right] \leqslant 4\mathbb{E}\,|X_k|^2.$$

Proof. The second statement follows from the first one since if $(X_k)_{k\in\mathbb{N}}$ is a martingale, then $(|X_k|)_{k\in\mathbb{N}}$ is a non-negative submartingale – see Remark 1.2. Let us thus focus on the first statement. Naturally if $\mathbb{E}\, X_k^2 = \infty$, this is not a very interesting statement, so let us assume that $\mathbb{E}\, X_j^2 < \infty$ for all $j \leqslant k$ (if $\mathbb{E}\, X_j^2 = \infty$ for some $j \leqslant k$ then $\mathbb{E}\, X_k^2$ by the submartingale property). As $\sup_{0\leqslant j\leqslant k} X_j^2 \leqslant \sum_{j=0}^k X_j^2$, we see that also $\mathbb{E}\left[\sup_{0\leqslant j\leqslant k} X_j^2\right] < \infty$ in this case. Using Doob's martingale inequality (Theorem 2.2), Fubini, and Cauchy-Schwarz, we now have[1]

[1]The first identity here uses the fact that for any non-negative random variable Y and $p > 0$, $\mathbb{E}\, Y^p = \int_0^\infty p\lambda^{p-1}\mathbb{P}(Y \geqslant \lambda)d\lambda$. This follows from Fubini: $\int_0^\infty p\lambda^{p-1}\mathbb{P}(Y \geqslant \lambda)d\lambda = \mathbb{E}\int_0^\infty p\lambda^{p-1}\mathbf{1}_{\{Y\geqslant\lambda\}}d\lambda = \mathbb{E}\int_0^Y p\lambda^{p-1}d\lambda = \mathbb{E}\, Y^p$.

$$
\begin{aligned}
\mathbb{E}\left(\sup_{0\leqslant j\leqslant k} X_j^2\right) &= 2\int_0^\infty \lambda \mathbb{P}\left(\sup_{0\leqslant j\leqslant k} X_j > \lambda\right) d\lambda \\
&\leqslant 2\int_0^\infty \lambda \frac{\mathbb{E}\, X_k \mathbf{1}_{\{\sup_{0\leqslant j\leqslant k} X_j > \lambda\}}}{\lambda} d\lambda \\
&= 2\mathbb{E}\left[X_k \int_0^{\sup_{0\leqslant j\leqslant k} X_j} d\lambda\right] \\
&= 2\mathbb{E}\, X_k \sup_{0\leqslant j\leqslant k} X_j \\
&\leqslant 2\sqrt{\mathbb{E}\, X_k^2 \mathbb{E}\left(\sup_{0\leqslant j\leqslant k} X_j^2\right)}.
\end{aligned}
$$

If $X_k = 0$ almost surely (implying that $X_j = 0$ almost surely by non-negativity and the sub-martingale property), we have nothing to prove, so let us assume that $\mathbb{E}\, X_k^2 > 0$ which implies also that $\mathbb{E}\,(\sup_{0\leqslant j\leqslant k} X_j^2) > 0$. We can thus divide by $\sqrt{\mathbb{E}\left(\sup_{0\leqslant j\leqslant k} X_j^2\right)}$ yielding

$$
\sqrt{\mathbb{E}\left(\sup_{0\leqslant j\leqslant k} X_j^2\right)} \leqslant 2\sqrt{\mathbb{E}\, X_k^2}
$$

from which the result follow by squaring. □

We conclude this section with the following addition to the martingale convergence theorem.

Theorem 2.5. *Let $(X_k)_{k\in\mathbb{N}}$ be a martingale and $\sup_{k\in\mathbb{N}} \mathbb{E}\, X_k^2 < \infty$. Then X_k converges to a limit X almost surely and in L^2.*

Proof. We now have (by the trivial inequality $\max(X_k, 0) \leqslant |X_k|$ and Cauchy-Schwarz)

$$
\sup_{k\in\mathbb{N}} \mathbb{E}\, \max(X_k, 0) \leqslant \sup_{k\in\mathbb{N}} \mathbb{E}\,|X_k| \leqslant \sqrt{\sup_{k\in\mathbb{N}} \mathbb{E}\, X_k^2} < \infty.
$$

Thus by the martingale convergence theorem (Theorem 1.5), X_k converges almost surely to some limit X. Let us now prove convergence in L^2 – namely that $\lim_{k\to\infty} \mathbb{E}\,|X_k - X|^2 = 0$. The idea is that we want to show this by dominated convergence, which is justified e.g. if we prove that $\mathbb{E}\,[\sup_{k\in\mathbb{N}} |X_k - X|^2] < \infty$. Noting that $|X_k - X|^2 \leqslant 4\sup_{k\in\mathbb{N}} |X_k|^2$ by the triangular inequality (and the fact that $\sup_k |X_k| \geqslant |X|$), we see that it is actually sufficient to prove that $\mathbb{E}\,[\sup_{k\in\mathbb{N}} |X_k|^2] < \infty$. Now from Theorem 2.4, we have for $n \in \mathbb{N}$

$$
\mathbb{E}\left(\sup_{0\leqslant k\leqslant n} |X_k|^2\right) \leqslant 4\mathbb{E}\,|X_n|^2 \leqslant 4\sup_{k\in\mathbb{N}} \mathbb{E}\, X_k^2 < \infty
$$

so letting $n \to \infty$ we find (by the monotone convergence theorem)

$$
\mathbb{E}\left(\sup_{k\in\mathbb{N}} |X_k|^2\right) \leqslant 4\sup_{k\in\mathbb{N}} \mathbb{E}\, X_k^2 < \infty.
$$

As mentioned, L^2-convergence now follows from the dominated convergence theorem. □

Remark 2.6. *All of the results and proofs of this section extend easily to L^p-versions of the statements (so replace X_k^2 by $|X_k|^p$) for arbitrary $p \in (1, \infty)$. Though not to $p = 1$!*

We now turn to uniform integrability.

3. Uniform integrability and martingales

The martingale convergence theorem is an important part of the theory of martingales in that it shows that martingales have a rich limit theory: one needs rather weak assumptions on the martingale to ensure that a limit exists. Nothing in the martingale convergence theorem though ensures that the limit is non-trivial. Uniform integrability is a condition that can sometimes be used to prove that a limit provided by the martingale convergence theorem is non-trivial. Also it will lead to a particularly nice representation of a martingale in terms of its limit. In this section we briefly review the main results.

A family of real-valued random variables $(X_i)_{i\in I}$ (note that I need not be countable) is called *uniformly integrable* if

$$\lim_{M\to\infty}\left[\sup_{i\in I}\mathbb{E}\left(|X_i|\mathbf{1}_{\{|X_i|>M\}}\right)\right]=0.$$

We record here a basic exercise related to the definition of uniform integrability. We will need its result shortly.

Exercise 3.1. *Let $(X_i)_{i\in I}$ be a family of uniformly integrable random variables. Show that $\sup_{i\in I}\mathbb{E}|X_i|<\infty$.*

Let us now see how knowing uniform integrability can strengthen the martingale convergence theorem.

Theorem 3.2. *Let $(X_k)_{k\in\mathbb{N}}$ be a uniformly integrable submartingale (with respect to some filtration) on some probability space $(\Omega,\mathcal{F},\mathbb{P})$. Then the there exists a random variable X on the probability space $(\Omega,\mathcal{F},\mathbb{P})$ such that $X_k\to X$ almost surely and $\lim_{k\to\infty}\mathbb{E}|X_k-X|=0$ (in particular, for a martingale $\mathbb{E}X=\mathbb{E}X_0$).*

Proof. Let us begin by noting that from Exercise 3.1, $\sup_k\mathbb{E}|X_k|<\infty$, so (since $\max(0,x)\leqslant|x|$) we see from the martingale convergence theorem (Theorem 1.5) that there exists a random variable X such that $X_k\to X$ almost surely and $\mathbb{E}|X|<\infty$. Let us turn to the second claim. The proof of it will make use of uniform integrability, and for this, we introduce a parameter $M>0$ and the bounded continuous function

$$\varphi_M(x)=\begin{cases}-M, & x\leqslant -M\\ x, & |x|<M\\ M, & x>M\end{cases}.$$

Note that if $x\geqslant M$, then $|x-\varphi_M(x)|=|x-M|\leqslant|x|$ and similarly if $x<-M$, then $|x-\varphi_M(x)|=|x+M|\leqslant|x|$. Combining these remarks with the fact that $\varphi_M(x)=x$ for $|x|\leqslant M$, we see that $|x-\varphi_M(x)|\leqslant|x|\mathbf{1}_{\{|x|\geqslant M\}}$. Thus by the triangular inequality

$$\begin{aligned}|X_k-X|&\leqslant|X_k-\varphi_M(X_k)|+|\varphi_M(X_k)-\varphi_M(X)|+|\varphi_M(X)-X|\\&\leqslant\mathbf{1}_{\{|X_k|>M\}}|X_k|+|\varphi_M(X_k)-\varphi_M(X)|+\mathbf{1}_{\{|X|>M\}}|X|\\&=\Delta_{k,1}(M)+\Delta_{k,2}(M)+\Delta_3(M).\end{aligned}$$

First by uniform integrability of $(X_k)_{k\in\mathbb{N}}$, we have for each $k\in\mathbb{N}$

$$0\leqslant\mathbb{E}\,\Delta_{k,1}(M)=\mathbb{E}\left(\mathbf{1}_{\{|X_k|>M\}}|X_k|\right)\leqslant\sup_{k\in\mathbb{N}}\mathbb{E}\left(\mathbf{1}_{\{|X_k|>M\}}|X_k|\right)\stackrel{M\to\infty}{\longrightarrow}0. \tag{3.1}$$

Then by the dominated convergence theorem (using the fact that by continuity of φ_M, almost sure convergence of X_k to X implies almost sure convergence of $\varphi_M(X_k)$ to $\varphi_M(X)$ for each fixed $M>0$), we have for each fixed $M>0$

$$\lim_{k\to\infty}\mathbb{E}\,\Delta_{k,2}(M)=0. \tag{3.2}$$

For the last term, we use the fact that the martingale convergence theorem ensures that $\mathbb{E}\,|X| < \infty$ so by the dominated convergence theorem

$$0 \leqslant \lim_{M\to\infty} \mathbb{E}\,\Delta_3(M) = \mathbb{E}\left(\lim_{M\to\infty} \mathbf{1}_{\{|X|>M\}}|X|\right) = 0. \tag{3.3}$$

Combining (3.1), (3.2), and (3.3), we find that

$$\lim_{k\to\infty} \mathbb{E}\,|X_k - X| \leqslant \lim_{M\to\infty} \limsup_{k\to\infty} [\mathbb{E}\,\Delta_{k,1}(M) + \mathbb{E}\,\Delta_{k,2}(M) + \mathbb{E}\,\Delta_3(M)] = 0,$$

which was the claim.

The final remark for martingales follows from noting that for a martingale, $0 \leqslant |\mathbb{E}\,X_0 - \mathbb{E}\,X| = |\mathbb{E}\,X_k - \mathbb{E}\,X| \leqslant \mathbb{E}\,|X_k - X| \to 0$. □

If $(X_k)_{k\in\mathbb{N}}$ is a martingale, this theorem implies a particularly nice representation for the martingale in terms of its limit.

Corollary 3.3. *Let $(X_k)_{k\in\mathbb{N}}$ be a uniformly integrable martingale with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then there exists a random variable X on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $X_k \to X$ almost surely, $\lim_{k\to\infty} \mathbb{E}\,|X_k - X| = 0$, and $X_k = \mathbb{E}\,(X|\mathcal{F}_k)$ for all $k \in \mathcal{F}_k$.*

Proof. The convergence statements are identical to the ones in Theorem 3.2, so we only need to focus on the representation $X_k = \mathbb{E}\,(X|\mathcal{F}_k)$. First of all note that as $\mathbb{E}\,|X| < \infty$, the conditional expectation is well defined and the statement is meaningful. According to the definition of conditional expectation, what we actually want to prove is that for each $A \in \mathcal{F}_k$, $\mathbb{E}\,\mathbf{1}_A X_k = \mathbb{E}\,\mathbf{1}_A X$. To do this, note that by the martingale property, we have for each $n > k$, $\mathbb{E}\,(X_n|\mathcal{F}_k) = X_k$, or in other words that for each $n > k$, $\mathbb{E}\,\mathbf{1}_A X_n = \mathbb{E}\,\mathbf{1}_A X_k$. We will thus be done if we can prove that $\lim_{n\to\infty} \mathbb{E}\,\mathbf{1}_A X_n = \mathbb{E}\,\mathbf{1}_A X$. To prove this, note that for any $A \in \mathcal{F}$,

$$|\mathbb{E}\,\mathbf{1}_A X_n - \mathbb{E}\,\mathbf{1}_A X| \leqslant \mathbb{E}\,|X_n - X| \overset{n\to\infty}{\longrightarrow} 0$$

as we knew already that $\lim_{k\to\infty} \mathbb{E}\,|X_k - X| = 0$. We are thus done since as we already said, we have for all $n > k$ and $A \in \mathcal{F}_k$

$$\mathbb{E}\,\mathbf{1}_A X_k = \mathbb{E}\,\mathbf{1}_A X_n \overset{n\to\infty}{\longrightarrow} \mathbb{E}\,\mathbf{1}_A X$$

which was equivalent to $X_k = \mathbb{E}\,(X|\mathcal{F}_k)$. □

Remark 3.4. *In fact, in Theorem 3.2 and Corollary 3.3, all of these facts are equivalent: a submartingale is uniformly integrable if and only if it converges in L^1, which is equivalent to it converging almost surely and in L^1. For a martingale, uniform integrability is also equivalent to the representation $X_k = \mathbb{E}\,(X|\mathcal{F}_k)$ for some X for which $\mathbb{E}\,|X| < \infty$. Proving these facts is not particularly hard, but typically one uses these results in the direction we have proved them in. For more details, see e.g.* [1, Chapter 5.5].

References

[1] R. Durret: Probability: theory and examples. Fourth edition. Cambridge Series in Statistical and Probabilistic Mathematics, 31. Cambridge University Press, Cambridge, 2010. x+428 pp.

PREDICTABLE PROCESSES

The "integrands" in our stochastic integral are required to be "predictable" in the precise sense defined below. The interpretation — in stock market terms — is that our decisions about our portfolio must be made with information available at the time, and we furthermore can not adjust the portfolio instantaneously when we see a change in a stock price. The mathematical reason for the requirement of predictability is to guarantee that integration against (local) martingales produces (local) martingales.

Here we view the $\mathbb{R}$-valued stochastic process $H = (H_t)_{t\in(0,\infty)}$ as a function of two variables — namely for a fixed $t>0$, the value H_t is a random variable

$$H_t : \Omega \longrightarrow \mathbb{R} ,$$

so H itself depends on both $t>0$ and $\omega\in\Omega$, i.e., it is a function

$$\begin{aligned} H : \Omega \times (0,\infty) &\longrightarrow \mathbb{R} \\ (\omega, t) &\longmapsto H_t(\omega) . \end{aligned}$$

Definition: The predictable σ-algebra is the σ-algebra $\mathcal{P}$ on $\Omega \times (0,\infty)$ generated by sets of the form

$$E \times (s,t]$$

where $0 \leq s < t$ and $E \in \mathcal{F}_s$.
A predictable process is a $\mathcal{P}$-measurable function $H : \Omega \times (0,\infty) \longrightarrow \mathbb{R}$.

Most of the time the following proposition saves us from worrying about measurability with respect to $\mathcal{P}$.

Proposition 2.4: If $H = (H_t)_{t\in(0,\infty)}$ is adapted and continuous (i.e. the paths $t \mapsto H_t(\omega)$ are continuous for all $\omega \in \Omega$), then H is predictable.

Proof: We "approximate" H by piecewise constant processes $H^{(n)}$ which are constant between times of the form $k \cdot 2^{-n}$, $k \in \mathbb{Z}_{\geq 0}$.

Define $H^{(n)}$ by

$$H^{(n)}_t(\omega) = \sum_{k=1}^{\infty} \mathbb{1}_{((k-1)2^{-n},\, k\cdot 2^{-n}]}(t) \cdot H_{(k-1)2^{-n}}(\omega)$$

i.e., if for $t > 0$ we denote

$$t_n^- = \sup\, [0,t) \cap 2^{-n}\mathbb{Z}_{\geq 0}$$

(the largest time of the form $k \cdot 2^{-n}$ which is strictly smaller than t), then

$$H^{(n)}_t(\omega) = H_{t_n^-}(\omega).$$

Clearly $t_n^- \uparrow t$ as $n \to \infty$, so by continuity of H we have

$$H^{(n)}_t(\omega) \longrightarrow H_t(\omega) \quad \text{as } n \to \infty.$$

If we show that $H^{(n)}$ is $\mathcal{P}$-measurable, then H will also be $\mathcal{P}$-measurable as the pointwise limit of the $H^{(n)}$.

But if we look at one term in the sum that defines $H^{(n)}$,

$$\mathbb{1}_{((k-1)2^{-n},\, k\cdot 2^{-n}]}(t) \cdot H_{(k-1)2^{-n}}(\omega)$$

and notice that $H_{(k-1)2^{-n}}$ is $\mathcal{F}_{(k-1)2^{-n}}$-measurable since H is adapted, then the $\mathcal{P}$-measurability of $H^{(n)}$ becomes clear. This finishes the proof. $\square$

Remark In the proof we actually only used the left-continuity of H, $H_{t_n^-}(\omega) \to H_t(\omega)$ as $t_n^- \uparrow t$.

Example: Brownian motion $B = (B_t)_{t \in \mathbb{R}_+}$ is continuous (and adapted to its natural filtration), and therefore predictable.

STOCHASTIC INTEGRATION W.R.T. MARTINGALES BOUNDED IN L^2

We first construct the integral

$$\int H_s \, dM_s$$

assuming that $M \in \mathcal{M}_c^2$

$$\mathcal{M}_c^2 := \Big\{ M = (M_t)_{t \in \mathbb{R}_+} \;\Big|\; \sup_{t \geq 0} \mathbb{E}[M_t^2] < \infty, \; M \text{ continuous martingale} \Big\}$$

Recall that $\mathcal{M}_c^2$ is a complete space w.r.t. the norm

$$\|M\|_{\mathcal{M}_c^2} = \sqrt{\mathbb{E}[M_\infty^2]} = \|M_\infty\|_{\mathcal{L}^2(\mathbb{P})}$$

where M_∞ is the limit of

$$M_t \xrightarrow[t \to \infty]{a.s.} M_\infty \qquad \text{and} \qquad M_t \xrightarrow[t \to \infty]{\mathcal{L}^2} M_\infty .$$

The idea is to approximate H by "simple processes" $H^{(n)} \longrightarrow H$ (in a suitable sense which depends on M), for which the integral is easy to define. Then we observe that the integrals $\int H_s^{(n)} \, dM_s$ for the approximating sequence also live in $\mathcal{M}_c^2$ and form a Cauchy sequence there. By completeness, we can define the integral $\int H_s \, dM_s$ as the limit of this sequence, if we also verify that the limit does not depend on the chosen approximation $H^{(n)} \longrightarrow H$.

Simple processes

Def: A simple process is a function $H : \Omega \times (0,\infty) \longrightarrow \mathbb{R}$ of the form

$$H_t(\omega) = \sum_{k=1}^{m} Z_k(\omega) \cdot \mathbb{1}_{(t_{k-1}, t_k]}(t)$$

where $m \in \mathbb{N}$, $0 = t_0 < t_1 < \dots < t_m < \infty$ and each Z_k is a bounded $\mathcal{F}_{t_{k-1}}$-measurable random variable.

We denote by $\mathcal{S}$ the collection of all simple processes. It is obvious that $\mathcal{S}$ is a vector space and

$$\mathcal{S} \subset b\mathcal{P} \subset m\mathcal{P}.$$

simple processes — bounded predictable processes — predictable processes

Def: For $H = \sum_{k=1}^{m} Z_k \cdot \mathbb{1}_{(t_{k-1}, t_k]} \in \mathcal{S}$ and $M \in \mathcal{M}_c^2$ define $H \cdot M = ((H \cdot M)_t)_{t \in \mathbb{R}_+}$ by

$$(H \cdot M)_t = \sum_{k=1}^{m} Z_k \cdot \left(M_{t \wedge t_k} - M_{t \wedge t_{k-1}} \right).$$

(This is our stochastic integral $\int_0^t H_s \, dM_s$ of a simple process $H \in \mathcal{S}$.)

Recall that for τ a stopping time and $X = (X_t)_{t \in \mathbb{R}_+}$ a process, we defined the stopped proces X^τ by $X_t^\tau = X_{t \wedge \tau}$.

Proposition 3.3: Let $H \in \mathcal{S}$ and $M \in \mathcal{M}_c^2$ and let τ be a stopping time. Then

(i): $(H \cdot M^\tau) = (H \cdot M)^\tau$

(ii): $H \cdot M \in \mathcal{M}_c^2$.

(iii): $\mathbb{E}\left[(H\cdot M)_\infty^2\right] = \mathbb{E}\left[\int_0^\infty H_s^2 \, \underbrace{d\langle M,M\rangle_s}_{\text{quadratic variation process of } M}\right]$.

This random variable is defined as the integral w.r.t. the increasing (finite variation) process $\langle M,M\rangle$:

$$\omega \longmapsto \int_0^\infty H_s(\omega)^2 \, d\langle M,M\rangle_s(\omega)$$

Proof: (i): For any $t \geq 0$ we have, if $H = \sum_k Z_k \mathbb{1}_{(t_{k-1},t_k]}$

$$\begin{aligned}(H\cdot M^\tau)_t &= \sum_{k=1}^m Z_k \cdot \left(M^\tau_{t\wedge t_k} - M^\tau_{t\wedge t_{k-1}}\right) \\ &= \sum_{k=1}^m Z_k \left(M_{t\wedge t_k \wedge \tau} - M_{t\wedge t_{k-1}\wedge\tau}\right) \\ &= (H\cdot M)_{t\wedge\tau} \quad = (H\cdot M)^\tau_t .\end{aligned}$$

(ii) The continuity of $H\cdot M$ is clear from the defining formula and continuity of M.

Clearly $H\cdot M$ is adapted (the value at time t is a finite sum of $\mathcal{F}_t$-measurable terms) and integrable (the value is a finite sum of terms which are a bounded random variable times integrable r.v.).

To check the martingale property of $H\cdot M$ assume first that $t_{k-1} \leq s \leq t \leq t_k$.

In that case $(H\cdot M)_t - (H\cdot M)_s = Z_k \cdot (M_t - M_s)$.

Therefore

$$\mathbb{E}[(H\cdot M)_t \mid \mathcal{F}_s] = \mathbb{E}[(H\cdot M)_s + Z_k\cdot(M_t - M_s) \mid \mathcal{F}_s]$$

$$= (H\cdot M)_s + Z_k \cdot \underbrace{\mathbb{E}[M_t - M_s \mid \mathcal{F}_s]}_{=0 \text{ since } M \text{ is mgale}} = (H\cdot M)_s .$$

The above extends straightforwardly to all $0 \le s < t$, so $H\cdot M$ is a martingale.

It remains to show that $H\cdot M$ is bounded in $\mathcal{L}^2$. For this, first observe the following "orthogonality relation": for $1 \le j < k \le m$

$$\mathbb{E}[Z_j\cdot(M_{t_j} - M_{t_{j-1}})\cdot Z_k \cdot (M_{t_k} - M_{t_{k-1}})]$$

$$= \mathbb{E}\Big[Z_j\cdot(M_{t_j} - M_{t_{j-1}})\cdot Z_k \cdot \underbrace{\mathbb{E}[M_{t_k} - M_{t_{k-1}} \mid \mathcal{F}_{t_{k-1}}]}_{=0}\Big] = 0.$$

For $t \ge t_m$, by expanding the square and using this orthogonality, we get

$$\mathbb{E}[(H\cdot M)_t^2] = \mathbb{E}\Big[\Big(\sum_{k=1}^m Z_k\cdot(M_{t_k} - M_{t_{k-1}})\Big)^2\Big]$$

$$= \sum_{k=1}^m \mathbb{E}[Z_k^2\,(M_{t_k} - M_{t_{k-1}})^2].$$

Since H is simple, we have $|Z_k| \le C$ for some $C < \infty$ and all k, so we get

$$\mathbb{E}[(H\cdot M)_t^2] \le C^2 \sum_{k=1}^m \mathbb{E}[(M_{t_k} - M_{t_{k-1}})^2]$$

$$= C^2 \cdot \sum_{k=1}^m \mathbb{E}[M_{t_k}^2 - M_{t_{k-1}}^2] \quad \left(\begin{array}{c}\text{Lemma 2.14} \\ \text{"basic trick"}\end{array}\right)$$

$$= C^2\, \mathbb{E}[M_{t_m}^2 - M_{t_0}^2]$$

$$\le C^2 \cdot \sup_{t\ge 0} \mathbb{E}[M_t^2] < \infty .$$

The case $t \le t_m$ follows also, since Lemma 2.14 and its Corollary give that $\mathbb{E}[(H\cdot M)_t^2]$ is increasing in t.

(iii): If $H = \sum_{k=1}^{m} Z_k \cdot \mathbb{1}_{(t_{k-1}, t_k]}$, then clearly $H \cdot M$ stays constant after time t_m, so

$$(H \cdot M)_\infty = \sum_{k=1}^{m} Z_k \cdot (M_{t_k} - M_{t_{k-1}}).$$

We again use Lemma 2.14 ("basic trick")

$$\mathbb{E}\left[(H \cdot M)_\infty^2\right] = \sum_{k=1}^{m} \mathbb{E}\left[Z_k^2 \left(M_{t_k} - M_{t_{k-1}}\right)^2\right]$$

$$= \sum_{k=1}^{m} \mathbb{E}\Big[Z_k^2 \cdot \underbrace{\mathbb{E}\left[(M_{t_k} - M_{t_{k-1}})^2 \,\middle|\, \mathcal{F}_{t_{k-1}}\right]}_{= \mathbb{E}\left[M_{t_k}^2 - M_{t_{k-1}}^2 \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \text{ (by "basic trick")}}\Big]$$

Then recall that the quadratic variation process $\langle M, M \rangle$ is such that $M - \langle M, M \rangle$ is a martingale. This gives

$$0 = \mathbb{E}\left[M_{t_k}^2 - \langle M, M \rangle_{t_k} - M_{t_{k-1}}^2 + \langle M, M \rangle_{t_{k-1}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right]$$

i.e.

$$\mathbb{E}\left[M_{t_k}^2 - M_{t_{k-1}}^2 \,\middle|\, \mathcal{F}_{t_{k-1}}\right]$$

$$= \mathbb{E}\left[\langle M, M \rangle_{t_k} - \langle M, M \rangle_{t_{k-1}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right].$$

Inserting this in the earlier calculation gives

$$\mathbb{E}\left[(H \cdot M)_\infty^2\right] = \sum_{k=1}^{m} \mathbb{E}\left[Z_k^2 \cdot \left(\langle M, M \rangle_{t_k} - \langle M, M \rangle_{t_{k-1}}\right)\right]$$

$$= \sum_{k=1}^{m} \mathbb{E}\left[\int_{t_{k-1}}^{t_k} H_s^2 \, d\langle M, M \rangle_s\right]$$

$$= \mathbb{E}\left[\int_0^\infty H_s^2 \, d\langle M, M \rangle_s\right]. \qquad \square$$

Itô isometry

Having constructed the stochastic integral $H \cdot M$ for all simple processes $H \in \mathcal{S}$ and all $M \in \mathcal{M}_c^2$, our next goal is to approximate a general predictable process H by simple processes and take a limit.

We now keep $M \in \mathcal{M}_c^2$ fixed. We define the norm (which depends on M) of a predictable process H by

$$\|H\|_M = \left(\mathbb{E}\left[\int_{(0,\infty)} H_s^2 \, d\langle M,M\rangle \right] \right)^{1/2}.$$

We denote by $L^2(M)$ the space of all predictable processes H for which $\|H\|_M < \infty$.

__Exercise__ Show that $\|\cdot\|_M$ is a (pseudo-) norm on $L^2(M)$, i.e., it satisfies

- $\|cH\|_M = |c| \cdot \|H\|_M$ for $c \in \mathbb{R}$, $H \in L^2(M)$
- $\|H + \tilde{H}\|_M \leq \|H\|_M + \|\tilde{H}\|_M$ for $H, \tilde{H} \in L^2(M)$

In order to be able to approximate an arbitrary $H \in L^2(M)$ by simple processes, we need the following "density result".

__Proposition 3.8:__ We have $\mathcal{S} \subset L^2(M)$ and for any $H \in L^2(M)$ there exists a sequence $H^{(1)}, H^{(2)}, \ldots \in \mathcal{S}$ such that

$$\|H - H^{(n)}\|_M \xrightarrow[n \to \infty]{} 0.$$

In the proof we use the following observation

<u>Theorem 3.7</u> If $M \in \mathcal{M}_c^2$, then $M^2 - \langle M,M\rangle$ is a uniformly integrable martingale and $\mathbb{E}[\langle M,M\rangle_\infty] = \mathbb{E}[(M_\infty - M_0)^2]$.

<u>Proof</u>: $t \mapsto \langle M,M\rangle_t$ is increasing, so the limit $\langle M,M\rangle_\infty := \lim_{t\to\infty} \langle M,M\rangle_t$ exists and equals $\sup_{t\geq 0} \langle M,M\rangle_t$.

The quadratic variation of $t \mapsto M_t - M_0$ is the same as that of $t \mapsto M_t$, so we may assume $M_0 = 0$ without loss of generality.

For $n \in \mathbb{N}$, let τ_n be the stopping time

$$\tau_n = \inf\{ t \geq 0 \mid \langle M,M\rangle_t \geq n \}.$$

Then the local martingale $M^2 - \langle M,M\rangle$ stopped at time τ_n satisfies

$$\left| M^2_{t\wedge\tau_n} - \langle M,M\rangle_{t\wedge\tau_n} \right| \leq \sup_{t\geq 0} M_t^2 + n$$

and the RHS upper bound is integrable by Doob's L^2-maximal inequality. Therefore this stopped process is a martingale and we get

$$0 = \mathbb{E}[M_0^2 + \langle M,M\rangle_0] = \mathbb{E}[M^2_{t\wedge\tau_n} - \langle M,M\rangle_{t\wedge\tau_n}]$$

$$\Longrightarrow \mathbb{E}[M^2_{t\wedge\tau_n}] = \mathbb{E}[\langle M,M\rangle_{t\wedge\tau_n}]$$

and by dominated convergence as $t \to \infty$

$$\mathbb{E}[M^2_{\tau_n}] = \mathbb{E}[\langle M,M\rangle_{\tau_n}]$$

and again by dominated and monotone convergence as $n \to \infty$

$$\mathbb{E}[M^2_\infty] = \mathbb{E}[\langle M,M\rangle_\infty].$$

Hence $|M_t^2 - \langle M,M\rangle_t| \leq \sup_{t\geq 0} M_t^2 + \langle M,M\rangle_\infty$,
where the RHS upper bound is integrable.
This shows that $M^2 - \langle M,M\rangle$ is a uniformly integrable martingale (Corollary 2.10). $\square$

<u>Proof of Proposition 3.8:</u> By Theorem 3.7, since $M \in \mathcal{M}_c^2$, we have $\mathbb{E}[\langle M,M\rangle_\infty] < \infty$.
A simple process $H \in \mathcal{S}$ is bounded, $|H_t| \leq C$ $\forall t$, so we get

$$\|H\|_M^2 = \mathbb{E}\left[\int_0^\infty H_s^2 \, d\langle M,M\rangle_s\right]$$
$$\leq \mathbb{E}\left[\int_0^\infty C^2 \, d\langle M,M\rangle_s\right] \leq C^2\, \mathbb{E}[\langle M,M\rangle_\infty] < \infty.$$

This shows $\mathcal{S} \subset L^2(M)$.

To prove that $\mathcal{S}$ is dense in $L^2(M)$, we use a monotone class argument. Let $\mathcal{H}$ be the collection of all predictable processes G for which there exists $H^{(1)}, H^{(2)}, \ldots \in \mathcal{S}$ such that $\|G - H^{(n)}\|_M \longrightarrow 0$ as $n \to \infty$.
We claim that $\mathcal{H}$ is a monotone class which contains the indicator functions of a π-system which generates $\mathcal{P}$.

Let $0 \leq r < s$ and $E \in \mathcal{F}_r$. Then the process $G_t(\omega) = \mathbb{1}_E(\omega)\, \mathbb{1}_{(r,s]}(t)$ is in $\mathcal{H}$ (it is a simple process itself!). This shows that $\mathcal{H}$ contains the indicator functions of $E \times (r,s]$, and sets of this form constitute a π-system which generates $\mathcal{P}$.

Suppose then that $G^{(1)}, G^{(2)}, \ldots \in \mathcal{H}$ are non-negative and $G^{(m)} \uparrow G$ where G is bounded. Then by dominated convergence theorem we get

$$\| G - G^{(m)} \|_M^2 = \mathbb{E}\Big[\int_0^\infty \underbrace{(G_s - G_s^{(m)})^2}_{\longrightarrow 0} \, d\langle M,M\rangle_s \Big] \xrightarrow[m\to\infty]{} 0$$

(dominating random variable : const $\times \langle M,M\rangle_\infty$) .

By assumption, for every m there exists a sequence $H^{(m,1)}, H^{(m,2)}, \ldots \in \mathcal{S}$ s.t.
$\| G^{(m)} - H^{(m,n)} \|_M \longrightarrow 0$ as $n \longrightarrow \infty$.

Fix $\varepsilon > 0$ and take m large enough s.t.

$$\| G - G^{(m)} \|_M < \frac{\varepsilon}{2}$$

and then take n large enough s.t.

$$\| G^{(m)} - H^{(m,n)} \|_M < \frac{\varepsilon}{2} .$$

Triangle inequality gives

$$\| G - H^{(m,n)} \|_M \leq \| G - G^{(m)} \|_M + \| G^{(m)} - H^{(m,n)} \|_M$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon .$$

Since $H^{(m,n)} \in \mathcal{S}$ and this is possible for any $\varepsilon > 0$, there exists a sequence of simple processes approximating G, so $G \in \mathcal{H}$.

To show that $\mathcal{H}$ is a monotone class, it remains to prove that it is a vector space and contains the constant function 1.

For vector space property, suppose that $G, \tilde{G} \in \mathcal{H}$ and $c, \tilde{c} \in \mathbb{R}$. Pick approximating sequences $H^{(1)}, H^{(2)}, \ldots \in \mathcal{S}$

s.t. $\| G - H^{(n)} \|_M \longrightarrow 0$ and $\tilde{H}^{(1)}, \tilde{H}^{(2)}, \ldots \in \mathcal{S}$

s.t. $\| \tilde{G} - \tilde{H}^{(n)} \|_M \longrightarrow 0$ as $n \to \infty$.

By triangle inequality and homogeneity (exercises)

$$\| c \cdot G + \tilde{c} \cdot \tilde{G} - (c \cdot H^{(n)} + \tilde{c} \cdot \tilde{H}^{(n)}) \|_M$$

$$\leq |c| \cdot \| G - H^{(n)} \|_M + |\tilde{c}| \cdot \| \tilde{G} - \tilde{H}^{(n)} \|_M \xrightarrow[n \to \infty]{} 0.$$

This shows that $cG + \tilde{c}\tilde{G} \in \mathcal{H}$.

To show that the constant process 1 belongs to $\mathcal{H}$, note that

$$\| 1 - \mathbb{1}_{(0,n]} \|_M^2 = \mathbb{E}\left[\int_0^\infty (1 - \mathbb{1}_{(0,n]}(s))^2 \, d\langle M, M \rangle_s \right]$$

$$\xrightarrow[n \to \infty]{} 0 \quad \text{by DCT as before.}$$

We have shown that $\mathcal{H}$ is a monotone class which contains the indicator functions of $E \times (s,t]$ $(0 \leq s < t, \; E \in \mathcal{F}_s)$, a π-system which generates $\mathcal{P}$.

Monotone class theorem implies that $\mathcal{H}$ contains all bounded predictable processes, then.

Finally, if $H \in L^2(M)$ is not necessarily bounded, then truncate it: for $m \in \mathbb{N}$, $G^{(m)} = H \cdot \mathbb{1}_{\{|H| \leq m\}}$ is bounded and predictable so we find $H^{(m,1)}, H^{(m,2)}, \ldots \in \mathcal{S}$ s.t.

$\| G^{(m)} - H^{(m,n)} \|_M \xrightarrow[n \to \infty]{} 0$.

Dominated convergence theorem shows that

$$\| G - G^{(m)} \|_M^2 = \mathbb{E}\left[\int_0^\infty (G_s - G_s^{(m)})^2 \, d\langle M, M \rangle_s \right]$$

$$\xrightarrow[m \to \infty]{} 0.$$

Fix $\varepsilon > 0$ and choose m large enough s.t. $\| G - G^{(m)} \|_M < \frac{\varepsilon}{2}$. Then choose n large enough so that $\| G^{(m)} - H^{(m,n)} \|_M < \frac{\varepsilon}{2}$. Triangle inequality gives

$$\| G - H^{(n,m)} \|_M \leq \| G - G^{(m)} \|_M + \| G^{(m)} - H^{(m,n)} \|_M$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon .$$

Since this works for any $\varepsilon > 0$ and $H^{(n,m)} \in \mathcal{S}$, we can find a sequence of simple processes which approximates $G \in L^2(M)$ is the norm $\| \cdot \|_M$.

$\square$

The construction of stochastic integrals with respect to $M \in \mathcal{M}_c^2$ is concluded by the following:

Theorem 3.9 (Itô isometry)

There exists a unique linear mapping

$$I : L^2(M) \longrightarrow \mathcal{M}_c^2$$

such that for $H \in \mathcal{S}$ we have $I(H) = H \bullet M$ and I has the isometry property

$$\| I(H) - I(\tilde{H}) \|_{\mathcal{M}_c^2} = \| H - \tilde{H} \|_M$$

for all $H, \tilde{H} \in L^2(M)$.

Proof: On the dense subspace $\mathcal{S} \subset L^2(M)$ we set $I(H) = H \bullet M$. By Proposition 3.3 (iii) for any $H, \tilde{H} \in \mathcal{S}$ we have

$$\underbrace{\mathbb{E}\left[((H - \tilde{H}) \cdot M)_\infty^2 \right]}_{= \| H \cdot M - \tilde{H} \cdot M \|_{\mathcal{M}_c^2}^2} = \underbrace{\mathbb{E}\left[\int_0^\infty (H_s - \tilde{H}_s) \, d\langle M, M \rangle_s \right]}_{= \| H - \tilde{H} \|_M^2}$$

This shows the isometry property on $\mathcal{S}$.

The uniqueness of the extension of I to $L^2(M)$ is clear: for $H \in L^2(M)$ we can take $H^{(1)}, H^{(2)}, \dots \in \mathcal{S}$ s.t. $\| H - H^{(n)} \|_M \longrightarrow 0$ as $n \to \infty$ so isometry property requires

$$\| I(H) - I(H^{(n)}) \|_{\mathcal{M}_c^2} = \| H - H^{(n)} \|_M \longrightarrow 0,$$

i.e. $I(H)$ must be the limit of $I(H^{(n)}) = H^{(n)} \cdot M$ as $n \to \infty$ in $\mathcal{M}_c^2$.

The isometry property is also used to prove existence of the extension. Let $H \in L^2(M)$ and take $H^{(1)}, H^{(2)}, \dots \in \mathcal{S}$ s.t. $\| H - H^{(n)} \|_M \longrightarrow 0$ as $n \to \infty$. Then the sequence $H^{(1)}, H^{(2)}, \dots$ is Cauchy:

$$\lim_{r \to \infty} \sup_{n, n' \geq r} \| H^{(n)} - H^{(n')} \|_M = 0.$$

The isometry property on $\mathcal{S}$ implies, however,

$$\| H^{(n)} - H^{(n')} \|_M = \| I(H^{(n)}) - I(H^{(n')}) \|_{\mathcal{M}_c^2}$$

so we get

$$\lim_{r \to \infty} \sup_{n, n' \geq r} \| I(H^{(n)}) - I(H^{(n')}) \|_{\mathcal{M}_c^2} = 0.$$

In other words, $I(H^{(1)}), I(H^{(2)}), \dots \in \mathcal{M}_c^2$ is a Cauchy sequence in $\mathcal{M}_c^2$. Since $\mathcal{M}_c^2$ is complete, this sequence has a limit $I(H) := \lim_{n \to \infty} I(H^{(n)})$.

The definition of $I(H)$ does not depend on the chosen approximation: if $K^{(1)}, K^{(2)}, \dots \in \mathcal{S}$

is another approximation, $\| H - K^{(n)} \|_M \to 0$, then

$$\| I(H^{(n)}) - I(K^{(n)}) \|_{\mathcal{M}_c^2} = \| H^{(n)} - K^{(n)} \|_M$$

$$= \| (H - K^{(n)}) - (H - H^{(n)}) \|_M$$

$$\le \| H - K^{(n)} \|_M + \| H - H^{(n)} \|_M \xrightarrow[n \to \infty]{} 0 ,$$

so the limits of $I(H^{(n)})$ and $I(K^{(n)})$ are equal.

A similar calculation shows that this extension $I : L^2(M) \to \mathcal{M}_c^2$ has the isometry property on all of $L^2(M)$.

$\square$

<u>Remark</u> A simple process $H \in \mathcal{S}$ can be approximated by the constant sequence $H^{(1)} = H$, $H^{(2)} = H$, ... so we have $I(H) = H \cdot M$ consistently.

From here on we use the following notations interchangeably for all $H \in L^2(M)$

$$I(M) = H \cdot M$$

$$(I(M))_t = (H \cdot M)_t = \int_0^t H_s \, dM_s .$$

The Itô isometry defines the stochastic integral, when

$M \in \mathcal{M}_c^2$ and $H \in L^2(M)$

$\Updownarrow$ $\Updownarrow$

M continuous martingale, H predictable process

$\sup_{t \ge 0} \mathbb{E}[M_t^2] < \infty$ $\mathbb{E}\left[\int H_s^2 \, d\langle M, M \rangle_s \right] < \infty .$

Recall that we defined integration against $M \in \mathcal{M}_c^2$ (continuous martingale bounded in $\mathcal{L}^2$) as follows:

- For simple processes $H \in \mathcal{S}$,
$$H_t(\omega) = \sum_{k=1}^{m} Z_k(\omega) \cdot \mathbb{1}_{(t_{k-1}, t_k]}(t)$$
where $0 = t_0 < t_1 < \dots < t_m$ and for each $k = 1, \dots, m$, Z_k is a bounded $\mathcal{F}_{t_{k-1}}$-measurable random variable, we set
$$(H \bullet M)_t = \sum_{k=1}^{m} Z_k \cdot \left(M_{t \wedge t_k} - M_{t \wedge t_{k-1}} \right)$$

- For $H \in L^2(M)$, i.e., if H is predictable and
$$\|H\|_M = \left(\mathbb{E}\left[\int_0^\infty H_s^2 \, d\langle M, M\rangle_s \right] \right)^{1/2} < \infty,$$
we used an approximation $H^{(1)}, H^{(2)}, \dots \in \mathcal{S}$ s.t. $\|H - H^{(n)}\|_M \longrightarrow 0$ and set
$$H \bullet M = \lim_{n \to \infty} H^{(n)} \bullet M \qquad (\text{limit in } \mathcal{M}_c^2)$$
which exists and is unique by virtue of Itô isometry.

From here on, we use notations
$$(H \bullet M)_t = \int_0^t H_s \, dM_s$$
interchangeably.

In order to extend to integration against local martingales, we "localize" by choosing a sequence of stopping times. We need to observe the following.

Proposition 3.10 Let $M \in \mathcal{M}_c^2$ and $H \in L^2(M)$ and let τ be a stopping time.
Then $H \cdot \mathbb{1}_{(0,\tau]} \in L^2(M)$ and $H \in L^2(M^\tau)$ and we have
$$(H \bullet M)^\tau = (H \cdot \mathbb{1}_{(0,\tau]}) \bullet M = H \bullet (M^\tau).$$

Sketch: $\mathbb{1}_{(0,\tau]}$ is adapted (note: $\{\mathbb{1}_{(0,\tau]}(t) = 1\} = \{\tau \geq t\} \in \mathcal{F}_t$) and left-continuous, so it is predictable.
Since H is also predictable, the product $H \cdot \mathbb{1}_{(0,\tau]}$ is, too. Also since $|H \cdot \mathbb{1}_{(0,\tau]}| \leq |H|$, we have $\|H \cdot \mathbb{1}_{(0,\tau]}\|_M \leq \|H\|_M < \infty$, so $H \cdot \mathbb{1}_{(0,\tau]} \in L^2(M)$. To check $H \in L^2(M^\tau)$, note
$$\mathbb{E}\left[\int_0^\infty H_s^2 \, d\langle M^\tau, M^\tau\rangle_s\right]$$
$$= \mathbb{E}\left[\int_0^\infty H_s^2 \, d\langle M, M\rangle_{s\wedge\tau}\right] \quad \longleftarrow \text{(Exercise: quadratic variation and stopping)}$$
$$= \mathbb{E}\left[\int_0^\tau H_s^2 \, d\langle M, M\rangle_s\right]$$
$$\leq \mathbb{E}\left[\int_0^\infty H_s^2 \, d\langle M, M\rangle\right] = \|H\|_M^2 < \infty.$$

The equalities are proven in three steps:

1st: assuming H is a simple process and τ has only finitely many possible values

2nd: assuming H is simple, and approximating a general stopping time τ with a sequence $\tau_n \to \tau$ where each τ_n has only finitely many possible values

3rd: for general H, approximating it with a sequence $H^{(n)} \to H$ of simple processes in the norm $\|\cdot\|_M$.

For details, see Berestycki's notes. $\square$

STOCHASTIC INTEGRATION W.R.T LOCAL MARTINGALES

Recall that a process $M = (M_t)_{t \in \mathbb{R}_+}$ is a local martingale if there exists a sequence $\sigma_1, \sigma_2, \ldots$ of stopping times $\sigma_n \uparrow \infty$ s.t. $(M_t^{\sigma_n} - M_0)_{t \in \mathbb{R}_+}$ is a martingale for each $n \in \mathbb{N}$.

We also say that a process $H = (H_t)_{t \in (0,\infty)}$ is locally bounded if there exists a sequence $\sigma_1', \sigma_2', \ldots$ of stopping times $\sigma_n' \uparrow \infty$ and a sequence $c_1, c_2, \ldots \in \mathbb{R}$ of constants s.t.

$$\sup_{t \geq 0} \left| H_t \cdot \mathbb{1}_{(0,\sigma_n']}(t) \right| \leq c_n \qquad \text{(almost surely).}$$

In particular, if $H = (H_t)_{t \in (0,\infty)}$ is adapted and continuous, then we may set $c_n = n$ and $\sigma_n' = \inf\{t > 0 \mid |H_t| \geq n\}$ to see that H is locally bounded (it is predictable by Proposition 2.4).

We then define the integral of a predictable locally bounded process $H = (H_t)_{t \in (0,\infty)}$ against a continuous local martingale $M \in \mathcal{M}_{c,loc}$ as follows. Let $\sigma_1', \sigma_2', \ldots$ be stopping times as above and $\tau_1', \tau_2', \ldots$ defined by

$$\tau_n' = \inf\{t \geq 0 \mid |M_t - M_0| \geq n\}$$

and set $\tau_n = \sigma_n' \wedge \tau_n'$. We still have $\tau_n \uparrow \infty$, and we define

$$(H \bullet M)_t = \left((H \cdot \mathbb{1}_{(0,\tau_n]}) \bullet (M^{\tau_n} - M_0) \right)_t \quad \text{when } t \leq \tau_n.$$

This definition is consistent :

- The RHS stochastic integral is well defined, because $M^{\tau_n} - M_0$ is a bounded continuous martingale, so $M^{\tau_n} - M_0 \in \mathcal{M}_c^2$, and $H \cdot \mathbb{1}_{(0,\tau_n]}$ is bounded and predictable, so $H \cdot \mathbb{1}_{(0,\tau_n]} \in L^2(M^{\tau_n} - M_0)$.

These are consequences of Proposition 3.10.

- For n large enough so that $\tau_n \geq t$, the RHS does not depend on n,
- The definition also does not depend on the sequence of stopping times used to "localize" (reduce) H and M.
- The definition agrees with earlier definition if $M \in \mathcal{M}_c^2$ and $H \in L^2(M)$.

The equalities of Proposition 3.10. continue to hold for this generalized definition.

<u>Proposition 3.11:</u> Suppose that H is a locally bounded predictable process and $M \in \mathcal{M}_{c,loc}$ is a continuous local martingale. Then for all stopping times τ we have

$$(H \bullet M)^\tau = (H \cdot \mathbb{1}_{(0,\tau]}) \bullet M = H \bullet (M^\tau).$$

Proposition 3.14 (Only the special case of local martingales)

Let $M \in \mathcal{M}_{c,loc}$ and let H be locally bounded, adapted and left-continuous. Then we have

$$\sum_{k=1}^{\lfloor 2^n t \rfloor} H_{(k-1)2^{-n}} \cdot \left(M_{k2^{-n}} - M_{(k-1)2^{-n}}\right) \xrightarrow[n\to\infty]{\mathbb{P}} \int_0^t H_s \, dM_s$$

for any $t \geq 0$.

Proof: Assume first that $M \in \mathcal{M}_c^2$ and $|H_t| \leq c \quad \forall t$ (the general case is reduced to this by localization).

For each $n \in \mathbb{N}$, define a process $H^{(n)}$ by

(essentially a simple process)

$$H_t^{(n)} = H_{t_n^-}. \quad \left(\begin{array}{l}\text{as before,} \\ t_n^- := \sup [0,t) \cap 2^{-n}\mathbb{Z}\end{array}\right)$$

By left-continuity of H, we have $H_t^{(n)} \to H_t$ as $n \to \infty$. The stochastic integral of $H^{(n)}$ is

$$\left(H^{(n)} \bullet M\right)_t = \sum_{k=1}^{\lfloor 2^n t \rfloor} H_{(k-1)2^{-n}} \left(M_{k2^{-n}} - M_{(k-1)2^{-n}}\right)$$
$$+ H_{t_n^-} \cdot \left(M_t - M_{t_n^-}\right).$$

The second term tends to zero a.s., since M is continuous and H is bounded. It therefore remains to show that $\left(H^{(n)} \bullet M\right)_t \xrightarrow[n\to\infty]{\mathbb{P}} (H \bullet M)_t$.

For this, observe that

$$\| H^{(n)} - H \|_M^2 = \mathbb{E}\left[\int_0^\infty (H_s^{(n)} - H_s) \, d\langle M,M\rangle_s\right] \xrightarrow[n\to\infty]{} 0$$

by dominated convergence theorem. ⊛

Details: $\langle M,M\rangle_\infty < \infty$ a.s. and $|H_s^{(n)} - H_s| \leq 2c$ so bounded convergence gives a.s.

$$\int_0^\infty (H_s^{(n)} - H_s) \, d\langle M,M\rangle_s \xrightarrow[n\to\infty]{} 0$$

But also $\left| \int_0^\infty (H_s^{(n)} - H_s) \, d\langle M,M\rangle_s \right| \leq 2c \cdot \langle M,M\rangle_\infty$ and $\langle M,M\rangle_\infty \in \mathcal{L}^1(\mathbb{P})$, so we can take expected values and use **dominated conv.**

The Itô isometry then shows

$$\| H^{(n)} \bullet M - H \bullet M \|_{\mathcal{M}_c^2} = \| H^{(n)} - H \|_M \xrightarrow[n\to\infty]{} 0 .$$

Doob's L^2-maximal inequality gives

$$\mathbb{E}\Big[\sup_{t \geq 0} \big((H^{(n)} \bullet M)_t - (H \bullet M)_t \big)^2 \Big]$$

$$\leq 4 \cdot \| H^{(n)} \bullet M - H \bullet M \|_{\mathcal{M}_c^2}^2 \longrightarrow 0$$

In particular, for any $\varepsilon > 0$, Markov's inequality yields, for any $t \geq 0$,

$$\mathbb{P}\Big[\big| (H^{(n)} \bullet M)_t - (H \bullet M)_t \big| \geq \varepsilon \Big]$$

$$= \mathbb{P}\Big[\big| (H^{(n)} \bullet M)_t - (H \bullet M)_t \big|^2 \geq \varepsilon^2 \Big]$$

$$\leq \frac{1}{\varepsilon^2} \mathbb{E}\Big[\big((H^{(n)} \bullet M)_t - (H \bullet M)_t \big)^2 \Big] \leq \frac{4}{\varepsilon^2} \| H^{(n)} \bullet M - H \bullet M \|_{\mathcal{M}_c}^2$$

$$\xrightarrow[n\to\infty]{} 0 .$$

This shows that $(H^{(n)} \bullet M)_t \xrightarrow[n\to\infty]{\mathbb{P}} (H \bullet M)_t$ as claimed.

Consider then the general case: $M \in \mathcal{M}_{c,loc}$ and H locally bounded, adapted, left continuous. Take a sequence $\tau_1, \tau_2, \ldots$ of stopping times which localizes both M and H. Fix $t \geq 0$, $\varepsilon > 0$, and $\delta > 0$. Since $\tau_m \uparrow \infty$ as $m \to \infty$, we can choose a large m s.t. $\mathbb{P}[\tau_m \geq t] < \frac{\delta}{2}$. On the event $\{\tau_m \geq t\}$ the stochastic integral is by definition

$$(H \bullet M)_t(\omega) = \left((H \mathbb{1}_{(0,\tau_m]}) \bullet M^{\tau_m} \right)_t (\omega)$$

$$(\text{for } \omega \text{ s.t. } \tau_m(\omega) \geq t)$$

But $\tilde{M} = M^{\tau_m} \in \mathcal{M}_c^2$, and $\tilde{H} = H \cdot \mathbb{1}_{(0,\tau_m]}$ is bounded, so by the earlier case we have for large enough $n \in \mathbb{N}$ that

$$\mathbb{P}\left[\left| (\tilde{H} \bullet \tilde{M})_t - \sum_{k=1}^{\lfloor 2^n t \rfloor} \tilde{H}_{(k-1)2^{-n}} \left(\tilde{M}_{k2^{-n}} - \tilde{M}_{(k-1)2^{-n}} \right) \right| \geq \varepsilon \right] < \frac{\delta}{2}.$$

On the event $\{\tau_m \geq t\}$ we also have

$$\sum_{k=1}^{\lfloor 2^n t \rfloor} \tilde{H}_{(k-1)2^{-n}} \left(\tilde{M}_{k2^{-n}} - \tilde{M}_{(k-1)2^{-n}} \right) = \sum_{k=1}^{\lfloor 2^n t \rfloor} H_{(k-1)2^{-n}} \left(M_{k2^{-n}} - M_{(k-1)2^{-n}} \right).$$

Therefore we can conclude by union bound:

$$\mathbb{P}\left[\left| (H \bullet M)_t - \sum_{k=1}^{\lfloor 2^n t \rfloor} H_{(k-1)2^{-n}} \left(M_{k2^{-n}} - M_{(k-1)2^{-n}} \right) \right| \geq \varepsilon \right]$$

$$\leq \underbrace{\mathbb{P}[\tau_m < t]}_{< \delta/2} + \underbrace{\mathbb{P}\left[\tau_m \geq t \text{ and } \left| (\tilde{H} \bullet \tilde{M})_t - \sum_k \cdots \right| \geq \varepsilon \right]}_{< \delta/2}$$

$$< \delta.$$

Since $\varepsilon > 0$, $\delta > 0$ were arbitrary, this shows the asserted convergence in probability. $\square$

Theorem 3.12 (Quadratic variation of stochastic integral)

Let $M \in \mathcal{M}_{c,loc}$ and let H be a locally bounded predictable process. Then $H \bullet M \in \mathcal{M}_{c,loc}$ is a continuous local martingale with quadratic variation given by

$$\langle H \bullet M , H \bullet M \rangle_t = \int_0^t H_s^2 \, d\langle M,M\rangle_s .$$

Proof: Let $\tau_1, \tau_2, \ldots$ be a sequence of stopping times which "localizes" (reduces) both H and M, so that $H \cdot \mathbb{1}_{(0,\tau_n]}$ is bounded and $M^{\tau_n} \in \mathcal{M}_c^2$. By Proposition 3.11 we have that

$$(H \bullet M)^{\tau_n} = (H \cdot \mathbb{1}_{(0,\tau_n]}) \bullet M^{\tau_n} \in \mathcal{M}_c^2 .$$

This localization shows that $H \bullet M$ is a continuous local martingale, $H \bullet M \in \mathcal{M}_{c,loc}$.

For the calculation of quadratic variation, assume first (by "localization") that $M \in \mathcal{M}_c^2$ and H is bounded. The isometry property of Theorem 3.9 then gives for any stopping time τ that

$$\mathbb{E}\left[(H \bullet M)_\tau^2\right] = \mathbb{E}\left[\left((H \cdot \mathbb{1}_{(0,\tau]}) \bullet M\right)_\infty^2\right]$$

$$= \mathbb{E}\left[\int_0^\infty H_s^2 \cdot \mathbb{1}_{(0,\tau]}(s)^2 \cdot d\langle M,M\rangle_s\right]$$

$$= \mathbb{E}\left[\int_0^\tau H_s^2 \, d\langle M,M\rangle_s\right].$$

By optional stopping characterization of martingales, Theorem 2.7, (iv) $\Rightarrow$ (i), this implies that $(H \bullet M)_t^2 - \int_0^t H_s^2 \, d\langle M,M\rangle_s$ is a martingale.

Moreover, the process $t \longmapsto \int_0^t H_s^2 \, d\langle M,M\rangle_s$ is increasing and continuous (a.s.). These properties uniquely determine the quadratic variation, so we conclude

$$\langle H\bullet M, H\bullet M\rangle_t = \int_0^t H_s^2 \, d\langle M,M\rangle_s .$$

It remains to lift the assumptions that M is bounded in $\mathcal{L}^2$ and H is bounded. In the general case we first "localize" by stopping times τ_n, $n=1,2,\ldots$. We have

$$\langle H\bullet M, H\bullet M\rangle_t = \lim_{n\to\infty} \langle H\bullet M, H\bullet M\rangle_t^{\tau_n}$$

$$= \lim_{n\to\infty} \langle (H\bullet M)^{\tau_n}, (H\bullet M)^{\tau_n}\rangle_t$$

$$= \lim_{n\to\infty} \langle (H\cdot \mathbb{1}_{(0,\tau_n]})\bullet M^{\tau_n}, (H\cdot \mathbb{1}_{(0,\tau_n]})\bullet M^{\tau_n}\rangle_t$$

by the case treated before

$$= \lim_{n\to\infty} \int_0^t H_s^2 \cdot \mathbb{1}_{(0,\tau_n]}(s)^2 \cdot d\langle M^{\tau_n}, M^{\tau_n}\rangle_s$$

by monotone convergence

$$= \int_0^t H_s^2 \, d\langle M,M\rangle_s \qquad \square$$

The following result is a "chain rule" or "associativity" of stochastic integration.

Theorem 3.13 Let H, K be locally bounded predictable processes and $M \in \mathcal{M}_{c,loc}$. Then we have

$$H\bullet(K\bullet M) = (H\cdot K)\bullet M$$

Concise (abuse of) notation: It is convenient to agree to write $dX_t = H_t \, dM_t$ if

$$X_t - X_0 = \int_0^t H_s \, dM_s \qquad \text{for all } t.$$

(The first equation of "stochastic differentials" is just a shorthand notation for the actual meaning in terms of stochastic integration)

Interpretation of Theorem 3.13 In the above notation, the assertion becomes $H_t \cdot d\left(\int_0^t K_s dM_s\right) = (H_t K_t)\cdot dM_t$

which contains the special case $(H=1)$

$$d\left(\int_0^t K_s\, dM_s\right) = K_t\, dM_t$$

and the fact that we are allowed to multiply such differentials consistently with locally bounded predictable processes.

<u>Proof of Theorem 3.13.</u> We only do the case when H and K are bounded and $M \in \mathcal{M}_c^2$. The general case is obtained by localization.

First consider the case that $H, K \in \mathcal{S}$ are simple processes. By refining time-partitions, we may assume that

$$H_t = \sum_{k=1}^m Z_k \cdot \mathbb{1}_{(t_{k-1}, t_k]}(t)$$

$$K_t = \sum_{k=1}^m W_k \cdot \mathbb{1}_{(t_{k-1}, t_k]}(t)$$

where $0 = t_0 < t_1 < \dots < t_m$ and Z_k, W_k are bounded $\mathcal{F}_{t_{k-1}}$-measurable.

But since both sides are bilinear in H and K, it suffices to consider the cases

$$H_t = Z \cdot \mathbb{1}_{(a,b]}(t), \qquad K_t = W \cdot \mathbb{1}_{(c,d]}(t)$$

where either $(a,b] = (c,d]$ or $(a,b] \cap (c,d] = \emptyset$. (and Z and W are bounded $\mathcal{F}_a$- and $\mathcal{F}_c$-measurable).

In the latter case both sides of the asserted equality are zero, so we consider $(a,b] = (c,d]$.

In that case we observe

$$(K \bullet M)_t = W \cdot (M_{t\wedge b} - M_{t \wedge a})$$

and then

$$\begin{aligned}(H \bullet (K \bullet M))_t &= Z \cdot \left((K\bullet M)_{t\wedge b} - (K \bullet M)_{t\wedge a}\right) \\ &= Z \cdot W \cdot (M_{t\wedge b} - M_{t\wedge a}) = ((HK)\bullet M)_t .\end{aligned}$$

This is the asserted equality, so we now conclude that it holds for all $H, K \in \mathcal{S}$ simple processes.

For general bounded predictable H, K, we find approximating sequences $H^{(1)}, H^{(2)}, \ldots \in \mathcal{S}$ and $K^{(1)}, K^{(2)}, \ldots \in \mathcal{S}$ such that $|H_t^{(n)}| \leq c_H$, $|K_t^{(n)}| \leq c_K$ $\forall n$ $\forall t$ and

$$\| H^{(n)} - H \|_M \longrightarrow 0$$

$$\| K^{(n)} - K \|_M \longrightarrow 0 \qquad \text{as } n \to \infty.$$

We first obtain an upper bound on $\| H \|_{K \bullet M}$ as follows:

$$\| H \|_{K \bullet M}^2 = \mathbb{E}\left[\int_0^\infty H_s^2 \, d\underbrace{\langle K \bullet M, K \bullet M \rangle_s}_{= \int_0^s K_u^2 \, d\langle M, M \rangle_u} \right]$$

exercise: associativity for finite variation integrals

$$= \mathbb{E}\left[\int_0^\infty H_s^2 K_s^2 \, d\langle M, M \rangle_s \right]$$

$$\leq \min \left\{ c_H^2 \cdot \| K \|_M^2 , \; c_K^2 \cdot \| H \|_M^2 \right\}.$$

For simple processes we have established

$$H^{(n)} \bullet (K^{(n)} \bullet M) = (H^{(n)} K^{(n)}) \bullet M$$

and we now want to take the limit as $n \to \infty$. For the RHS,

$$\| (H^{(n)} K^{(n)}) \bullet M - (HK) \bullet M \|_{\mathcal{M}_c^2}^2$$

$$= \mathbb{E}\left[\int_0^\infty (H_s^{(n)} K_s^{(n)} - H_s K_s) \, d\langle M, M \rangle_s \right] \xrightarrow[n \to \infty]{} 0$$

by dominated convergence.

For the LHS,

$$\| H^{(n)} \bullet (K^{(n)} \bullet M) - H \bullet (K \bullet M) \|_{\mathcal{M}_c^2}$$

$$\le \| (H^{(n)} - H) \bullet (K^{(n)} \bullet M) \|_{\mathcal{M}_c^2} + \| H \bullet ((K^{(n)} - K) \bullet M) \|_{\mathcal{M}_c^2}$$

Itô isometry

$$\le \| H^{(n)} - H \|_{K^{(n)} \bullet M} + \| H \|_{(K^{(n)} - K) \bullet M}$$

$$\le c_K \cdot \underbrace{\| H^{(n)} - H \|_M}_{\to 0} + c_H \cdot \underbrace{\| K^{(n)} - K \|_M}_{\to 0} \xrightarrow[n \to \infty]{} 0.$$

This finishes the proof. $\square$

QUADRATIC COVARIATION

As a tool, we need a close cousin of quadratic variation, which is defined for a pair of processes and is bilinear in them.

Def: Let $M, N \in \mathcal{M}_{c,loc}$. Define, for $t \geq 0$,

$$\langle M,N\rangle_t = \frac{1}{4}\Big(\underbrace{\langle M+N, M+N\rangle_t}_{\text{quadratic variation of } M+N} - \underbrace{\langle M-N, M-N\rangle_t}_{\text{quadratic variation of } M-N}\Big)$$

We call $\langle M,N\rangle = (\langle M,N\rangle_t)_{t\in\mathbb{R}_+}$ the quadratic covariation process of M and N.

The following properties are straightforward consequences of the definition and Theorems 3.5 and 3.7 about quadratic variation. (Uniqueness uses Theorem 2.13., too.)

Theorem 3.16 (Quadratic covariation)

(i): $\langle M,N\rangle$ is the unique (up to indistinguishability) finite variation process such that $M\cdot N - \langle M,N\rangle$ is a continuous local mgale.

(ii): $$\sum_{k=1}^{\lfloor 2^n t\rfloor}\left(M_{k2^{-n}} - M_{(k-1)2^{-n}}\right)\left(N_{k2^{-n}} - N_{(k-1)2^{-n}}\right) \xrightarrow[n\to\infty]{\mathbb{P}} \langle M,N\rangle_t .$$

(iii): If $M, N \in \mathcal{M}_c^2$, then $MN - \langle M,N\rangle$ is a uniformly integrable martingale.

(iv): $\langle M,N\rangle = \langle N,M\rangle$ and the quadratic covariation is bilinear in M and N.

Theorem 3.17 (Kunita-Watanabe identity)

Let $M, N \in \mathcal{M}_{c,loc}$ and let H be a locally bounded predictable process. Then we have

$$\langle H \bullet M, N \rangle_t = \int_0^t H_s \, d\langle M,N \rangle_s$$

stochastic integral

finite variation integral

Proof: Assume that $M, N \in \mathcal{M}_c^2$ and $|H_t| \leq c \quad \forall t$ (in the general case we "localize", to reduce to this).

Note that $t \longmapsto \int_0^t H_s \, d\langle M,N\rangle_s$ is of finite variation (exercise), so by Theorem 3.16 (i), it suffices to show that

$t \mapsto (H \bullet M)_t \cdot N_t - \int_0^t H_s \, d\langle M,N\rangle_s$ is a cont. loc. mgale.

Continuity is clear, and we in fact show that this is a martingale by using the optional stopping characterization (Thm 2.7.).

So we want to show that for all (bounded) stopping times τ we have

$$\mathbb{E}\left[(H \bullet M)_\tau N_\tau - \int_0^\tau H_s \, d\langle M,N\rangle_s \right] = 0.$$

By considering the stopped processes H^τ, M^τ, N^τ instead, it suffices to show that

$$\mathbb{E}\left[(H \bullet M)_\infty N_\infty \right] = \mathbb{E}\left[\int_0^\infty H_s \, d\langle M,N\rangle_s \right] \quad (*).$$

Assume first that $H = Z \cdot \mathbb{1}_{(s,t]}$ with Z bounded, $\mathcal{F}_s$-measurable. Then the LHS of $(*)$ is

$$\begin{aligned}
\text{LHS} &= \mathbb{E}\left[Z \cdot (M_t - M_s) \cdot N_\infty \right] \\
&= \mathbb{E}\Big[Z \cdot M_t \cdot \underbrace{\mathbb{E}[N_\infty \mid \mathcal{F}_t]}_{= N_t} \Big] - \mathbb{E}\Big[Z \cdot M_s \cdot \underbrace{\mathbb{E}[N_\infty \mid \mathcal{F}_s]}_{= N_s} \Big] \\
&= \mathbb{E}\left[Z \cdot (M_t N_t - M_s N_s) \right] \\
&= \mathbb{E}\Big[Z \cdot \underbrace{\mathbb{E}\left[M_t N_t - M_s N_s \mid \mathcal{F}_s \right]}_{\substack{= \mathbb{E}[\langle M,N\rangle_t - \langle M,N\rangle_s \mid \mathcal{F}_s] \\ \text{since } M\cdot N - \langle M,N\rangle \text{ is a martingale}}} \Big] \\
&= \mathbb{E}\left[Z \cdot \left(\langle M,N\rangle_t - \langle M,N\rangle_s \right) \right] \\
&= \mathbb{E}\left[\int_s^t Z \cdot d\langle M,N\rangle_s \right] = \mathbb{E}\left[\int_0^\infty H_s \, d\langle M,N\rangle_s \right] = \text{RHS}.
\end{aligned}$$

This proves ⊛ in the case of $H = Z \cdot \mathbb{1}_{(s,t]}$.
Both sides are linear, so we get ⊛ also for any simple process $H \in \mathcal{S}$.
If H is predictable and bounded, then we can find $H^{(1)}, H^{(2)}, \ldots \in \mathcal{S}$ s.t. $|H^{(n)}_t| \leq C \quad \forall n \ \forall t$, and $\|H - H^{(n)}\|_M \xrightarrow[n\to\infty]{} 0$.

We conclude by the calculation

$$\begin{aligned}
&\mathbb{E}\left[(H \bullet M)_\infty N_\infty \right] \\
&= \mathbb{E}\left[\left(\lim_{n\to\infty} (H^{(n)} \bullet M)_\infty \right) \cdot N_\infty \right] && \text{(def. of integral and ...)} \\
&= \lim_{n\to\infty} \mathbb{E}\left[(H^{(n)} \bullet M)_\infty \cdot N_\infty \right] && \text{(dominated convergence)} \\
&= \lim_{n\to\infty} \mathbb{E}\left[\int_0^\infty H^{(n)}_s \, d\langle M,N\rangle_s \right] && \text{(⊛ for simple process)} \\
&= \mathbb{E}\left[\int_0^\infty H_s \, d\langle M,N\rangle_s \right]. && \text{(dominated convergence)}
\end{aligned}$$

This proves ⊛. □

STOCHASTIC INTEGRATION W.R.T. SEMIMARTINGALES

We have defined integration w.r.t. two essentially complementary types of continuous stochastic processes:

- continuous local martingales
- processes of finite variation

As a final step, we combine the two.

Def: A stochastic process $X = (X_t)_{t \in \mathbb{R}_+}$ is a continuous semimartingale if it can be written as
$$X_t = M_t + A_t$$
where $M = (M_t)_{t \in \mathbb{R}_+}$ is a continuous local martingale and $A = (A_t)_{t \in \mathbb{R}_+}$ is a finite variation process.

(Recall: Our convention is that finite variation processes are started from zero, $A_0 = 0$.)

The decomposition $X = M + A$ of a semimartingale to a local mgale and a finite variation process is called the Doob-Meyer decomposition. Note that the decomposition is unique (up to indistinguishability): if
$$X = M + A \quad \text{and} \quad X = \tilde{M} + \tilde{A}$$
where $M, \tilde{M} \in \mathcal{M}_{c,loc}$ and $A, \tilde{A}$ are finite variation processes, then
$$M - \tilde{M} = (X - A) - (X - \tilde{A}) = \tilde{A} - A$$
is both a local martingale (LHS expression) and a finite variation process (RHS expression), so by Theorem 2.13 it is zero (up to indistinguishability) — i.e. $M = \tilde{M}$ and $A = \tilde{A}$.

The most general setup for integration that we will use is the following: suppose that

- $H = (H_t)_{t \in (0,\infty)}$ is a locally bounded predictable process
- $X = (X_t)_{t \in \mathbb{R}_+}$ is a continuous semimartingale with Doob-Meyer decomp. $X = M + A$.

Then we define

$$(H \bullet X)_t = \int_0^t H_s \, dX_s = \underbrace{\int_0^t H_s \, dM_s}_{\text{integral w.r.t. local mgale}} + \underbrace{\int_0^t H_s \, dA_s}_{\text{integral w.r.t finite variation process}} .$$

$$= (H \bullet M)_t + (H \bullet A)_t .$$

Note that $H \bullet M$ is itself a continuous local martingale (see definition of integration w.r.t. $M \in \mathcal{M}_{c,loc}$) and $H \bullet A$ is itself a finite variation process (exercise on associativity of finite variation integration).

Therefore $H \bullet X$ is a continuous semimartingale with Doob-Meyer decomposition

$$H \bullet X = H \bullet M + H \bullet A$$

where $X = M + A$ is the Doob-Meyer decomposition of X.

<u>Quadratic variation and covariation of semimartingales</u>

Let X, $\tilde{X}$ be two continuous semimartingales with Doob-Meyer decompositions $X = M + A$, $\tilde{X} = \tilde{M} + \tilde{A}$. We define $\langle X \rangle = \langle M \rangle$ and $\langle X, \tilde{X} \rangle = \langle M, \tilde{M} \rangle$, i.e. the quadratic variation and covariation of semimgales is that of their local martingale parts.

ITÔ'S FORMULA

The most important practical formula about stochastic integration is a "change of variables" result known as Itô's formula. We first present a one-variable version and later a multivariable generalization.

(The result should be compared with the following: if $a : \mathbb{R}_+ \to \mathbb{R}$ is a finite variation function and $f : \mathbb{R} \to \mathbb{R}$ continuously differentiable, then

$$f(a_t) - f(a_0) = \int_0^t f'(a_s)\, da_s .$$

In stochastic integration there is a "higher order correction term" coming from quadratic variation.)

Theorem 3.22 (Itô's formula, one-variable case)

Let $X = (X_t)_{t \in \mathbb{R}_+}$ be a continuous semimartingale and $f : \mathbb{R} \to \mathbb{R}$ a twice continuously differentiable function ($f \in C^2(\mathbb{R};\mathbb{R})$). Then we have, for all $t \geq 0$

$$f(X_t) - f(X_0) = \int_0^t f'(X_s)\, dX_s + \frac{1}{2} \int_0^t f''(X_s)\, d\langle X, X\rangle_s .$$

Remark: Why is this a "change of variables formula"? The convenient abuse of notation is to write the formula in terms of "stochastic diferentials" as

$$d\, f(X_t) = f'(X_t) \cdot dX_t + \frac{1}{2} f''(X_t)\, d\langle X,X\rangle_t .$$

The precise meaning of this is, of course, just the above formula for $f(X_t) - f(X_0)$. But we now interpret this as a change of variables from X_t to $f(X_t)$.

We postpone the complete proof to a later stage. The following sketch could, however, be turned into a rigorous proof with some amount of work.

Sketch of a proof of Itô's formula:

Let us divide the interval $[0,t]$ to small pieces by $0 = t_0 < t_1 < \cdots < t_m = t$. Then we can write a telescopic sum

$$f(X_t) - f(X_0) = \sum_{j=1}^{m} \left(f(X_{t_j}) - f(X_{t_{j-1}}) \right).$$

By continuity, the terms $f(X_{t_j}) - f(X_{t_{j-1}})$ are small, if the division is fine, and by assumed "smoothness" of f, it makes sense to form a Taylor approximation

$$f(X_{t_j}) - f(X_{t_{j-1}}) \approx f'(X_{t_{j-1}}) \cdot (X_{t_j} - X_{t_{j-1}}) + \frac{1}{2} f''(X_{t_{j-1}}) \cdot (X_{t_j} - X_{t_{j-1}})^2 + \text{error terms}.$$

The first terms in the approximations add up to

$$\sum_{j=1}^{m} f'(X_{t_{j-1}}) \cdot (X_{t_j} - X_{t_{j-1}}) \approx \int_0^t f'(X_s)\, dX_s$$

(for example according to Proposition 3.14, we have convergence in probability as the mesh of the division tends to zero).

The second terms in the approximations add up to

$$\sum_{j=1}^{m} \frac{1}{2} f''(X_{t_{j-1}}) \cdot (X_{t_j} - X_{t_{j-1}})^2 \approx \frac{1}{2} \int_0^t f''(X_s)\, d\langle X, X \rangle_s.$$

This explains the form of Itô's formula. One strategy of a proof is to control the convergence of these approximations and of error terms.

With straightforward modifications to this sketch of proof, it is not difficult to see that the multivariable version should be:

Theorem 3.22 (Itô's formula)

Let $X^{(1)}, X^{(2)}, \dots, X^{(d)}$ be continuous semimgales and $f : \mathbb{R}^d \to \mathbb{R}$ a twice continuously differentiable function ($f \in C^2(\mathbb{R}^d;\mathbb{R})$). Then we have

$$f(X_t^{(1)}, \dots, X_t^{(d)}) - f(X_0^{(1)}, \dots, X_0^{(d)})$$
$$= \sum_{i=1}^{d} \int_0^t (\partial_i f)(X_s^{(1)}, \dots, X_s^{(d)}) \cdot dX_s^{(i)}$$
$$+ \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t (\partial_i \partial_j f)(X_s^{(1)}, \dots, X_s^{(d)}) \cdot d\langle X^{(i)}, X^{(j)} \rangle_s$$

($\partial_i f$ and $\partial_i \partial_j f$ denote partial derivatives of $f : \mathbb{R}^d \to \mathbb{R}$.)

We postpone the proof and first illustrate some typical applications.

APPLICATIONS OF ITÔ'S FORMULA

In typical applications, Itô's formula is used to find a (local) martingale, which is then used in optional stopping theorem to draw a probabilistic conclusion.

We are already quite familiar with how simple martingales for Brownian motion can be used, but let us rephrase some known applications

Examples of (local) martingales for Brownian motion

Let $B = (B_t)_{t \in \mathbb{R}_+}$ be standard Brownian motion.

1°) B itself is a martingale

(We could apply Itô's formula with $f(x) = x$ to see that it is at least a local martingale, but this is slightly silly...)

If $a \leq 0 \leq b$, and $\tau = \inf \{ t \geq 0 \mid B_t \in \{a, b\} \}$ is the hitting time of endpoints of $[a,b]$, then optional stopping for the bounded martingale B^τ gives (since $\tau < \infty$ a.s.)

$$0 = \mathbb{E}[B_0] = \mathbb{E}[B_\tau]$$
$$= b \cdot \mathbb{P}[B_\tau = b] + a \cdot \underbrace{\mathbb{P}[B_\tau = a]}_{= 1 - \mathbb{P}[B_\tau = b]}$$
$$= (b-a)\, \mathbb{P}[B_\tau = b] + a,$$

which allows us to solve for $\mathbb{P}[B_\tau = b]$ and get Gambler's ruin formula

$$\mathbb{P}[B_\tau = b] = \frac{-a}{b-a}.$$

2°) $B_t^2 - t$ is a (local) martingale

(Consider $X_t^{(1)} = B_t$ and $X_t^{(2)} = t$ and $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x,s) = x^2 - s$.

Then

$$B_t^2 - t = f(B_t, t) - f(0,0)$$
$$= \int_0^t \underbrace{(\partial_x f)(B_s, s)}_{=2B_s} \, dB_s + \frac{1}{2} \int_0^t \underbrace{(\partial_x^2 f)(B_s,s)}_{=2} \, d\langle B,B\rangle_s$$
$$+ \int_0^t \underbrace{(\partial_s f)(B_s,s)}_{=-1} \cdot ds .$$

Since $\langle B,B\rangle_s = s$, the last two terms cancel. The first term is a local martingale, because it is the integral of a predictable process against the local martingale B.)

If $a \leq 0 \leq b$ and τ is the hitting time of endpoints of $[a,b]$ as above, then the fact that $B_{t\wedge\tau}^2 - (t \wedge \tau)$ is a true martingale yields, for any $t \geq 0$,

$$0 = B_0^2 - 0 = \mathbb{E}\left[B_{t\wedge\tau}^2 - (t\wedge\tau) \right],$$

which gives

$$\mathbb{E}[(t\wedge\tau)] = \mathbb{E}[B_{t\wedge\tau}^2].$$

Letting $t \to \infty$, the LHS becomes (monotone conv.)

$$\mathbb{E}[(t\wedge\tau)] \xrightarrow[t\to\infty]{} \mathbb{E}[\tau]$$

and the RHS becomes (bounded conv.)

$$\mathbb{E}[B_{t\wedge\tau}^2] \longrightarrow b^2 \cdot \mathbb{P}[B_\tau = b] + a^2 \mathbb{P}[B_\tau = a].$$

Combining with gambler's ruin we can solve

$$\mathbb{E}[\tau] = b^2 \cdot \frac{-a}{b-a} + a^2 \cdot \frac{b}{b-a} = -ab.$$

3°) Let again $X^{(1)}_t = B_t$ and $X^{(2)}_t = t$.

Consider a polynomial $f : \mathbb{R}^2 \to \mathbb{R}$ of the form $f(x,s) = \alpha \cdot x^4 + \beta \cdot x^2 s + s^2$.

($\alpha, \beta \in \mathbb{R}$ coefficients)

Itô's formula tells that

$$f(B_t, t) = f(B_t, t) - f(B_0, 0)$$

$$= \int_0^t (\partial_x f)(B_s, s)\, dB_s \quad \leftarrow \text{local mgale}$$

$$\left.\begin{aligned} &+ \frac{1}{2} \int_0^t (\partial_{xx} f)(B_s, s)\, d\langle B, B\rangle_s \\ &+ \int_0^t (\partial_s f)(B_s, s)\, ds . \end{aligned}\right\} \text{can be made to cancel}$$

It is possible to choose the coefficients $\alpha, \beta \in \mathbb{R}$ in such a way that the last two terms cancel and thus $f(B_t, t)$ is a local martingale.

For $a \leq 0 \leq b$, letting τ be the hitting time of the endpoints of $[a,b]$, we have that $\big(f(B_{t\wedge\tau}, t\wedge\tau)\big)_{t \in \mathbb{R}_+}$ is a local mgale which is bounded on the time interval $[0,t]$. Optional stopping can then be used to calculate

$$\mathbb{E}[\tau^2] \quad \text{and} \quad \mathrm{Var}(\tau).$$

(exercise).

Recurrence and transience of d-dim. Brownian motion

Let us consider the d-dimensional Brownian motion started from $(x^{(1)}, \dots, x^{(d)}) \in \mathbb{R}^d$.
It is the process

$$t \longmapsto (X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(d)})$$

constructed from d independent standard Brownian motions $B^{(1)}, B^{(2)}, \dots, B^{(d)}$ via

$$X_t^{(j)} = x^{(j)} + B_t^{(j)} \qquad \text{for } j = 1, 2, \dots, d.$$

Of course the coordinates $X^{(1)}, X^{(2)}, \dots, X^{(d)}$ are continuous semimartingales, so for any

$$f : \mathbb{R}^d \longrightarrow \mathbb{R} \qquad \text{twice cont. dif}$$

we have, by Itô's formula

$$f(X_t^{(1)}, \dots, X_t^{(d)}) - f(X_0^{(1)}, \dots, X_0^{(d)})$$

$$= \sum_{i=1}^{d} \int_0^t (\partial_i f)(X_s^{(1)}, \dots, X_s^{(d)})\, dX_s^{(i)} \qquad dX_s^{(i)} = dB_s^{(i)}$$

$$+ \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t (\partial_i \partial_j f)(X_s^{(1)}, \dots, X_s^{(d)})\, \underbrace{d\langle X^{(i)}, X^{(j)} \rangle_s}$$

$$d\langle X^{(i)}, X^{(j)} \rangle_s = d\langle B^{(i)}, B^{(j)} \rangle_s = \delta_{ij} \cdot ds$$

(exercise: covariations of independent Brownian motions)

$$= \sum_{i=1}^{d} \int_0^t (\partial_i f)(X_s^{(1)}, \dots, X_s^{(d)})\, dB_s^{(i)} \qquad \longleftarrow \text{local mgale}$$

$$+ \frac{1}{2} \sum_{i=1}^{d} \int_0^t (\partial_i^2 f)(X_s^{(1)}, \dots, X_s^{(d)})\, ds .$$

$$\left(= \int_0^t \nabla f(X_s) \cdot d\vec{B}_s + \frac{1}{2} \int_0^t (\Delta f)(X_s)\, ds \qquad \text{in shorthand notation} \right)$$

The first term is a local martingale (as an integral w.r.t. local mgale). The second term vanishes if

$$\Delta f = \sum_{i=1}^{d} \partial_i^2 f = 0 ,$$

i.e. if f is a harmonic function.

Let us try a radial function

$$f(x^{(1)}, \ldots, x^{(d)}) = g(r(x^{(1)}, \ldots, x^{(d)}))$$

where $r(x^{(1)}, \ldots, x^{(d)}) = \|\bar{x}\| = \sqrt{(x^{(1)})^2 + \cdots + (x^{(d)})^2}$ and $g : [0,\infty) \to \mathbb{R}$. We will in fact allow $g : (0,\infty) \to \mathbb{R}$, so the function f is defined except at the origin,

$$f : \mathbb{R}^d \setminus \{(0,\ldots,0)\} \longrightarrow \mathbb{R}$$

but we will stop before hitting the origin, so this is not an issue. For harmonicity, we have to require

$$0 = (\Delta f)(x^{(1)}, \ldots, x^{(d)}) = \sum_{j=1}^{d} (\partial_j^2 f)(x^{(1)}, \ldots, x^{(d)}) .$$

To apply the chain rule, calculate first

$$\partial_j \, r(x^{(1)}, \ldots, x^{(d)}) = \partial_j \|\bar{x}\| = \partial_j \Big(\sum_{i=1}^{d} (x^{(i)})^2 \Big)^{1/2}$$

$$= 2x^{(j)} \cdot \frac{1}{2} \Big(\sum_{i=1}^{d} (x^{(i)})^2 \Big)^{-1/2} = \frac{x^{(j)}}{\|\bar{x}\|} .$$

Then

$$\partial_j f(x^{(1)}, \ldots, x^{(d)}) = g'(r(x^{(1)}, \ldots, x^{(d)})) \cdot \partial_j r(x^{(1)}, \ldots, x^{(d)})$$

$$= g'(r(x^{(1)}, \ldots, x^{(d)})) \cdot \frac{x^{(j)}}{\|\bar{x}\|}$$

and

$$\partial_j^2 f(x^{(1)},\dots,x^{(d)}) = g''(r(x^{(1)},\dots,x^{(d)}))\cdot\frac{(x^{(j)})^2}{\|\vec{x}\|^2} + g'(r(x^{(1)},\dots,x^{(d)}))\cdot\left(\frac{1}{\|\vec{x}\|} - \frac{(x^{(j)})^2}{\|\vec{x}\|^3}\right)$$

so

$$\Delta f(x^{(1)},\dots,x^{(d)}) = \sum_{j=1}^{d}(\partial_j^2 f)(x^{(1)},\dots,x^{(d)})$$

$$= g''(r(x^{(1)},\dots,x^{(d)})) + g'(r(x^{(1)},\dots,x^{(d)}))\cdot\left(\frac{d}{\|\vec{x}\|} - \frac{1}{\|\vec{x}\|}\right).$$

We therefore want g to solve

$$g''(r) + \frac{d-1}{r}\cdot g'(r) = 0$$

and for $d \neq 2$ a solution is $g(r) = r^{2-d}$

(Indeed, then $g''(r) = (2-d)(1-d)\,r^{-d}$ and $g'(r) = (2-d)\,r^{1-d}$)

(The other linearly independent solution is constant, which does not give us very interesting local martingales, so let us disregard that.)

At this stage we conclude that for $d \geq 3$, the process defined by

$$M_t = f(X_t^{(1)},\dots,X_t^{(d)}) = R_t^{2-d}$$

where $R_t = \sqrt{(X_t^{(1)})^2 + \dots + (X_t^{(d)})^2}$

is a local martingale (if stopped before hitting the origin).

Let now $0 < \varepsilon \leq \|\vec{X}_0\| \leq \ell$, and consider the stopping times

$$\tau_\varepsilon = \inf\{t \geq 0 \mid R_t = \varepsilon\}$$

$$\tau_\ell = \inf\{t \geq 0 \mid R_t = \ell\}.$$

and $\tau = \tau_\varepsilon \wedge \tau_\ell$.

Then the stopped local m'gale M^τ is bounded, and thus actually a true martingale. We have $\tau < \infty$ a.s. (this follows from noticing that e.g. the first coordinate process exits the interval $[-\ell, \ell]$ in a.s. finite time). Optional stopping theorem gives

$$\|\vec{X}_0\|^{2-d} = R_0^{2-d} = M_0 = \mathbb{E}[M_\tau]$$
$$= \varepsilon^{2-d} \cdot \mathbb{P}[R_\tau = \varepsilon] + \ell^{2-d} \cdot \underbrace{\mathbb{P}[R_\tau = \ell]}_{= 1 - \mathbb{P}[R_\tau = \varepsilon]}$$
$$= \mathbb{P}[R_\tau = \varepsilon] \cdot \left(\varepsilon^{2-d} - \ell^{2-d}\right) - \ell^{2-d}$$

So we can solve

$$\mathbb{P}[R_\tau = \varepsilon] = \frac{\|\vec{X}_0\|^{2-d} - \ell^{2-d}}{\varepsilon^{2-d} - \ell^{2-d}}.$$

$\boxed{\text{Note: By letting } \ell \to \infty \text{ we get } \mathbb{P}[\tau_\varepsilon < \infty] = \left(\frac{\|\vec{X}_0\|}{\varepsilon}\right)^{2-d}}$

Note that as $\varepsilon \downarrow 0$ the RHS tends to 0 (we use $d > 2$). This indicates that the process can not reach the origin:

Proposition: If $\vec{X}_0 \neq \vec{0} \in \mathbb{R}^d$, $d \geq 3$, then

$$\mathbb{P}[\vec{X}_t = \vec{0} \text{ for some } t \geq 0] = 0.$$

Proof: The stopping time $\tau_0 = \inf\{t \geq 0 \mid \vec{X}_t = \vec{0}\}$ is (by continuity of paths of $t \mapsto \vec{X}_t$) the increasing limit (choose $\varepsilon = 1/n$) of

$$\tau_{1/n} \uparrow \tau_0 \qquad \text{as } n \to \infty.$$

For any $\ell \geq \|\vec{X}_0\|$ we have

$$\mathbb{P}[\tau_0 < \tau_\ell] = \mathbb{P}[\tau_{1/n} < \tau_\ell \quad \forall n \in \mathbb{N}]$$

$$\leq \mathbb{P}[\tau_{1/n} < \tau_\ell] \qquad \text{(for any } n\in\mathbb{N}\text{)}$$

$$= \frac{\|\vec{x}_0\|^{2-d} - \ell^{2-d}}{(1/n)^{2-d} - \ell^{2-d}} \xrightarrow[n\to\infty]{} 0$$

so $\mathbb{P}[\tau_0 < \tau_\ell] = 0$. By union bound

$$\mathbb{P}[\exists \ell \in \mathbb{N} : \tau_0 < \tau_\ell] \leq \sum_{\ell\in\mathbb{N}} \underbrace{\mathbb{P}[\tau_0 < \tau_\ell]}_{=0} = 0.$$

But this implies that $\tau_0 = +\infty$ almost surely, since if $\tau_0(\omega) < \infty$ then there exists (by continuity of $t \mapsto R_t(\omega)$ and compactness of $[0, \tau_0(\omega)]$) some ℓ s.t. $\tau_0(\omega) < \tau_\ell(\omega)$. $\square$

<u>Corollary</u> Let $\vec{B} = (B^{(1)}, \dots, B^{(d)})$ be the d-dimensional Brownian motion with $d \geq 3$. Then for any $t_0 > 0$ we have the non-recurrence

$$\mathbb{P}[\vec{B}_t = \vec{0} \text{ for some } t \geq t_0] = 0.$$

<u>Proof</u>: Apply Markov-property at time t_0, and note that $\vec{B}_{t_0} \neq \vec{0}$ almost surely. This says that the "future after t_0" is a BM started away from the origin, and we saw that it never hits the origin. $\square$

<u>Remark</u> The same "non-recurrence" holds also for $d=2$. The next "transience" result, however requires $d>2$.

We can in fact get something better.

Theorem 4.12 (b): If $d \geq 3$, then $R_t = \|\vec{B}_t\| \to \infty$ almost surely as $t \to \infty$.

Proof Consider $\tau_n = \inf\{t \geq 0 \mid R_t = n\}$ for $n = 2, 3, 4, \ldots$. Define the events

$$A_n = \{\forall\, t \geq \tau_n : R_t > \sqrt{n}\}.$$

By strong Markov property at the stopping time τ_n, we get

$$\mathbb{P}[A_n^c] = \mathbb{P}_{\vec{B}_{\tau_n}}[\tau_{\sqrt{n}} < \infty] = \left(\frac{n}{\sqrt{n}}\right)^{2-d}$$

earlier calculation of $\lim\limits_{\ell \to \infty} \mathbb{P}[\tau_\varepsilon < \tau_\ell]$ with $\varepsilon = \sqrt{n}$.

$$= n^{1-d/2} \xrightarrow[n \to \infty]{} 0 \qquad \text{since } d > 2.$$

Therefore by reverse Fatou's lemma

$$\mathbb{P}[\limsup_n A_n] \geq \limsup_n \mathbb{P}[A_n]$$

$$= \limsup_n (1 - \mathbb{P}[A_n^c]) = 1.$$

In other words, A_n occurs infinitely often, almost surely. Since each τ_n is finite, a.s., this implies that for arbitrary $r_0 > 0$ and all large enough t we have $R_t > r_0$. Thus $R_t \to \infty$.

$\square$

By contrast, for one-dimensional Brownian motion we have:

Theorem 4.12 (a) For the standard BM $B = (B_t)_{t \in \mathbb{R}}$ we have for any $t_0 > 0$: $\mathbb{P}[B_t = 0 \text{ for some } t \geq t_0] = 1$.

PROOF OF ITÔ'S FORMULA

It is possible to prove Itô's formula with the straightforward approach sketched after its statement. We adopt a different strategy, however. A key component of this strategy is

Theorem 3.21 (Integration by parts)

Let $X = (X_t)_{t \in \mathbb{R}_+}$ and $Y = (Y_t)_{t \in \mathbb{R}_+}$ be continuous semimartingales. Then

$$X_t Y_t - X_0 Y_0 = \int_0^t X_s \, dY_s + \int_0^t Y_s \, dX_s + \langle X, Y \rangle_t .$$

Proof: Since both sides are continuous in t, it suffices to consider $t = k_0 \cdot 2^{-n_0}$ for some $k_0 \in \mathbb{Z}_{\geq 0}$, $n_0 \in \mathbb{N}$.

Note that

$$X_u Y_u - X_s Y_s = X_s (Y_u - Y_s) + Y_s (X_u - X_s) + (X_u - X_s) \cdot (Y_u - Y_s)$$

so for $n \geq n_0$

$$X_t Y_t - X_0 Y_0 = \sum_{k=1}^{k_0 \cdot 2^{n-n_0}} \left(X_{k2^{-n}} Y_{k2^{-n}} - X_{(k-1)2^{-n}} Y_{(k-1)2^{-n}} \right)$$

$$= \sum_{k=1}^{k_0 \cdot 2^{n-n_0}} \Big\{ X_{(k-1)2^{-n}} \left(Y_{k2^{-n}} - Y_{(k-1)2^{-n}} \right) + Y_{(k-1)2^{-n}} \left(X_{k2^{-n}} - X_{(k-1)2^{-n}} \right) + \left(X_{k2^{-n}} - X_{(k-1)2^{-n}} \right) \left(Y_{k2^{-n}} - Y_{(k-1)2^{-n}} \right) \Big\}$$

$$\xrightarrow[n \to \infty]{\mathbb{P}} \int_0^{k_0 2^{-n_0}} X_s \, dY_s + \int_0^{k_0 2^{-n_0}} Y_s \, dX_s + \langle X, Y \rangle_{k_0 2^{-n_0}}$$

by Proposition 3.14 (and a similar result for finite variation integrals) and the lemma below. $\square$

Lemma For two continuous semimartingales $X = (X_t)_{t \in \mathbb{R}_+}$ and $\tilde{X} = (\tilde{X}_t)_{t \in \mathbb{R}_+}$ and any $t \geq 0$ we have

$$\sum_{k=1}^{\lfloor 2^n t \rfloor} \left(X_{k2^{-n}} - X_{(k-1)2^{-n}}\right)\left(\tilde{X}_{k2^{-n}} - \tilde{X}_{(k-1)2^{-n}}\right) \xrightarrow[n \to \infty]{\mathbb{P}} \langle X, \tilde{X} \rangle_t$$

Proof: Let $X = M + A$ and $\tilde{X} = \tilde{M} + \tilde{A}$ be the Doob-Meyer decompositions of the two semimartingales, and recall that by definition their quadratic covariation is

$$\langle X, \tilde{X} \rangle_t = \langle M, \tilde{M} \rangle_t .$$

Proposition 3.16. establishes the asserted result for continuous local martingales $M, \tilde{M}$ (as a consequence of Theorem 3.5 about quadratic variation). The remaining task is to verify that adding finite variation processes does not alter the limit.

Write

$$\begin{aligned}
&\left(X_{k2^{-n}} - X_{(k-1)2^{-n}}\right)\left(\tilde{X}_{k2^{-n}} - \tilde{X}_{(k-1)2^{-n}}\right) \\
= {} & \left(M_{k2^{-n}} - M_{(k-1)2^{-n}}\right)\left(\tilde{M}_{k2^{-n}} - \tilde{M}_{(k-1)2^{-n}}\right) \\
&+ \left(M_{k2^{-n}} - M_{(k-1)2^{-n}}\right)\left(\tilde{A}_{k2^{-n}} - \tilde{A}_{(k-1)2^{-n}}\right) \\
&+ \left(A_{k2^{-n}} - A_{(k-1)2^{-n}}\right)\left(\tilde{M}_{k2^{-n}} - \tilde{M}_{(k-1)2^{-n}}\right) \\
&+ \left(A_{k2^{-n}} - A_{(k-1)2^{-n}}\right)\left(\tilde{A}_{k2^{-n}} - \tilde{A}_{(k-1)2^{-n}}\right) .
\end{aligned}$$

The first term alone gives the desired result:

$$\sum_{k=1}^{\lfloor 2^n t \rfloor} \left(M_{k2^{-n}} - M_{(k-1)2^{-n}}\right)\left(\tilde{M}_{k2^{-n}} - \tilde{M}_{(k-1)2^{-n}}\right) \xrightarrow[n \to \infty]{\mathbb{P}} \langle M, \tilde{M} \rangle_t = \langle X, \tilde{X} \rangle_t$$

by Proposition 3.16.

The last three terms contribute nothing, all essentially for the same reason. Let us handle explicitly the second term.

Let $\tilde{V}$ denote the total variation process of the finite variation process $\tilde{A}$. Then we have

$$\left| \sum_{k=1}^{\lfloor 2^n t \rfloor} (M_{k2^{-n}} - M_{(k-1)2^{-n}})(\tilde{A}_{k2^{-n}} - \tilde{A}_{(k-1)2^{-n}}) \right|$$

$$\leq \underbrace{\sup_{1 \leq k \leq \lfloor 2^n t \rfloor} \left| M_{k2^{-n}} - M_{(k-1)2^{-n}} \right|}_{\xrightarrow[n\to\infty]{a.s.} 0} \cdot \underbrace{\sum_{k=1}^{\lfloor 2^n t \rfloor} \left| \tilde{A}_{k2^{-n}} - \tilde{A}_{(k-1)2^{-n}} \right|}_{\leq \tilde{V}_t < \infty} .$$

by (uniform) continuity of paths $s \mapsto M_s(\omega)$ on $s \in [0,t]$ compact interval

This shows that the second term tends to 0 as $n \to \infty$ almost surely, and a similar argument works for the last two terms.

$\square$

Let us first consider the one-variable case of Itô's formula, and start by proving that the formula holds for all polynomial functions.

<u>Lemma</u>: Let $X = (X_t)_{t \in \mathbb{R}_+}$ be a continuous semimartingale. Then for all polynomial functions $p : \mathbb{R} \to \mathbb{R}$ $\left(p(x) = c_n x^n + c_{n-1} x^{n-1} + \dots + c_1 x + c_0 \right)$ we have

$$p(X_t) - p(X_0) = \int_0^t p'(X_s)\, dX_s + \frac{1}{2} \int_0^t p''(X_s)\, d\langle X, X \rangle_s . \quad (\star)$$

Proof: Let $\mathcal{H}$ be the collection of those polynomial functions p for which ⊛ holds.

Clearly $\mathcal{H}$ is a vector space (both sides of ⊛ are linear in p). In order to show that $\mathcal{H}$ contains all polynomials, it is therefore sufficient to show that it contains all monomial functions $p(x) = x^n$ $(n \in \mathbb{N})$.

The case $n=0$ is clear: for $p(x) = 1$ both sides of ⊛ vanish.

The case $n=1$ is also clear: for $p(x) = x$ the formula ⊛ states only that

$$X_t - X_0 = \int_0^t 1 \cdot dX_s$$

which is true by definition of integrals (the finite variation part and local mgale part are defined separately, but the above formula is obvious in both cases).

In view of the above observations, it suffices to show that if $p, q \in \mathcal{H}$, then also $p \cdot q \in \mathcal{H}$. (Then we may build all monomials by taking products of the $n=1$ case.)

Suppose therefore that $p, q \in \mathcal{H}$ and denote $P_t = p(X_t)$ and $Q_t = q(X_t)$. Integration by parts (Theorem 3.21) yields

$$P_t Q_t - P_0 Q_0 = \int_0^t P_s \, dQ_s + \int_0^t Q_s \, dP_s + \langle P, Q \rangle_t .$$

Let us consider the first term on the RHS. Since $q \in \mathcal{H}$ by assumption, we have by ⊛:

$$Q_t - Q_0 = \int_0^t q'(X_s)\, dX_s + \frac{1}{2}\int_0^t q''(X_s)\, d\langle X,X\rangle_s .$$

(in shorthand notation: $dQ_t = q'(X_t)\, dX_t + \frac{1}{2} q''(X_t) \cdot d\langle X,X\rangle_t$)

By associativity of stochastic integration, $P \bullet Q = P \bullet (1 \bullet Q)$ this implies

$$\int_0^t P_s\, dQ_s = \int_0^t P_s \cdot q'(X_s)\, dX_s + \frac{1}{2}\int_0^t P_s \cdot q''(X_s)\, d\langle X,X\rangle_s$$

$$= \int_0^t p(X_s)\, q'(X_s)\, dX_s + \frac{1}{2}\int_0^t p(X_s) \cdot q''(X_s)\, d\langle X,X\rangle_s .$$

The second term on the RHS is handled completely analogously, and we get

$$\int_0^t Q_s\, dP_s = \int_0^t p'(X_s)\, q(X_s)\, dX_s + \frac{1}{2}\int_0^t p''(X_s)\, q(X_s)\, d\langle X,X\rangle_s .$$

For the third term on the RHS we first observe that

$$\langle P, Q\rangle = \langle p'(X) \bullet X , q'(X) \bullet X \rangle$$

since the finite variation parts $\frac{1}{2} p''(X) \bullet \langle X,X\rangle$ and $\frac{1}{2} q''(X) \bullet \langle X,X\rangle$ do not affect the quadratic covariation. Then apply the Kunita-Watanabe identity (Proposition 3.17) twice:

$$\langle P, Q\rangle = \langle p'(X) \bullet X , q'(X) \bullet X\rangle$$

$$= p'(X) \bullet \langle X, q'(X) \bullet X\rangle$$

$$= p'(X) \bullet \big(q'(X) \bullet \langle X,X\rangle \big)$$

and finally associativity of finite variation integrals to get

$$\langle P, Q\rangle = p'(X) q'(X) \bullet \langle X, X\rangle \qquad \text{i.e.}$$

$$\langle P, Q\rangle_t = \int_0^t p'(X_s) \cdot q'(X_s)\, d\langle X,X\rangle_s .$$

Combining the three calculations, we have shown that

$$P_t Q_t - P_0 Q_0$$
$$= \int_0^t \underbrace{\left(p(X_s)q'(X_s) + p'(X_s)q(X_s)\right)}_{(p\cdot q)'(X_s)} dX_s$$
$$+ \int_0^t \underbrace{\left(\tfrac{1}{2}p(X_s)q''(X_s) + \tfrac{1}{2}p''(X_s)q(X_s) + p'(X_s)q'(X_s)\right)}_{=\frac{1}{2}(p\cdot q)''(X_s)} d\langle X,X\rangle_s .$$

This shows that $p\cdot q \in \mathcal{H}$. $\square$

<u>Proof of Itô's formula (Theorem 3.22) in the one-variable case:</u>

We must prove that for $X = (X_t)_{t\in\mathbb{R}_+}$ a continuous semimartingale and $f \in C^2(\mathbb{R};\mathbb{R})$

$$f(X_t) - f(X_0) = \int_0^t f'(X_s)\,dX_s + \frac{1}{2}\int_0^t f''(X_s)\,d\langle X,X\rangle_s .$$

To this end, we approximate f by polynomials, for which the formula is known to hold by the above lemma, and we check that each term for the approximations tends to the corresponding term.

Specifically, by Weierstrass approximation theorem, we may find a sequence $p_1, p_2, \ldots$ of polynomial functions such that for every $r>0$ we have

$$\Delta_n^{(r)} := \max\Big\{ \sup_{|x|\le r} |f(x) - p_n(x)| ,\ \sup_{|x|\le r} |f'(x) - p_n'(x)| ,\ \sup_{|x|\le r} |f''(x) - p_n''(x)| \Big\} \xrightarrow[n\to\infty]{} 0 .$$

Let $X = M + A$ be the Doob-Meyer decomposition of the semimartingale X and let V be the total variation process of the finite variation process A.

Choose stopping times

$$\tau_r = \inf\{t \geq 0 \mid |X_t| + V_t + \langle M,M\rangle_t \geq r\}.$$

Then $\tau_r \uparrow +\infty$ as $r \to \infty$, so it is sufficient to prove the assertion on the time intervals $[0,\tau_r]$, i.e., for the stopped process X^{τ_r}.

By the previous lemma we have

$$(*): \quad p_n(X_t^{\tau_r}) - p_n(X_0^{\tau_r}) = \int_0^t p_n'(X_s^{\tau_r})\, dX_s^{\tau_r} + \frac{1}{2}\int_0^t p_n''(X_s^{\tau_r})\, d\langle X,X\rangle_s^{\tau_r}.$$

On the event $\{|X_0| \geq r\}$ we have $\tau_r = 0$ so the LHS vanishes, whereas on the complementary event we have $|X_t^{\tau_r}| \leq r$, so we have

$$\left|\left(p_n(X_t^{\tau_r}) - p_n(X_t^{\tau_n})\right) - \left(f(X_t^{\tau_r}) - f(X_0^{\tau_r})\right)\right| \leq 2\Delta_n^{(r)} \xrightarrow[n\to\infty]{} 0$$

so the LHS has the desired limit as $n \to \infty$.

We next show the same for the two terms on the RHS.

The first term on the RHS in fact consists of two parts

$$\int_0^t p_n'(X_s^{\tau_r})\, dM_s^{\tau_r} \quad + \int_0^t p_n'(X_s^{\tau_r})\, dA_s^{\tau_r}.$$

To handle the first of these, note that $|M^{\tau_r} \mathbb{1}_{\{\tau_r > 0\}}| \leq r$ so (apart from an event where no integration is done) we have $M^{\tau_r} \in \mathcal{M}_c^2$.

Also the integrands are bounded (apart from similar event), so we calculate

$$\| (p_n'(X)\cdot M)^{\tau_r} - (f'(X)\cdot M)^{\tau_r} \|^2_{\mathcal{M}_c^2}$$

$$= \mathbb{E}\Big[\int_0^{\tau_r} \underbrace{\big(p_n'(X_s) - f(X_s)\big)^2}_{\leq (\Delta_n^{(r)})^2} \, d\langle M,M\rangle_s^{\tau_r} \Big] \qquad \text{(Itô isometry)}$$

$$\leq (\Delta_n^{(r)})^2 \cdot \mathbb{E}\big[\langle M,M\rangle_{\tau_r}\big] \leq (\Delta_n^{(r)})^2 \cdot r \xrightarrow[n\to\infty]{} 0.$$

This shows that

$$\int_0^t p_n'(X_s^{\tau_r})\, dM_s^{\tau_r} \xrightarrow[n\to\infty]{} \int_0^t f'(X_s^{\tau_r})\, dM_s^{\tau_r}.$$

To handle the second term, which is of finite variation, we estimate

$$\Big| \int_0^t p_n'(X_s^{\tau_r})\, dA_s^{\tau_r} - \int_0^t f'(X_s^{\tau_r})\, dA_s^{\tau_r} \Big|$$

$$\leq \int_0^t \underbrace{|p_n'(X_s^{\tau_r}) - f'(X_s^{\tau_r})|}_{\leq \Delta_n^{(r)}} \, dV_s^{\tau_r} \leq \Delta_n^{(r)} \cdot V_{\tau_r} \leq \Delta_n^{(r)} \cdot r \xrightarrow[n\to\infty]{} 0.$$

This shows

$$\int_0^t p_n'(X_s^{\tau_r})\cdot dA_s^{\tau_r} \xrightarrow[n\to\infty]{} \int_0^t f'(X_s^{\tau_r})\, dA_s^{\tau_r}.$$

A completely similar argument leads to

$$\frac{1}{2}\int_0^t p_n''(X_s^{\tau_r})\, d\langle X,X\rangle_s^{\tau_r} \xrightarrow[n\to\infty]{} \frac{1}{2}\int_0^t f''(X_s^{\tau_r})\, d\langle X,X\rangle_s^{\tau_r}.$$

Combining the above, we may take the limit $n\to\infty$ in $\circledast$ to obtain the desired Itô's formula up to stopping time τ_r. Since $\tau_r \uparrow \infty$ as $r\to\infty$, this is sufficient.

$\square$

ILLUSTRATION OF THE PRACTICAL RECEIPE FOR APPLYING ITÔ'S FORMULA TO SOLVE PROBABILISTIC PROBLEMS

Let us try to clarify how Itô's formula is almos always used in practice. To exemplify the general receipe, we look at the following concrete questions.

Let $B = (B_t)_{t \in \mathbb{R}_+}$ be a standard Brownian motion (or perhaps a Brownian motion started from a point $x \in \mathbb{R}$).

Let $[a,b] \subset \mathbb{R}$ be an interval containing the starting point, $x \in [a,b]$, and let

$$\tau = \inf \{ t \geq 0 \mid B_t \notin (a,b) \}$$

be the exit time from the interval (τ is a stopping time – hitting a closed set $(a,b)^c$).

Fix also a subinterval $[\alpha, \beta] \subset [a,b]$ and consider the time spent by B on this subinterval before exit from $[a,b]$:

$$T = \int_0^\tau \chi(B_s)\, ds$$

$$\text{where} \qquad \chi(x) = \begin{cases} 1 & \text{if } x \in [\alpha,\beta] \\ 0 & \text{if } x \notin [\alpha,\beta] \end{cases}.$$

This T is a random variable. We may ask, for example:

$\boxed{Q1}$: What is the expected time $\mathbb{E}_x[T]$ spent on $[\alpha,\beta]$ before exiting $[a,b]$?

$\boxed{Q2}$: What is the law of T, for example Laplace transform $\varphi(\theta) = \mathbb{E}_x[e^{-\theta T}]$?

We illustrate a "standard receipe" to solve these questions. The plan is to use Itô's formula combined with optional stopping, but how do we decide which semimartingales $X^{(1)}, \dots, X^{(d)}$ to use and which function $f : \mathbb{R}^d \to \mathbb{R}$ to use to get a martingale

$$M_t = f(X_t^{(1)}, \dots, X_t^{(d)})$$

to plug into optional stopping formula ?

$\boxed{Q1}$: Which semimartingales and which f ?

Intuition : We want to solve

$$g(x) = \mathbb{E}_x[T].$$

(In particular $g(a) = 0$, $g(b) = 0$)

There is a trivial way to construct a martingale related to this expected value :

$$M_t = \mathbb{E}_x[T \mid \mathcal{F}_t]$$

(tower property of conditional expected value implies that such a process has martingale property).

Let us write this M_t explicitly :

$$M_t = \mathbb{E}_x\left[\int_0^\tau \chi(B_s)\,ds \;\middle|\; \mathcal{F}_t\right]$$

$$= \mathbb{E}_x\left[\int_0^{t\wedge\tau} \chi(B_s)\,ds + \int_{t\wedge\tau}^{\tau} \chi(B_s)\,ds \;\middle|\; \mathcal{F}_t\right]$$

$$= \underbrace{\int_0^{t\wedge\tau} \chi(B_s)\,ds}_{\text{this part is } \mathcal{F}_t\text{-measurable}} + \underbrace{\mathbb{E}_{B_{t\wedge\tau}}\left[\int_0^{\hat\tau} \chi(\hat B_s)\,ds\right]}_{\text{Markov property says that future } \hat B \text{ is a Brownian started from } B_{t\wedge\tau}.}$$

If we denote (generalizing def. of T)

$$T_t = \int_0^{t\wedge\tau} \chi(B_s)\, ds \, ,$$

then we have arrived at

$$M_t = T_{t\wedge\tau} + g(B_{t\wedge\tau}) .$$

This suggests the following choice ...

We use semimartingales

$$X_t^{(1)} = B_{t\wedge\tau}$$

$$X_t^{(2)} = T_t = \int_0^{t\wedge\tau} \chi(B_s)\, ds$$

(increasing continuous adapted process $\Rightarrow$ semimgale)

and function

$$f : \mathbb{R}^2 \to \mathbb{R}$$

of the form

$$f(x,u) = u + g(x)$$

where $g : \mathbb{R} \to \mathbb{R}$ is a (yet unknown) function and we consider the process

$$M_t := f(B_{t\wedge\tau}, T_{t\wedge\tau}) = T_{t\wedge\tau} + g(B_{t\wedge\tau}) ,$$

which should be a (local) martingale.

Itô's formula says

$$M_t - M_0 = f(B_{t\wedge\tau}, T_{t\wedge\tau}) - f(B_0, T_0)$$

$$= \int_0^{t\wedge\tau} (\partial_x f)(B_s, T_s)\, dB_s$$

$$+ \frac{1}{2} \int_0^{t\wedge\tau} (\partial_x^2 f)(B_s, T_s)\, d\langle B, B\rangle_s$$

$$+ \int_0^{t\wedge\tau} (\partial_u f)(B_s, T_s)\, dT_s \, .$$

We observe

$$d\langle B,B\rangle_s = ds$$
$$dT_s = \chi(B_s)\cdot ds$$

($T_t = \int_0^{t\wedge\tau} \chi(B_s)\,ds$ and associativity of finite variation integrals)

and

$$(\partial_x f)(x,u) = g'(x)$$
$$(\partial_x^2 f)(x,u) = g''(x)$$
$$(\partial_u f)(x,u) = 1$$

so the formula simplifies to

$$M_t - M_0 = \int_0^{t\wedge\tau} g'(B_s)\,dB_s + \int_0^{t\wedge\tau}\left(\tfrac{1}{2}g''(B_s) + \chi(B_s)\right)ds .$$

The first term here is a local martingale (which we want) but the second is a finite variation process (which we do not want). We get rid of the second term if g solves

$$\tfrac{1}{2}g''(x) + \chi(x) = 0 \qquad \forall x\in[a,b].$$

So let us solve this differential equation, which becomes (piecewise)

$$\begin{cases} g''(x) = 0 & \text{if } \quad a \le x < \alpha \\ g''(x) = -2 & \text{if } \quad \alpha \le x \le \beta \\ g''(x) = 0 & \text{if } \quad \beta < x \le b . \end{cases}$$

Piecewise these second order differential equations are easy to solve.

$$\left(\begin{array}{lll} g''(x) = 0 & \Longrightarrow & g(x) = \sigma x + \mu \\ g''(x) = -2 & \Longrightarrow & g(x) = -x^2 + c_1 x + c_0 \end{array}\right)$$

Taking into account desired boundary conditions $g(a)=0$, $g(b)=0$ we have

$$g(x) = \begin{cases} \sigma_a \cdot (x-a) & \text{if} \quad x \in [a,\alpha) \\ -x^2 + c_1 x + c_0 & \text{if} \quad x \in [\alpha,\beta] \\ \sigma_b \cdot (x-b) & \text{if} \quad x \in (\beta,b] \, . \end{cases}$$

There are four remaining unknowns $\sigma_a, \sigma_b, c_0, c_1 \in \mathbb{R}$ which can be solved from the requirements of

- continuity at $x=\alpha$ and $x=\beta$:

$$\sigma_a \cdot (\alpha - a) = -\alpha^2 + c_1\alpha + c_0$$
$$\sigma_b \cdot (\beta - b) = -\beta^2 + c_1\beta + c_0$$

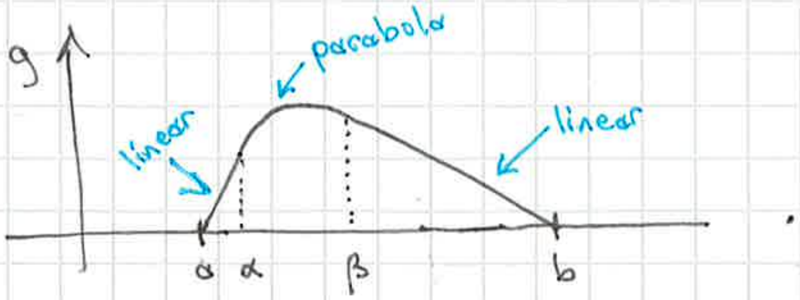
- continuity of derivative at $x=\alpha$ and $x=\beta$:

$$\sigma_a = -2\alpha + c_1$$
$$\sigma_b = -2\beta + c_1$$

The solution is

$$\sigma_a = \frac{\beta-\alpha}{b-a} \cdot (2b - \alpha - \beta)$$
$$\sigma_b = \frac{\beta-\alpha}{b-a} \cdot (2a - \alpha - \beta)$$
$$c_1 = \frac{1}{b-a}\left(\alpha^2 - \beta^2 + 2b\beta - 2a\alpha\right)$$
$$c_0 = \frac{1}{b-a}\left(a\beta^2 - b\alpha^2 + 2ab\alpha - 2ab\beta\right)$$

and the function g looks like:



Now we know exactly the form of $f:\mathbb{R}^2\to\mathbb{R}$ we want to use. We thus define the process

$$M_t = T_{t\wedge\tau} + g(B_{t\wedge\tau})$$

where g is the above explicit function. The calculation we made with Itô's formula guarantees that M is a local martingale. ⊛ Also g is bounded and $|T_{t\wedge\tau}| \le t$, so this M is in fact a true martingale.

Strictly speaking there is a discontinuity in g'', so we should have used a slightly smoothened χ_ε to really get $f\in C^2(\mathbb{R}^2;\mathbb{R})$ but let us not get stuck.

It remains to apply optional stopping to M:

$$0+g(x) = M_0 = \mathbb{E}_x[M_{t\wedge\tau}]$$

$$= \mathbb{E}_x[T_{t\wedge\tau}] + \mathbb{E}_x[g(B_{t\wedge\tau})].$$

Monotone convergence theorem gives

$$\mathbb{E}_x[T_{t\wedge\tau}] \xrightarrow[t\to\infty]{} \mathbb{E}_x[T_\tau] = \mathbb{E}_x[T].$$

Also $\tau<\infty$ almost surely and thus $B_\tau\in\{a,b\}$ a.s., so since $g(a)=0=g(b)$ we have

$$g(B_{t\wedge\tau}) \xrightarrow[t\to\infty]{a.s.} 0.$$

Bounded convergence theorem gives

$$\mathbb{E}_x[g(B_{t\wedge\tau})] \xrightarrow[t\to\infty]{} 0.$$

We conclude

$$\boxed{g(x) = \mathbb{E}_x[T].}$$

This solves $\boxed{Q1}$.

$\boxed{Q2}$: Which semimartingales and which f ?

<u>Intuition:</u> We want to solve, for $\theta \geq 0$,

$$h(x) = \mathbb{E}_x\left[e^{-\theta T}\right]. \leftarrow \begin{pmatrix} \text{In particular} \\ h(a)=1, \ h(b)=1 \end{pmatrix}$$

A martingale can be constructed by

$$N_t = \mathbb{E}_x\left[e^{-\theta T} \,\middle|\, \mathcal{F}_t\right].$$

Write this explicitly:

$$N_t = \mathbb{E}_x\left[\exp\left(-\theta \cdot \int_0^{t\wedge\tau} \chi(B_s)\,ds - \theta \cdot \int_{t\wedge\tau}^{\tau} \chi(B_s)\,ds\right) \,\middle|\, \mathcal{F}_t\right]$$

$$= \exp\left(-\theta \cdot \int_0^{t\wedge\tau} \chi(B_s)\,ds\right) \cdot \mathbb{E}_{B_{t\wedge\tau}}\left[\exp\left(-\theta \int_0^{\hat{\tau}} \chi(\hat{B}_s)\,ds\right)\right]$$

$\mathcal{F}_t$-measurable

$$= e^{-\theta T_{t\wedge\tau}} \cdot h(B_{t\wedge\tau}).$$

Markov property

This suggests the following choice...

Use semimartingales

$$X_t^{(1)} = B_t \qquad X_t^{(2)} = T_t = \int_0^{t\wedge\tau} \chi(B_s)\,ds$$

again, and function $f: \mathbb{R}^2 \to \mathbb{R}$ of the form

$$f(x,u) = e^{-\theta u} \cdot h(x)$$

where $h: \mathbb{R} \to \mathbb{R}$ is a (yet unknown) function.

Consider process

$$N_t = f(B_{t\wedge\tau}, T_{t\wedge\tau}) = e^{-\theta T_{t\wedge\tau}} \cdot h(B_{t\wedge\tau}).$$

In order to prepare for Itô's formula, calculate

$$\partial_x f(x,u) = e^{-\theta u} \cdot h'(x)$$

$$\partial_x^2 f(x,u) = e^{-\theta u} \cdot h''(x)$$

$$\partial_u f(x,u) = -\theta \cdot e^{-\theta u} \cdot h(x) \quad .$$

Now Itô's formula gives

$$
\begin{aligned}
N_t - N_0 &= f(B_{t\wedge\tau}, T_{t\wedge\tau}) - f(B_0, T_0) \\
&= \int_0^{t\wedge\tau} (\partial_x f)(B_s, T_s)\, dB_s \\
&\quad + \frac{1}{2}\int_0^{t\wedge\tau} (\partial_x^2 f)(B_s, T_s)\, d\langle B, B\rangle_s + \int_0^{t\wedge\tau} (\partial_u f)(B_s, T_s)\, dT_s \\
&= \int_0^{t\wedge\tau} e^{-\theta T_s} h'(B_s)\, dB_s \\
&\quad + \int_0^{t\wedge\tau} e^{-\theta T_s}\left(\frac{1}{2} h''(B_s) - \theta \cdot h(B_s)\cdot \chi(B_s)\right) ds .
\end{aligned}
$$

The first term here is a local martingale (as we want), but the second term is a finite variation process (which we do not want). We get rid of the second term if h solves

$$\frac{1}{2} h''(x) - \theta\cdot\chi(x)\cdot h(x) = 0 \qquad \forall x\in[a,b].$$

After looking for a solution piecewise on $[a,\alpha)$ and $[\alpha,\beta]$ and $(\beta,b]$, and taking into account the desired boundary values $h(a)=1$, $h(b)=1$, we arrive at

$$
h(x) = \begin{cases}
1 + \rho_a\cdot(x-a) & \text{if } x\in[a,\alpha) \\
A_+\cdot e^{x\sqrt{2\theta}} + A_-\cdot e^{-x\sqrt{2\theta}} & \text{if } x\in[\alpha,\beta] \\
1 + \rho_b\cdot(x-b) & \text{if } x\in(\beta,b]
\end{cases}
$$

where $\rho_a, \rho_b, A_+, A_- \in \mathbb{R}$ are parameters.

These parameters can be solved from

- continuity of h at $x=\alpha$ and $x=\beta$
- continuity of h' at $x=\alpha$ and $x=\beta$.

Exercise : solve these equations and observe that $h:[a,b] \to [0,1]$ (h is bounded, $0 \le h(x) \le 1 \quad \forall x$).

Then use this h to construct the process

$$N_t = e^{-\theta T_{t\wedge\tau}} \cdot h(B_{t\wedge\tau})$$

which is a local martingale (by Itô's formula) and bounded, so in fact a true martingale.

Use optional stopping theorem :

$$e^{-0} \cdot h(x) = N_0 = \mathbb{E}_x[N_{t\wedge\tau}]$$

$$= \mathbb{E}_x\Big[\underbrace{e^{-\theta T_{t\wedge\tau}}}_{\xrightarrow[t\to\infty]{} e^{-\theta T}} \cdot \underbrace{h(B_{t\wedge\tau})}_{\xrightarrow[t\to\infty]{} 1}\Big] \quad \text{since } h(a)=1=h(b)$$

$$\xrightarrow[t\to\infty]{BCT} \mathbb{E}_x[e^{-\theta T}].$$

We conclude

$$\boxed{h(x) = \mathbb{E}_x[e^{-\theta T}]}$$

This solves $\boxed{Q2}$

(the function h is explicit, but slightly complicated).