Lecture Notes
STOCHASTIC DIFFERENTIAL EQUATIONS

Andreas Löpker, TU/e Eindhoven

PREFACE

These are the lecture notes of the Mastermath course 'Stochastic Differential Equations' in the semester 2009/2010, held at Utrecht university.

The lecture notes follow in large parts (namely the chapters 2 to 7) the exciting textbook "Stochastic calculus and financial applications" by Michael Steele (see [🕮7]). The numbering of the sections was adopted from this book, so theorem and section numbering is not always continuous. Additionally, every subsection has its own number, indicated and referred to by double brackets $[\![\cdot]\!]$.

The notes start with a very brief introduction to probability and measure theory, before introducing martingales, Brownian motion and finally stochastic integrals.

In the Appendix A we give some additional information on topics that couldn't be discussed during the lecture. Appendices B and C consist of a collection of exercises with solutions that were handed out during the classes. The lecture notes close with a short list of literature for further study.

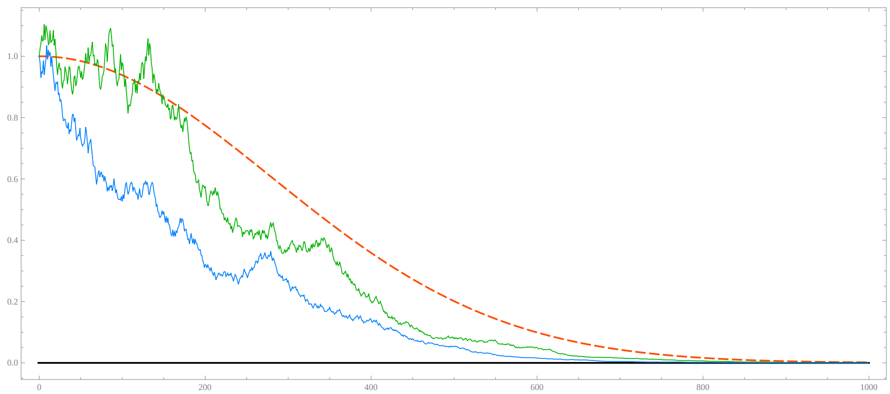
CONTENTS

MOTIVATION

Consider an ordinary differential equation

$$\frac{d}{dt}X_t = \mu(t, X_t), \quad t \geq 0, \tag{1}$$

where we write the unknown function as X_t instead of $X(t)$, since we will consider random processes later. For instance take $\mu(t,x) = -t \cdot x$, then $\frac{d}{dt}X_t = -t \cdot X_t$, with solution $X_t = X_0 \exp(-t^2/2)$ (see the orange curve in the figure below).



If we want to add a random disturbance, depending both on t and X_t, we take a random noise W_t, multiply it with some sufficiently nice function σ and add it to (1). We end up with an stochastic differential equation (SDE):

$$\frac{d}{dt}X_t = \mu(t, X_t) + \sigma(t, X_t)W_t. \tag{2}$$

Equation (2) can also be written as $dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)W_t\,dt$ or

$$X_t = X_0 + \int_0^t \mu(s, X_s)\,ds + \int_0^t \sigma(s, X_s)W_s\,ds.$$

If W_s is a nicely behaving process, then we may study this equation and find for example conditions for the existence and uniqueness of solutions.

Particularly interesting would be to use a *white noise* W_t, a stochastic process such that "$\frac{d}{dt}B_t = W_t$", where B_t is the famous Brownian motion. Then we could write

$$X_t = X_0 + \int_0^t \mu(s, X_s)\,ds + \int_0^t \sigma(s, X_s)\,dB_s \tag{3}$$

Clearly, if μ is continuous, the first integral is well defined. Unfortunately the dubious white noise process W_t does not exist as a stochastic process (although it has a Wikipedia article...). Brownian motion is not differentiable and it is not clear, if the right integral in (3) (the *stochastic integral*) exists or has any meaning at all. A considerable part of this course will be devoted to answer the question, whether we can actually give (3) a reasonable meaning.

1 MEASURE THEORY AND PROBABILITY

1.1 Measure spaces

⟦1⟧ **Measure space ingredients –** If we want to measure we clearly need

(1) objects to be measured: sets.

(2) a measure that assigns to each set a real number.

All sets to be measured are subsets of some basic set ω. The family of all sets that we can measure is denoted by $\mathscr{F}$

Usually $\mathscr{F}$ is smaller then the collection $\mathscr{P}(\omega)$ of all subsets of ω. This has some technical reasons that will be mentioned later.

⟦2⟧ **σ-field –** To be consistent with our intuition, $\mathscr{F}$ should have a certain structure.

A *σ-field* (or σ-field) in ω is a family of subsets of ω, such that

(1) $\omega \in \mathscr{F}$

(2) $A \in \mathscr{F} \Rightarrow A^c \in \mathscr{F}$ (where A^c is the complement of A)

(3) $A_1, A_2, \ldots \in \mathscr{F} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathscr{F}$

The sets in $\mathscr{F}$ are called *measurable sets*, $(\omega, \mathscr{F})$ is a *measurable space*.

⟦3⟧ **Generated σ-field –** We can produce σ-fields from collections $\mathscr{C}$ of subsets of ω. Just take the intersection of all σ-fields that contain $\mathscr{C}$ (prove that this is indeed a σ-field). This intersection is called the σ-field *generated* by $\mathscr{C}$ (written $\sigma(\mathscr{C})$).

⟦4⟧ **Example - Borel σ-field –** An important example is the following. Let $\mathscr{C}$ be the collection of all open sets in $[0,1]$ (could be another topological space), then $\mathscr{B}([0,1]) = \sigma(\mathscr{C})$ is called the Borel σ-field, the natural candidate to choose as an σ-field on $[0,1]$. It turns out that if you take the family $\mathscr{I}$ of all intervals of $[0,1]$, then $\sigma(\mathscr{I}) = \mathscr{B}([0,1])$.

[[5]] **Measures –** Once we have defined our measurable objects, how can we actually measure them?

A *measure* is a function $\mu : \mathscr{F} \to [0,\infty]$, such that

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

(2) For disjoint sets $A_1, A_2, \ldots \in \mathscr{F}$ we have

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu(A_i).$$

The triple $(\omega, \mathscr{F}, \mu)$ is called a *measure space*.

[[6]] **Null sets –** Sets A with $\mu(A) = 0$ are called *null sets*. If a property holds for all $\omega \in \omega$ except for ω in a null set, then we say the property holds almost everywhere (a.e.).

[[7]] **Completion –** Sometimes we stumble upon the situation that not all subsets of ω that are subsets of a null set are in $\mathscr{F}$. It is then possible to extend $\mathscr{F}$ to a σ-field $\overline{\mathscr{F}}$, such that any subset of a null set is again measurable (and has of course measure zero). This new σ-field is called the *completion* of $\mathscr{F}$.

[[8]] **Lebesgue measure –** Particularly interesting for practical purposes would be a measure on $\mathbb{R}$, such that the measure of an interval I is its length. There is indeed such a measure λ, defined for all sets $A \in \mathscr{B}(\mathbb{R})$. If we complete $\mathscr{B}(\mathbb{R})$ to $\mathscr{L}(\mathbb{R}) = \overline{\mathscr{B}(\mathbb{R})}$ we end up with the so called *Lebesgue measurable sets* and the measure λ can be extended to $\mathscr{L}(\mathbb{R})$.

[[9]] **An disturbing fact –** There are actually subsets of $\mathbb{R}$ that are not in $\mathscr{B}(\mathbb{R})$ and some are not even in $\mathscr{L}(\mathbb{R})$.☞ These sets can not be measured by λ. You can do strange things with nonmeasurable sets. If you are interested and brave enough check the notion "Banach-Tarski paradox" with some internet search engine. ☞ (1)

1.2 Measurable functions

[[10]] **Measurable functions –** Like continuous functions are 'nice' mappings between topological spaces, measurable functions are the 'natural' functions to use with measurable spaces.

A function f from a measurable space $(\omega_1, \mathscr{F}_1)$ to another measurable space $(\omega_2, \mathscr{F}_2)$ is *measurable* if

$$f^{-1}(B) = \{\omega \in \omega_1 : X(\omega) \in B\} \in \mathscr{F}_1,$$

for every $B \in \mathscr{F}_2$. We also say that f is measurable with respect to $\mathscr{F}_1$.

⟦11⟧ How can we decide whether a function is measurable? –
How many functions are actually measurable and do our favourite functions from calculus belong to this class? It is sometimes very difficult to verify the above definition. It turns out that the class of measurable functions is in fact very high and it takes considerable amount of work to construct functions that do not belong to it.

A function $f : (\omega, \mathscr{F}) \to \mathbb{R}$ (we omit the Borel σ-field) is measurable iff for all $y \in \mathbb{R}$ either $\{\omega : f(\omega) \geq y\} \in \mathscr{F}$ or $\{\omega : f(\omega) > y\} \in \mathscr{F}$ or $\{\omega : f(\omega) \leq y\} \in \mathscr{F}$ or $\{\omega : f(\omega) < y\} \in \mathscr{F}$.

You can perform the following operations on measurable functions

If $f_1, f_2, \ldots : (\omega, \mathscr{F}) \to \mathbb{R}$ are measurable, then the same holds for

$$|f_1|, f_1 + f_2, f_1 - f_2, f_1 \cdot f_2,$$
$$\sup_{n\in\mathbb{N}} f_n, \inf_{n\in\mathbb{N}} f_n,$$
$$\limsup_{n\to\infty} f_n (!), \liminf_{n\to\infty} f_n (!),$$
$$\max\{f_1, \ldots, f_n\}, \min\{f_1, \ldots, f_n\}.$$

1.3 Lebesgue integral

Once measures and measurable functions are designed, it is obvious to ask, whether integration can be defined in a suitable way. This is the case, and the definition below turns out to generate an integral that beats our old friend the Riemann integral in most regards.

The standard way to construct the so called *Lebesgue integral* is to use a three step construction that defines the integral first for a class of simple functions, then for nonnegative limits of them and finally for general measurable functions.

⟦12⟧ (I) Simple functions – A real-valued measurable function $f : (\omega, \mathscr{F}) \to \mathbb{R}$ is called *simple* if there are sets $A_i \in \mathscr{F}$, $i = 1, 2, \ldots, n$, such that

$$f(x) = \sum_{i=1}^{n} c_i \mathbb{1}_{A_i}(x), \tag{1.1}$$

where $\mathbb{1}_{A_i}(x) = 1$ if $x \in A_i$ and 0 else.

We define its integral (in the obvious way) by

$$\int f(x)\, d\mu(x) = \sum_{i=1}^{n} c_i \mu(A_i).$$

This integral does not depend on the choice of the sets A_i in (1.1)

⟦13⟧ (II) Non-negative measurable functions – Note that any nonnegative measurable functon $f : (\omega, \mathscr{F}) \to \mathbb{R}^+$ is an increasing limit of nonnegative simple functions $f_1, f_2, \ldots$.

We define

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

One can prove that for any other sequence $g_1, g_2, \ldots$ with $g_n \uparrow f$ the defined integral coincedes, i.e.

$$\lim_{n\to\infty}\int g_n(x)\,d\mu(x) = \lim_{n\to\infty}\int f_n(x)\,d\mu(x)$$

⟦14⟧ (III) Any measurable function –

Let f be some measurable function. If $\int |f|\,d\mu < \infty$ then we call f *integrable* and we define

$$\int f(x)\,d\mu(x) = \int f^+(x)\,d\mu(x) - \int f^-(x)\,d\mu(x),$$

where $f^+(x) = \max\{f(x), 0\}$ and $f^-(x) = \max\{-f(x), 0\}$. Note that $\int |f|\,d\mu < \infty$ iff $\int f^+\,d\mu < \infty$ and $\int f^-\,d\mu < \infty$.

⟦15⟧ Notation – We write

$$\int_A f(x)\,d\mu(x) = \int \mathbb{1}_A(x) f(x)\,d\mu(x)$$

if $A \in \mathscr{F}$ and $\int f(x)\,dx = \int f(x)\,d\lambda(x)$, if the integrating measure is the Lebesgue-measure.

⟦16⟧ Theorem (Properties) *The Lebesgue integral enjoys all the properties that we are used to see for the classical Riemann integral from calculus.*

(1) The integral is linear: $\int (af + bg)\,d\mu = a\int f\,d\mu + a\int g\,d\mu$.

(2) If $f \le g$ *then* $\int f\,d\mu \le \int g\,d\mu$.

(3) If $\mu(A) = 0$ *then* $\int_A f\,d\mu = 0$. *If* $\int_A f\,d\mu = 0$ *and* $f \ge 0$ *then* $\mu(A) = 0$.

If $f : I \to \mathbb{R}$, where I a closed interval, is Riemann integrable then it also Lebesgue integrable (w.r.t Lebesgue measure λ) and the value of the integrals coincide.

On the whole real line we have to be careful, e.g. the improper Riemann integral

$$\int_0^\infty \frac{\sin(x)}{x}\,dx$$

(defined as the limit of the proper integrals $\int_0^n \sin(x)/x\,dx$) exists and is finite. But since $\int_0^\infty |\sin(x)/x|\,dx = \infty$ the Lebesgue integral does not exist.

However, in most cases the Lebesgue integral wins the contest. Especially function which are not very smooth can often be Lebesgue integrated but not Riemann integrated. A famous example is $\int_0^1 \mathbb{1}_{\mathbb{Q}}(x)\,dx$, where the Riemann integral does not converge to a finite limit, but the Lebesgue integral exists and is equal to zero. ☞ ☞ (2)

⟦17⟧ **Three convergence theorems –**

The main advantage of Lebesgue integration is the fact that the integral behaves nicely with limits, as the next three cornerstones of integration theory show:

⟦18⟧ **Theorem (Lemma of Fatou)** *If $f_1, f_2, \ldots \geq 0$ are measurable then*

$$\liminf_{n\to\infty} \int f_n(x)\,d\mu(x) \geq \int \liminf_{n\to\infty} f_n(x)\,d\mu(x)$$

To remember Fatou's lemma, think of the functions $f_n(x) = 1 + \sin(2\pi x + n)$, for which the limit inferior is $f(x) = 0$, but the integrals $\int_0^1 f_n(x)\,dx$ are equal to one (note that f_n does not converge).

⟦19⟧ **Theorem (Monotone convergence (Levi))** *If $f, f_1, f_2 \ldots$ are measurable with $0 \leq f_n \uparrow f$, then*

$$\lim_{n\to\infty} \int f_n(x)\,d\mu(x) = \int f(x)\,d\mu(x).$$

⟦20⟧ **Theorem (Dominated convergence (Lebesgue))** *If $f, g, f_1, f_2, \ldots$ are measurable functions with $|f_n| \leq g$ and $f_n \to f$ and g is integrable then*

$$\lim_{n\to\infty} \int f_n(x)\,d\mu(x) = \int f(x)\,d\mu(x).$$

For some transformation rules see ☞. ☞ (3)

1.4 $\mathscr{L}^p$ spaces

⟦21⟧ **Definition ($\mathscr{L}^p$ space)** *A measurable function $f : (\omega, \mathscr{F}) \to \mathbb{R}$ is in $\mathscr{L}^p_\mu$, $p \geq 1$, if*

$$\int |f(x)|^p\,d\mu(x) < \infty.$$

$\mathscr{L}^p_\mu$ is a complete normed vector space with norm

$$||f||_p = \left(\int |f(x)|^p\,d\mu(x)\right)^{1/p}$$

We let $\mathscr{L}^\infty_\mu$ be the set of all $f : (\omega, \mathscr{F}) \to \mathbb{R}$, such that

$$||f||_\infty = \inf\{y : \mu(|f| > y) = 0\} < \infty.$$

If μ is a finite measure, i.e. $\mu(\omega) < \infty$ then

$$\mathscr{L}^{\infty} \subset \mathscr{L}^{p} \subset \mathscr{L}^{q} \subset \mathscr{L}^{1}$$

for $p > q$ (why isn't this true e.g. for the Lebesgue measure, which is certainly not finite?). For $f, f_1, f_2 \ldots \in \mathscr{L}^p$ we say that f_n converges to f in $\mathscr{L}^p$ if $||f_n - f||_p \to 0$.

⟦22⟧ **Special case** $p = 2$ **–** If $p = 2$ then $\mathscr{L}^p_\mu$ is a Hilbert space, i.e. a complete normed vector space with inner product. The inner product is simply defined by

$$\langle f, g \rangle = \int f(x) g(x) \, d\mu(x).$$

⟦23⟧ **Important Inequalities –**

The following famous inequalities are heavily used in connection with $\mathscr{L}^p$ spaces.

(1) Minkowski inequality:

$$||f + g||_p \leq ||f||_p + ||g||_p$$

(2) Hölder inequality: $1/p + 1/q = 1$, then

$$||fg||_1 \leq ||f||_p ||g||_q,$$

(3) Special case (Cauchy-[Bunyakovsky]-Schwarz)

$$||fg||_1 \leq ||f||_2 ||g||_2$$

1.5 Product measure and Fubini's theorem

Given two measure spaces $(\omega_1, \mathscr{F}_1, \mathbb{P}_1)$ and $(\omega_1, \mathscr{F}_1, \mathbb{P}_1)$, there is a measure space $(\omega, \mathscr{F}, \mathbb{P})$ such that $\omega = \omega_1 \times \omega_2$ and

$$\mathbb{P}(A \times B) = \mathbb{P}_1(A) \mathbb{P}_1(B), \quad A \in \mathscr{F}_1, B \in \mathscr{F}_2.$$

This is the product space with the product measure $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$ (we don't say anything about $\mathscr{F}$ here, which is constructed from $\mathscr{F}_1$ and $\mathscr{F}_2$ in a certain way).

⟦24⟧ **Theorem (Fubini)** *If either*

- $f : (\omega, \mathscr{F}) \to \mathbb{R}$ *is* $\mathbb{P}$ *integrable or*
- $\int \int |f(x,y)| d\mathbb{P}_1(x) d\mathbb{P}_2(y) < \infty$ *and* $\int \int |f(x,y)| d\mathbb{P}_1(y) d\mathbb{P}_2(x) < \infty$,

then

$$\int f(x,y) d\mathbb{P}(x,y) = \int \int f(x,y) d\mathbb{P}_1(x) d\mathbb{P}_2(y) = \int \int f(x,y) d\mathbb{P}_1(y) d\mathbb{P}_2(x).$$

1.6 Probability spaces

⟦25⟧ **Definition –** As we stated before, probability theory can be seen as a special discipline of measure theory with some important extra features (like independence). Here is the definition of a probability space.

A *probability space* is a measure space $(\omega, \mathscr{F}, \mathbb{P})$, where $\mathbb{P}$ is a finite measure with $\mathbb{P}(\omega) = 1$.

We interpret elements $A \in \mathscr{F}$ as events and $P(A)$ as the probability that A happens.

Measurable functions $X : (\omega, \mathscr{F}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ are called *random variables*.

The definition of random variables makes sense, since measurability of X guarantees that sets of the form $\{\omega \in \omega : X(\omega) \in B, B \in \mathscr{B}(\mathbb{R})\}$ are in $\mathscr{F}$ and hence $\mathbb{P}$-measurable. We shortly write

$$\mathbb{P}(X \text{ has property } A) = \mathbb{P}(\{\omega \in \omega : X(\omega) \text{ has property } A\}).$$

We say that a property holds with probability one (or almost surely) if

$$\mathbb{P}(\text{property holds}) = 1.$$

1.7 Expectation

The *expectation* of X is defined as

$$\mathbb{E}(X) = \int X(\omega) dP(\omega) = \int_{-\infty}^{\infty} x \, dF(x).$$

The properties of the expectation are of course inherited from the integral:

[[26]] Theorem (Properties of the integral)

(1) $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$,

(2) *If* $X \le Y$ *a.s. then* $\mathbb{E}(X) \le \mathbb{E}(Y)$,

(3) $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$,

Our three integration theorems now read:

Fatou:
If $0 \le X_n$ then $\mathbb{E}(\liminf X_n) \le \liminf \mathbb{E}(X_n)$.

Monotone convergence:
If $0 \le X_n \uparrow X$ then $\lim \mathbb{E}(X_n) = \mathbb{E}(X)$.

Dominated convergence:
If $X_n \to X$ almost surely and $|X_n| \le Y$ with $\mathbb{E}(Y) < \infty$ then $\lim \mathbb{E}(X_n) = \mathbb{E}(X) < \infty$.

We can also define the expectation of functions of X,

$$\mathbb{E}(g(X)) = \int g(X(\omega))\,dP(\omega) = \int_{-\infty}^{\infty} g(x)\,dF(x).$$

For example $\mathbb{E}(X^n)$ is called the nth *moment* of X. If $\mathbb{E}(X^2)$ exists, then

$$\mathrm{Var(X)} = \mathbb{E}(\mathrm{X}^2) - \mathbb{E}(\mathrm{X})^2.$$

is the variance of X and $\sigma(X) = \sqrt{\mathrm{Var(X)}}$ is the *standard deviation* of X.
The covariance of two random variables is defined as

$$\mathrm{Cov(X,Y)} = \mathbb{E}(\mathrm{XY}) - \mathbb{E}(\mathrm{X})\mathbb{E}(\mathrm{Y}).$$

The *Laplace transform* of X is the function $s \mapsto \mathbb{E}(e^{-sX})$, $s \ge 0$.
The *characteristic function* of X is the function $s \mapsto \mathbb{E}(e^{isX})$, $s \in \mathbb{R}$.

[[27]] Distribution function – The distribution function of X is given by

$$F(x) = \mathbb{P}(X \le x).$$

F is nondecreasing and right-continuous and it is possible to define the *Stieltjes integral* (see the section on 'bounded variation' in your favourite textbook on advanced calculus)

$$\mathbb{E}(g(X)) = \int g(x)\,dF(x) = \int g(X(\omega))\,dP(\omega)).$$

X is a *discrete random variable* if there is a set $C = \{c_1, c_2, \ldots\}$ of elements in ω, such that $\mathbb{P}(C) = 1$. In that case F is a step function and

$$\mathbb{E}(g(X)) = \int g(x)\,dF(x) = \sum_{i=1}^{\infty} c_i \mathbb{P}(X = c_i).$$

If $F(x) = \int_{-\infty}^{x} f(x)dx$ then F is called absolutely continous and f is the probability density function of X. In this case

$$\mathbb{E}(g(X)) = \int g(x)dF(x) = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

⟦28⟧ **Independence –** Two events A, B are *independent* if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. If $\mathscr{C}, \mathscr{D}$ are two collections of events, then $\mathscr{C}$ and $\mathscr{D}$ are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for all $A \in \mathscr{C}, B \in \mathscr{D}$.
Two random variables X and Y are independent if

$$H(x, y) = F(x)G(y)$$

where $H(x, y) = \mathbb{P}(X \le x, Y \le y)$ is the joint distribution function of X, Y and F and G are the distribution functions of X and Y respectively.

⟦29⟧ **Theorem (Borel-Cantelli lemma)** *Let $(A_i)_{i=1,2,\dots}$ be a sequence of events and let B be the event that A_i happens for infinitely many i. Then*

(1) $\sum_{i=1}^{\infty} \mathbb{P}(A_i) < \infty$ *implies* $\mathbb{P}(B) = 0$.

(2) *if the events are independent then* $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$ *implies* $\mathbb{P}(B) = 1$.

☞ ☞ (4)

⟦30⟧ **Inequalities –** Markov's inequality:

$$\mathbb{P}(X > \varepsilon) \le \frac{\mathbb{E}(X)}{\varepsilon}, \quad \varepsilon > 0.$$

Chebyshev's inequality:

$$\mathbb{P}(|X - \mathbb{E}(X)| > \varepsilon) \le \frac{\mathrm{Var(X)}}{\varepsilon^2}, \quad \varepsilon > 0.$$

1.8 Convergence of random variables

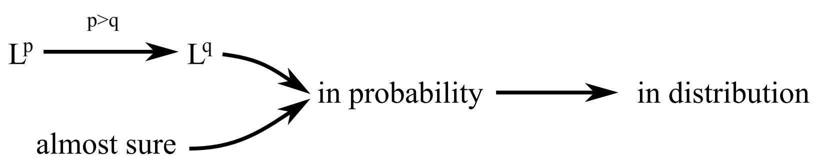
⟦31⟧ **Convergence –**

Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of random variables. We write $X_n \to X$, if we want to say that X_n converges to X as n tends to infinity.

Here is a formal definition of different types of convergence:

Types of convergence:

- Sure (pointwise) convergence: $X_n(\omega) \to X(\omega)$ for all $\omega \in \omega$,
- almost sure convergence: $X_n(\omega) \to X(\omega)$ for all $\omega \in A$, some $A \in \mathscr{F}$ with $P(A) = 1$,
- convergence in probability: $\mathbb{P}(|X_n - X| > \varepsilon) \to 0$ for all $\varepsilon > 0$, we write $\lim^{\mathrm{p}} X_n = X$,
- $\mathscr{L}^p$-convergence: $X_n, X \in \mathscr{L}^p$ and $\mathbb{E}(|X_n - X|^p) \to 0$, we write $\lim^{\mathscr{L}^p} X_n = X$,
- convergence in distribution (weakly): $\mathbb{P}(X_n \leq x) \to \mathbb{P}(X \leq x)$ at all points x where $x \mapsto \mathbb{P}(X \leq x)$ is continuous.



⟦32⟧ Criterion for almost sure convergence – X_n converges to X almost surely iff for all $\varepsilon > 0$

$$\mathbb{P}(\sup_{k \geq n} |X_k - X_n| \geq \varepsilon) \to 0, \quad n \to \infty.$$

1.9 Limit theorems

⟦33⟧ Strong law of large numbers – If $(X_n)_{n=1,2,\ldots}$ is a sequence of i.i.d. random variables with $\mathbb{E}(|X_1|) < \infty$ then

$$\frac{\sum_{i=1}^n X_i}{n} \to \mathbb{E}(X_1)$$

almost surely as $n \to \infty$.

⟦**34**⟧ **Central limit theorem –** If $(X_n)_{n=1,2,\dots}$ is a sequence of i.i.d. random variables with $\sigma = \sqrt{\mathrm{Var(X)}} < \infty$ then

$$\frac{\sum_{i=1}^{n} X_i - n\mathbb{E}(X)}{\sigma\sqrt{n}} \to \Phi(x),$$

where Φ is the normal distribution function with mean 0 and variance 1, i.e.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du.$$

⟦**35**⟧ **Conditional expectation –**

We postpone the exact definition of conditional expectation to chapter 4.

Let $X_1, X_2, \dots$ be sequence of random variables (a discrete time stochastic process). Let $\mathbb{E}(Y|X_1, X_2, \dots, X_n)$ denote the conditional expectation of Y, given $X_1, X_2, \dots$. We shortly write $\mathbb{E}(\cdot|\mathscr{F}_k)$ instead of $\mathbb{E}(\cdot|X_1, \dots, X_k)$ and say that $Y \in \mathscr{F}_n$ if $Y = f(X_1, X_2, \dots, X_n)$ for some measurable function f. Moreover we say $A \in \mathscr{F}_n$ if $\mathbb{1}_A \in \mathscr{F}_n$.

We have the following rules:

A) $\mathbb{E}(Y|\mathscr{F}_n) = Y$ and $\mathbb{E}(YZ|\mathscr{F}_n) = Y\mathbb{E}(Z|\mathscr{F}_n)$ if $Y \in \mathscr{F}_n$.
($X_1, X_2, \dots, X_n$ contains all information about Y)

B) $\mathbb{E}(Y|\mathscr{F}_n) = \mathbb{E}(Y)$ if Y is independent of $X_1, X_2, \dots, X_n$.
($X_1, X_2, \dots, X_n$ contains no information about Y)

C) $\mathbb{E}(\mathbb{E}(Y|\mathscr{F}_n)) = \mathbb{E}(Y)$. This is the *law of total probability*.

(Compare with ⟦61⟧)

2 DISCRETE TIME MARTINGALES

2.1 Classic examples

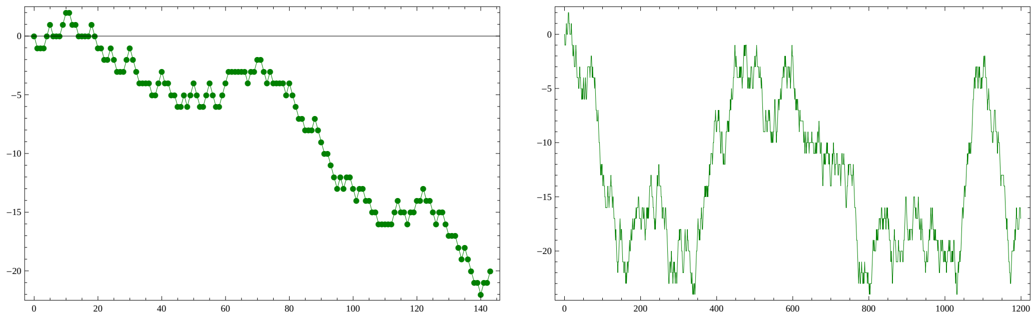


Figure 1: Coin tossing. Sample path of the process M_n.

⟦36⟧ **Coin tossing –** Consider a coin-tossing game, where a player can wins a Euro if head is tossed and loses 1 Euro if not. Let X_n denote the profit in the nth game (so $X_n \in \{-1,1\}$) and let M_n denote the players fortune after the nth coin toss.

(1) M_n is a function of $X_1,\ldots,X_n$: $M_n = \sum_{i=1}^n X_i$ (2) $\mathbb{E}(|M_n|) \le \sum_{i=1}^n \mathbb{E}(|X_i|) = n$.

(3) Since this is a fair game, we have $\mathbb{E}(X_n) = \mathbb{E}(X_{n-1})$. But more is true:

$$\mathbb{E}(M_n|\mathscr{F}_{n-1}) = M_{n-1}.$$

(this will be shown in an exercise, see also the example below).

⟦37⟧ **Martingales –** A sequence $\{M_k\}_{k\in\mathbb{N}}$ of random variables is a *martingale* with respect to $\{X_k\}_{k\in\mathbb{N}}$ if

(1) $M_n \in \mathscr{F}_n$,

(2) $\mathbb{E}(|M_n|) < \infty$,

(3) $\mathbb{E}(M_n|\mathscr{F}_{n-1}) = M_{n-1}$ for all $n \ge 1$.

Martingales have some very special properties as we will see. An obvious property is that martingales have a constant expectation (see exercise):

$$\mathbb{E}(M_n) = \mathbb{E}(M_1).$$

⟦38⟧ **Example 1 –** Let $\{X_k\}_{k\in\mathbb{N}}$ be a sequence of i.i.d. random variables (i.e. they are independent and have the same distribution) with $\mathbb{E}(X_1) < \infty$ and let $S_n = \sum_{i=1}^n X_i$. Then $M_n = S_n - n\mathbb{E}(X_1)$ is a martingale w.r.t. (X_k) (see exercise).

⟦39⟧ **Example 2 –** In the above situation, if $\mathbb{E}(X_1) = 0$ and $\mathrm{Var}(X_1) = \sigma^2$, then $M_n = S_n^2 - n\sigma^2$ is a martingale.

Proof. (1) Clearly M_n is a function of $X_1, \dots, X_n$.
(2) $\mathbb{E}(|M_n|) < \infty$ since $\mathbb{E}(S_n^2) = \mathbb{E}(\sum_{i,j=1}^n X_i X_j) = \sum_{i=1}^n \mathbb{E}(X_i^2) < \infty$.
(3) We have

$$\mathbb{E}(M_n|\mathscr{F}_{n-1}) = \mathbb{E}(S_n^2 - n\sigma^2|\mathscr{F}_{n-1}) = \mathbb{E}(S_n^2|\mathscr{F}_{n-1}) - n\sigma^2.$$

First $\mathbb{E}(S_n^2|\mathscr{F}_{n-1}) = \mathbb{E}(S_{n-1}^2 + 2S_{n-1}X_n + X_n^2|\mathscr{F}_{n-1})$. It is clear that, given $M_1, \dots, M_{n-1}$, the conditional expectation of the first term is $\mathbb{E}(S_{n-1}^2|\mathscr{F}_{n-1}) = S_{n-1}^2$. Moreover, using A), it follows from the independence of X_n from $X_1, \dots, X_n$

$$\mathbb{E}(S_{n-1}X_n|\mathscr{F}_{n-1}) = S_{n-1}\mathbb{E}(X_n|\mathscr{F}_{n-1}). = S_{n-1}\mathbb{E}(X_n) = 0$$

By the same reasoning we obtain $\mathbb{E}(X_n^2|\mathscr{F}_{n-1}) = \mathbb{E}(X_n^2) = \sigma^2$. Altogether

$$\mathbb{E}(M_n|\mathscr{F}_{n-1}) = S_{n-1}^2 + \sigma^2 - n\sigma^2 = M_{n-1}. \quad \square$$

⟦40⟧ **Example 3 –** If $\{X_k\}_{k\in\mathbb{N}}$ are non-negative and i.i.d. with $\mathbb{E}(X_1) = 1$ then $M_n = X_1 X_2 \cdots X_n$ is a martingale (see exercise)

⟦41⟧ **Definition 2.3 (Stopping times)** *A random variable τ with values in $\{0, 1, \dots\} \cup \{\infty\}$ is a stopping time w.r.t. $\{X_n\}_{n\in\mathbb{N}}$ if $\{\tau \le n\} \in \mathscr{F}_n$, $0 \le n$*

Let M_τ denote the value of the process $\{M_n\}$ at τ, or formally $M_\tau = \sum_{k=0}^\infty M_k \mathbb{1}_{\tau=k}$.

⟦42⟧ **Theorem 2.2 (Optional stopping theorem)** *Let $\{M_n\}$ be a martingale w.r.t. $\mathscr{F}_n$. Let $\{M_{n\wedge\tau}\}$ denote the stopped process, i.e.*

$$M_{n\wedge\tau} = \begin{cases} M_k & , \quad \tau \le n, \tau = k \\ M_n & , \quad \tau \ge n \end{cases}.$$

Then $\{M_{n\wedge\tau}\}$ is also a martingale w.r.t. $\mathscr{F}_n$

Proof. (1) $M_{n\wedge\tau} \in \mathscr{F}_n$?

$$M_{n\wedge\tau} \quad = \quad \underbrace{\mathbb{1}_{\tau\le n-1} \sum_{k=0}^{n-1} M_k \mathbb{1}_{\tau=k}}_{\in\mathscr{F}_{n-1}} + \underbrace{\mathbb{1}_{\tau\ge n}}_{\in\mathscr{F}_{n-1}} \underbrace{M_n}_{\in\mathscr{F}_n}.$$

(2) We obviously have $\mathbb{E}(|M_{n\wedge\tau}|) < \infty$.

(3) For the third property we use the representation from above

$$\begin{aligned}\mathbb{E}(M_{n\wedge\tau}|\mathscr{F}_{n-1}) &= \mathbb{E}(\mathbb{1}_{\tau\le n-1}\sum_{k=0}^{n-1} M_k \mathbb{1}_{\tau=k} + \mathbb{1}_{\tau\ge n} M_n|\mathscr{F}_{n-1}) \\ &= \mathbb{E}(\mathbb{1}_{\tau\le n-1}\sum_{k=0}^{n-1} M_k \mathbb{1}_{\tau=k}|\mathscr{F}_{n-1}) + \mathbb{E}(\mathbb{1}_{\tau\ge n} M_n|\mathscr{F}_{n-1})\end{aligned}$$

Using the martingale property of M_n,

$$\begin{aligned}&= \mathbb{1}_{\tau\le n-1}\sum_{k=0}^{n-1} M_k \mathbb{1}_{\tau=k} + \mathbb{1}_{\tau\ge n} M_{n-1} \\ &= \mathbb{1}_{\tau< n-1}\sum_{k=0}^{n-2} M_k \mathbb{1}_{\tau=k} + \mathbb{1}_{\tau\ge n-1} M_{n-1} = M_{(n-1)\wedge\tau}. \quad \square\end{aligned}$$

⟦43⟧ Coin tossing revisited –

How long does it take until the player's fortune is for the first time either A or $-B$, $-B<0<A$. We define

$$\tau = \min\{n : M_n = X_1 + \ldots + X_n \in \{A, -B\}\}.$$

Then

$$\{\tau \le n\} = \{X_1 \in \{A,-B\} \text{ or } X_2 \in \{A,-B\} \text{ or ... or } X_n \in \{A,-B\}\} \in \mathscr{F}_n,$$

where τ is a stopping time w.r.t. $\{M_n\}$. It follows that $\{M_{n\wedge\tau}\}$ is also a martingale and hence

$$\mathbb{E}(M_{n\wedge\tau}) = \mathbb{E}(M_1) = 0.$$

Since $\mathbb{P}(\tau<\infty)=1$ (see exercise) it follows that $M_{n\wedge\tau}\to M_\tau$ as $n\to\infty$ almost surely (actually $M_{n\wedge\tau}=M_\tau$ for large enough n) and $|M_{n\wedge\tau}|\le\max\{A,B\}$, so by dominated convergence

$$\mathbb{E}(M_\tau) = 0.$$

But $\mathbb{E}(M_\tau) = A\mathbb{P}(M_\tau = A) - B\mathbb{P}(M_\tau=-B)$ and hence

$$A\mathbb{P}(M_\tau = A) = B\mathbb{P}(M_\tau=-B) = B(1-\mathbb{P}(M_\tau = A)),$$

leading to the nice formulas

$$\mathbb{P}(M_\tau = A) = \frac{B}{A+B}, \quad \mathbb{P}(M_\tau = -B) = \frac{A}{A+B}. \tag{2.1}$$

2.4 Submartingales

⟦44⟧ **Submartingales –** A sequence $\{M_k\}_{k\in\mathbb{N}}$ of random variables is a *submartingale* with respect to $\{X_k\}_{k\in\mathbb{N}}$ if

(1) $M_n \in \mathscr{F}_n$,

(2) $\mathbb{E}(|M_n|) < \infty$,

(3) $\mathbb{E}(M_n|\mathscr{F}_{n-1}) \geq M_{n-1}$ for all $n \geq 1$.

submartingale.

If $\phi : \mathbb{R} \to \mathbb{R}$ is a convex function and M_n is a martingale then $\phi(M_n)$ is a submartingale, provided that $\mathbb{E}(|\phi(M_n)|) < \infty$ (see exercise 15). For example M_n^2 or $|M_n|$ are submartingales, if the integrability condition is fulfilled.

2.5 Doob's Inequalities

We present two important inequalities without proof.

Note first that if $\{M_n\}$ is a martingale and ϕ is a convex function then $\phi(M_n)$ is a submartingale if $\mathbb{E}(|\phi(M_n)|) < \infty$ for all $n \geq 0$. For example M_n^p and $|M_n|$ are submartingales.

⟦45⟧ **Theorem 2.4 (Doob's Maximal inequality)** *Let $\{M_n\}$ be a non-negative submartingale and define the maximal sequence*

$$M_n^* = \sup_{0\leq m\leq n} M_m.$$

Then for $\lambda > 0$

$$\lambda\mathbb{P}(M_n^* \geq \lambda) \leq \mathbb{E}(M_n \mathbb{1}_{M_n^*>\lambda}) \leq \mathbb{E}(M_n).$$

Moreover for any $p \geq 1$

$$\lambda^p\mathbb{P}(M_n^* \geq \lambda) \leq \mathbb{E}(M_n^p).$$

⟦46⟧ **Theorem 2.5 (Doob's $\mathscr{L}^p$ inequality)** *Let $\{M_n\}$ be a non-negative submartingale. Then for $p > 1$ and $n \geq 0$*

$$||M_n^*||_p \leq \frac{p}{p-1}||M_n||_p.$$

2.6 Martingale convergence

Together with optional stopping the convergence property is what makes martingales so valuable. Under quite loose conditions it is possible to show that $M_n \to M_\infty$ as $n \to \infty$ for some random variable M_∞.

We first show a convergence result that holds for (uniformly) $\mathscr{L}^2$-bounded martingales.

⟦47⟧ Theorem 2.6 ($\mathscr{L}^2$-bounded Martingale Convergence Theorem) *If $\{M_n\}$ is a martingale and $\mathbb{E}(M_n^2) \le B < \infty$ then there is a random variable $M_\infty \in \mathscr{L}^2$ with $\mathbb{E}(M_\infty^2) \le B < \infty$ and $M_n \to M$ almost surely and in $\mathscr{L}^2$, i.e.*

$$||M_n - M_\infty||_2 \to 0, \quad n \to \infty.$$

If X is $\mathscr{L}^2$-bounded then X is also $\mathscr{L}^1$-bounded (follows from $\mathrm{Var}(|X|) = \mathbb{E}(X^2) - \mathbb{E}(|X|)^2 \ge 0$). The opposite is not true.

Proof. Without loss of generality we let $M_1 = 0$. Otherwise take $\widehat{M}_k = M_k - M_1$. We are going to show that

$$\mathbb{P}(\sup_{k \ge m} |M_k - M_m| \ge \varepsilon) \to 0$$

as $m \to \infty$. Then M_k converges almost surely to a limit M_∞ and $\mathbb{E}(M_\infty^2) \le B$ by Fatou's lemma:

$$\mathbb{E}(\lim_{n\to\infty} M_n^2) \le \lim_{n\to\infty} \mathbb{E}(M_n^2) \le B.$$

Let $d_k = M_k - M_{k-1}$, then, if $j < k$

$$\begin{aligned} \mathbb{E}(d_k d_j | \mathscr{F}_{k-1}) &= \mathbb{E}(M_k M_j - M_{k-1} M_j - M_k M_{j-1} + M_{j-1} M_{k-1} | \mathscr{F}_{k-1}) \\ &= M_j M_{k-1} - M_{k-1} M_j - M_{k-1} M_{j-1} + M_{j-1} M_{k-1} = 0. \end{aligned}$$

So $\mathbb{E}(d_k d_j) = 0$ for $j \ne k$ and

$$\mathbb{E}(M_n^2) = \mathbb{E}\big(\big(\sum_{k=1}^n d_k\big)^2\big) = \sum_{k=1}^n \mathbb{E}(d_k^2).$$

Since $\mathbb{E}(M_n^2) \le B$ for all n, the sum on the right converges to

$$\sum_{k=1}^\infty \mathbb{E}(d_k^2) \le B$$

and it follows that $\mathbb{E}(d_k^2) \to 0$ as $k \to \infty$. The process $M_n' = (M_{k+m} - M_m)^2$ is a submartingale (exercise) and using Doob's inequality

$$\begin{aligned}\mathbb{P}(\sup_{k\geq 0}|M_{m+k}-M_m|\geq\varepsilon) &= \mathbb{P}(\sup M'_n\geq\varepsilon^2)\leq\frac{\mathbb{E}(M'_n)}{\varepsilon^2}=\frac{\mathbb{E}((M_{k+m}-M_m)^2)}{\varepsilon^2}\\ &= \frac{\mathbb{E}\Big(\big(\sum_{i=m+1}^{k+m}d_i\big)^2\Big)}{\varepsilon^2}=\frac{\sum_{i=m+1}^{k+m}\mathbb{E}\big({d_i}^2\big)}{\varepsilon^2}\\ &\leq \frac{\sum_{i=m+1}^{\infty}\mathbb{E}\big({d_i}^2\big)}{\varepsilon^2}\end{aligned}$$

which tends to zero as $m\to\infty$. Moreover,

$$\mathbb{E}\Big(\big(M_\infty-M_n\big)^2\Big)=\mathbb{E}\Big(\big(\sum_{k=n+1}^{\infty}d_k\big)^2\Big)=\sum_{k=n+1}^{\infty}\mathbb{E}(d_k^2)\to 0$$

as $n\to\infty$. □

Without proof we state the following important generalization of the last theorem for $\mathscr{L}^1$-bounded martingales.

[[48]] Theorem 2.8 ($\mathscr{L}^1$-bounded Martingale Convergence Theorem) *If $\{M_n\}$ is a martingale and $\mathbb{E}(|M_n|)\leq B<\infty$ then there is a random variable $M_\infty\in\mathscr{L}^1$ with $\mathbb{E}(|M_\infty|)\leq B<\infty$ and $M_n\to M$ almost surely.*

It is not clear, whether under the stated conditions in fact

$$\mathbb{E}\big(|M_n-M_\infty|\big)\to 0,\quad n\to\infty$$

3 BROWNIAN MOTION

3.0 Multivariate Gaussians and Brownian motion

[[49]] Definition (Multivariate Gaussians) *A d-dimensional random vector V has a multivariate Gaussian distribution with mean vector $\mu=(\mu_1,\dots,\mu_d)$ and covariance matrix*

$$\Sigma=\begin{pmatrix}\mathrm{Var}(V_1) & \mathrm{Cov}(V_1,V_2) & \dots & \mathrm{Cov}(V_1,V_d)\\ \mathrm{Cov}(V_1,V_2) & \mathrm{Var}(V_2) & \dots & \mathrm{Cov}(V_2,V_d)\\ \vdots & \vdots & \vdots & \vdots\\ \mathrm{Cov}(V_d,V_2) & \mathrm{Cov}(V_d,V_2) & \dots & \mathrm{Var}(V_d)\end{pmatrix}$$

if the density of V is given by

$$f(x)=f(x_1,x_2,\dots,x_d)=\frac{1}{\sqrt{(2\pi)^d\det\Sigma}}\exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big).$$

If V has a d-dimensional multivariate Gaussian distribution then each V_i, $i=1,2,\dots,d$ has a normal distribution with mean μ_i and standard deviation σ_i.

In general the V_i are not independent. However, if $V_1, V_2, \ldots, V_d$ are uncorrelated then Σ is a diagonal matrix and

$$f(x) = \frac{1}{\sqrt{(2\pi)^d \prod_{i=1}^d \sigma_i}} \exp\Big(-\frac{1}{2}\sum_{i=1}^d (x_i-\mu_i)^2/\sigma_1\Big) = \prod_{i=1}^d \frac{1}{\sqrt{2\pi\sigma_i}} \exp\Big(-\frac{1}{2\sigma_i}(x_i-\mu_i)^2\Big),$$

Hence $V_1, \ldots, V_d$ are independent normal random variables!

Characteristic function of Gaussian r.v.: If V has a d-dimensional multivariate Gaussian distribution then

$$\mathbb{E}(e^{i\theta^T V}) = e^{i\theta^T \mu - \frac{1}{2}\theta^T \Sigma \theta}.$$

In particular if $d = 1$

$$\mathbb{E}(e^{i\theta V}) = e^{i\theta\mu - \frac{\theta^2\sigma^2}{2}}.$$

⟦50⟧ **Definition 3.1 (Brownian motion)** *A standard Brownian motion on* $[0,T]$ *is a stochastic process* $\{B_t\}_{0\le t<T}$, $t \in [0,T]$, *such that*

(1) $B_0 = 0$

(2) Increments are independent, i.e.

$$B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_n}$$

are independent for any choice of $0 \le t_1 \le t_2 \le \ldots \le t_n < T$.

(3) $B_t - B_s$ *has a normal distribution with mean* 0 *and variance* $t-s$ *for* $0 \le s \le t < T$.

(4) $t \mapsto B_t$ *is a continuous function a.s.*

3.3 Wavelets

⟦51⟧ **The β functions –** Let

$$\beta(t) = \begin{cases} t & , \quad 0 \le t \le \frac{1}{2} \\ 1-t & , \quad \frac{1}{2} \le t \le 1 \\ 0 & , \quad \text{otherwise} \end{cases} \tag{3.1}$$

and define for $n = 2^j + k$, $j \ge 0$, $0 \le k < 2^j$,

$$\beta_n(t) = 2^{-j/2}\beta(2^j t - k)$$

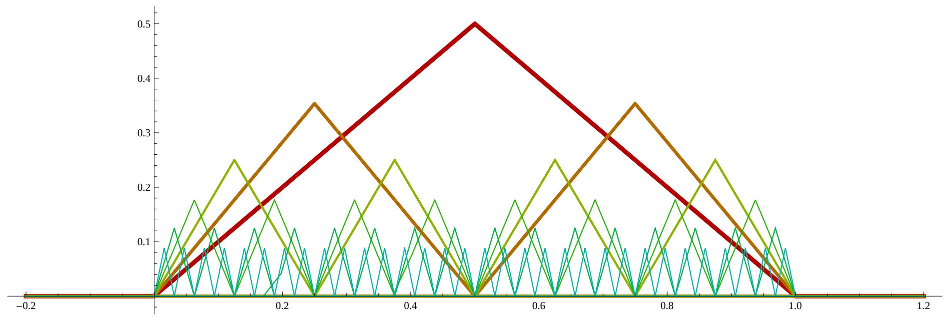


Figure 2: The functions $\beta_1, \ldots, \beta_{63}$.

[[52]] Theorem (Parseval's identity) *We have*

$$s \wedge t \quad = \quad \sum_{n=0}^{\infty} \beta_n(t)\beta_n(s).$$

Now let $Z_0, Z_1, \ldots$ be a sequence of i.i.d. normal random variables with mean 0 and variance 1 and let for $t \in [0,1]$

$$X_{t,m} = \sum_{n=0}^{m} \beta_n(t) Z_n$$

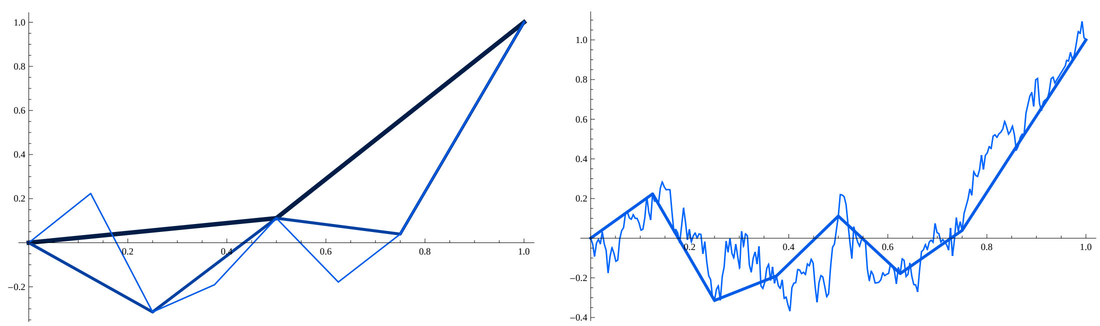


Figure 3: Approximation $X_{t,m}$ for $m = 1$, $m = 3$, $m = 7$ (left) and $m = 7$, $m = 255$ (right).

You can see in the above figures how the processes $X_{t,m}$ become more and more 'fractal' as n increases. The question is, whether $X_{t.m}$ actually converges as $n \to \infty$ (and then, in which sense).

3.4 Wavelet representation of Brownian motion

⟦53⟧ **Theorem 3.1a (Wavelet representation)** *The sum $X_{t,m}$ converges almost surely uniformly for $t \in [0,1]$ as $m \to \infty$. We denote the limit by*

$$X_t = \sum_{n=0}^{\infty} \beta_n(t) Z_n.$$

Proof. First note that

$$\sum_{n=M}^{\infty} \beta_n(t)|Z_n| \quad \leq \quad S \sum_{n=M}^{\infty} \sqrt{\log(n)}\beta_n(t), \tag{3.2}$$

where $S = \sup_{i \geq M} \frac{|Z_i|}{\sqrt{\log(i)}}$. We have for $M > 2^J$

$$\begin{aligned}
\sum_{n=M}^{\infty} \sqrt{\log(n)}\beta_n(t) &\leq \sum_{j=J}^{\infty} \sum_{k=0}^{2^j-1} \sqrt{\log(2^j+k)}\beta_{2^j+k}(t), \\
&\leq \sum_{j=J}^{\infty} \sqrt{j+1} \sum_{k=0}^{2^j-1} \beta_{2^j+k}(t), \\
&\leq \sum_{j=J}^{\infty} \sqrt{j+1}2^{-j/2}.
\end{aligned}$$

Since the rightmost sum in (3.2) tends to 0 uniformly as $J \to \infty$, it follows that we only have to show that S is a.s. finite. We have

$$\begin{aligned}
\mathbb{P}(|Z_n| \geq \sqrt{2\alpha\log(n)}) &= \frac{2}{\sqrt{2\pi}} \int_{\sqrt{2\alpha\log(n)}}^{\infty} e^{-u^2/2}\, du \\
&\overset{u=\sqrt{2\log w}}{=} \frac{2}{\sqrt{\pi}} \int_{n^\alpha}^{\infty} \frac{1}{\sqrt{\log(w)}} \frac{1}{w^2}\, dw \\
&\leq \frac{2}{\sqrt{\pi}} n^{-\alpha}.
\end{aligned}$$

For all $\alpha > 1$ we have $\sum_{n=1}^{\infty} \mathbb{P}(|Z_n| \geq \sqrt{2\alpha\log(n)}) < \infty$ and by Borel-Cantelli

$$\mathbb{P}(\frac{|Z_n|}{\sqrt{\log(n)}} \geq \sqrt{2\alpha} \text{ i.o.}) = 0.$$

Consequently S is finite with probability one. □

⟦54⟧ **Theorem 3.1b (Brownian motion representation)** *The process $X_t = \sum_{n=0}^{\infty} \beta_n(t) Z_n$ is a standard Brownian motion.*

Proof. Clearly $X_0 = 0$ and X_t is continuous, because it is the uniform limit of continuous functions. It remains to show that X_t has independent increments and that the increments $X_t - X_s$, $t > s$ have a normal distribution with mean 0 and variance $t - s$.

We calculate the *finite dimensional distributions*

$$\mathbb{P}(X_{t_1} \le x_1, \dots, X_{t_n} \le x_n).$$

Let $0 \le t_1 \le t_n < T$. We calculate the joint characteristic function

$$\begin{aligned}
\mathbb{E}(\exp\big(i \sum_{j=1}^{m} \theta_j X_{t_j}\big)) &= \mathbb{E}(\exp\big(i \sum_{j=1}^{m} \theta_j \sum_{n=0}^{\infty} \beta_n(t_j) Z_n\big)) \\
&= \prod_{n=0}^{\infty} \mathbb{E}(\exp\big(i Z_n \sum_{j=1}^{m} \theta_j \beta_n(t_j)\big)) \\
&= \prod_{n=0}^{\infty} \exp\big(-\frac{1}{2}(\sum_{j=1}^{m} \theta_j \beta_n(t_j))^2\big) \\
&= \exp\Big(-\frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \theta_j \theta_i \sum_{n=0}^{\infty} \beta_n(t_j) \beta_n(t_i)\Big) \\
&= \exp\Big(-\frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \theta_j \theta_i (t_i \wedge t_j)\Big),
\end{aligned}$$

which is the characteristic function of a multivariate Gaussian with mean 0 and covariance matrix $\Sigma = (t_i \wedge t_j)$.

We still have to show the independence of the increments. For $i < j$

$$\begin{aligned}
\mathbb{E}((X_{t_i} - X_{t_{i-1}})(X_{t_j} - X_{t_{j-1}})) &= \mathbb{E}(X_{t_i} X_{t_j}) - \mathbb{E}(X_{t_i} X_{t_{j-1}}) - \mathbb{E}(X_{t_{i-1}} X_{t_j}) + \mathbb{E}(X_{t_{i-1}} X_{t_{j-1}}) \\
&= t_i - t_i - t_{i-1} + t_{i-1} = 0,
\end{aligned}$$

so $X_{t_i} - X_{t_{i-1}}$ and $X_{t_j} - X_{t_{j-1}}$ are uncorrelated and hence (having a two-dimensional Gaussian distribution) independent. □

⟦55⟧ **Brownian motion on** $[0,\infty)$ **–** We have defined a standard Brownian motion on $[0,1]$. To obtain a standard Brownian motion on $[0,\infty)$ we simply glue together Brownian motions on $[n, n+1)$. We write for the resulting process B_t.

To be more precise, for $t \in [n, n+1)$ we let

$$B_t = \sum_{k=1}^{n} X_1^{(k)} + X_{t-n}^{(k+1)},$$

where $\{X_t^{(k)}\}_{k=1,2,\dots}$ are independent copies of our process X_t.

For other constructions using the β_i – in particular the construction of a Brownian bridge – see exercise 17.

4 CONTINUOUS TIME MARTINGALES

4.2 Conditional Expectation

⟦56⟧ **σ-fields and information –** Given a probability space $(\omega, \mathscr{F}, \mathbb{P})$ let $\mathscr{G}$ be a be sub-σ-fields of $\mathscr{F}$ (i.e. $A \in \mathscr{G} \Rightarrow A \in \mathscr{F}$).

Suppose that we know for any event $A \in \mathscr{G}$ whether $\omega \in A$ or not (A happened or not). Then there are still events in $\mathscr{F} \setminus \mathscr{G}$ for which we don't know this information. In a sense $\mathscr{G}$ contains a certain amount of information and any larger σ-field holds more information.

The smallest σ-field is $\{\emptyset, \omega\}$, giving us no information whether $\omega \in A$ for sets $A \neq \omega \in \mathscr{F}$.

⟦57⟧ **Example –** (compare with Exercise 1). Consider $\omega = [0,1]$ with $\mathscr{F} = \mathscr{B}(\omega)$ and $\mathbb{P} = \lambda$ the Lebesgue-measure. Let X be a random variable defined by

$$X(\omega) = \begin{cases} 1 & , \quad \omega \in [0, 1/3] \\ 2 & , \quad \omega \in (1/3, 2/3] \\ 3 & , \quad \omega \in (2/3, 1] \end{cases}.$$

The expectation of X is given by $\mathbb{E}(X) = \frac{1+2+3}{3} = \frac{6}{3} = 2$. The smallest σ-field such that X is measurable is $\sigma(X) = \sigma(\{[0,1/3], (1/3, 2/3]\})$. Now consider a second random variable Z

$$Z(\omega) = \begin{cases} 3 & , \quad \omega \in [0, 1/3] \\ 0 & , \quad \omega \in (1/3, 1] \end{cases}.$$

The smallest σ-field such that Z is measurable is $\mathscr{G} = \sigma(Z) = \sigma(\{[0,1/3]\})$.

Given $Z = 3$ the conditional expectation $\mathbb{E}(X|Z = 3)$ is obviously 1. Similarly $\mathbb{E}(X|Z = 0) = 5/2$, so we could define a random variable

$$Y(\omega) = \begin{cases} 1 & , \quad \omega \in [0, 1/3] \\ 5/2 & , \quad \omega \in (1/3, 1] \end{cases}.$$

⟦58⟧ **Conditional Expectation –** We observe the following two properties of Y:

- Y is $\mathscr{G}$-measurable.
- $\mathbb{E}(Y \mathbb{1}_A) = \mathbb{E}(X \mathbb{1}_A)$, $A \in \mathscr{G}$.

⟦59⟧ **Definition 4.1 (Conditional Expectation)** *Let X be an integrable random variable in the probability space $(\omega, \mathscr{F}, \mathbb{P})$. If $\mathscr{G}$ is a sub-σ-field of $\mathscr{F}$ and Y is a random variable such that*

- *Y is $\mathscr{G}$-measurable.*
- *$\mathbb{E}(Y \mathbb{1}_A) = \mathbb{E}(X \mathbb{1}_A)$, for all $A \in \mathscr{G}$.*

Then Y is called a conditional expectation of X w.r.t. $\mathscr{G}$. We write $Y = \mathbb{E}(X|\mathscr{G})$.

⟦60⟧ **Uniquenes and existence –** $\mathbb{E}(X|Z)$ is not unique. If $Y' = \mathbb{E}(X|Z)$ a.s. then Y' is also a conditional expectation, given Z. We say that Y and Y' are *versions* of the conditional expectation.

We don't prove that the conditional expectation $\mathbb{E}(X|Z)$ actually exists (see ☞). ☞ (5)

> ⟦61⟧ **Rules for conditional expectation (see Exercise 18) –**
>
> A) $\mathbb{E}(X|\mathscr{G}) = X$ and $\mathbb{E}(XZ|\mathscr{G}) = X\mathbb{E}(Z|\mathscr{G})$ if X is $\mathscr{G}$-measurable.
> ($\mathscr{G}$ contains all information about X)
>
> B) $\mathbb{E}(X|\mathscr{G}) = \mathbb{E}(X)$ if $\sigma(X)$ is independent of $\mathscr{G}$. ($\mathscr{G}$ contains no information about X)
>
> C) If $\mathscr{H} \subseteq \mathscr{G}$ then $\mathbb{E}(\mathbb{E}(X|\mathscr{G})|\mathscr{H}) = \mathbb{E}(X|\mathscr{H})$. In particular $\mathbb{E}(\mathbb{E}(X|\mathscr{G})) = \mathbb{E}(X)$.

4.3 Uniform integrability

> ⟦62⟧ **Definition 4.2 (Uniform Integrability)** *A collection $\mathscr{C}$ of random variables is uniformly integrable if*
>
> $$\sup_{Z \in C} \mathbb{E}(|Z| \mathbb{1}_{|Z|>x}) \to 0$$
>
> *as $x \to \infty$.*

☞

☞ (6)

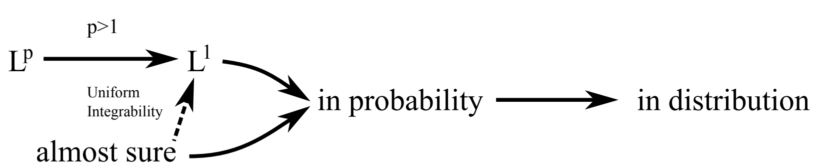
> ⟦63⟧ **Lemma 4.4 (Conditions for uniform integrability)** *If for all $Z \in \mathscr{C}$*
>
> $$\mathbb{E}(\phi(|Z|)) \leq B < \infty$$
>
> *for some constant B and some function ϕ with $\phi(x)/x \to \infty$ as $x \to \infty$ then $\mathscr{C}$ is uniformly integrable.*

☞

☞ (7)

⟦64⟧ **Why uniform integrability? –**

> ⟦65⟧ **Lemma 4.1 (Uniform integrability and $\mathscr{L}^1$-convergence)** *If $\{Z_n\}_{n\in\mathbb{N}}$ is a uniformly integrable sequence with $Z_n \to Z$ almost surely, then $X_n \to X$ also in $\mathscr{L}^1$ (so $Z \in \mathscr{L}^1$).*



Proof. By Fatou's lemma

$$\sup_{n\geq 1}\mathbb{E}(|Z_n|\mathbb{1}_{|Z_n|>x})\geq \liminf_{n\to\infty}\mathbb{E}(|Z_n|\mathbb{1}_{|Z_n|>x})\geq \mathbb{E}(|Z|\mathbb{1}_{|Z|>x})$$

and thus, writing $\rho(x)=\sup_{n\geq 1}\mathbb{E}(|Z_n|\mathbb{1}_{|Z_n|>x})$,

$$\mathbb{E}(|Z|)=\mathbb{E}(|Z|\mathbb{1}_{|Z|>x})+\mathbb{E}(|Z|\mathbb{1}_{|Z|>x})\leq \rho(x)+x.$$

and hence $Z\in\mathscr{L}^1$. Next

$$|Z_n-Z|=|Z_n-Z|\mathbb{1}_{|Z_n|\leq x}+|Z_n-Z|\mathbb{1}_{|Z_n|>x}\leq |Z_n-Z|\mathbb{1}_{|Z_n|\leq x}+|Z|\mathbb{1}_{|Z_n|>x}+|Z_n|\mathbb{1}_{|Z_n>x|}.$$

Clearly $|Z_n-Z|\mathbb{1}_{|Z_n|\leq x}\leq x+|Z|\in\mathscr{L}^1$ so by dominated convergence

$$\mathbb{E}(|Z_n-Z|\mathbb{1}_{|Z_n|\leq x})\to 0$$

as $n\to\infty$. For the second term we obtain $|Z|\mathbb{1}_{|Z_n|>x}\leq |Z|\in\mathscr{L}^1$ and hence

$$\mathbb{E}(|Z|\mathbb{1}_{|Z_n|>x})\to \mathbb{E}(|Z|\mathbb{1}_{|Z|>x})\leq \rho(x).$$

Finally $\mathbb{E}(|Z_n|\mathbb{1}_{|Z_n>x|})\leq\rho(x)$. It follows that $\limsup_{n\to\infty}\mathbb{E}(|Z_n-Z|)\leq 2\rho(x)$, but $\rho(x)$ can be made arbitrarily small if x is large enough. □

4.4 Continuous Time Martingales

⟦**66**⟧ **Filtration –** Let $\{\mathscr{F}_t\}_{t\geq 0}$ be a collection σ-fields, such that $\mathscr{F}_s\subseteq\mathscr{F}_t$ for $s\leq t$, then $\{\mathscr{F}_t\}_{t\geq 0}$ is called a *filtration* .

If for any $t\geq 0$, X_t is $\mathscr{F}_t$-measurable, then we say that the process $\{X_t\}_{t\geq 0}$ is *adapted* to the filtration $\mathscr{F}_t$.

⟦**67**⟧ **Definition (Martingales)** *A adapted process* $\{X_t\}_{t\geq 0}$ *is a martingale if*

(1) $\mathbb{E}(|X_t|)<\infty$,

(2) $\mathbb{E}(X_t|\mathscr{F}_s)=X_s$ *for all* $0\leq s\leq t<\infty$.

$\{X_t\}_{t\geq 0}$ *is a submartingale if instead of (2)* $\mathbb{E}(X_t|\mathscr{F}_s)\geq X_s$.

⟦**68**⟧ **Example (Poisson process) –** Let $\{N_t\}_{t\geq 0}$ be a Poisson process and define $X_t=N_t-\lambda t$. We show that X_t is a martingale (w.r.t. the filtration $\sigma(N_s:0\leq s\leq t)$).

Clearly $\mathbb{E}(|X_t|)<\infty$ and X_t is adapted. Moreover

$$\begin{aligned}\mathbb{E}(X_t|\mathscr{F}_s) &= \mathbb{E}(N_t-N_s+N_s|\mathscr{F}_s)-\lambda t=\mathbb{E}(N_t-N_s)+N_s-\lambda t\\ &= \lambda(t-s)+N_s-\lambda t=N_s-\lambda s=X_s.\end{aligned}$$

(see also Exercise 19).

⟦69⟧ **Stopping times –** A stopping time is a random variable τ with values in $[0,\infty)\cup\{\infty\}$ such that

$$\{\tau \le t\} \in \mathscr{F}_t, \quad \forall t \ge 0.$$

⟦70⟧ **Standard Brownian filtration –**

⟦71⟧ **Definition (Standard Brownian filtration)** *The standard Brownian filtration on $[0,T]$ is given by the smallest σ-field that contains $\sigma(B_s : s \le t)$ and the collection of subsets of null sets in $\sigma(B_s : s \le T)$.*
We extend $\mathbb{P}$ so that it assigns measure 0 to all subsets of null-sets.

The standard Brownian filtration fulfils the so called *standard conditions*:

⟦72⟧ **Definition (Usual conditions)** *A filtration $\mathscr{F}_t$ fulfils the usual conditions if*

(1) *$\mathscr{F}_0$ contains all subsets of null sets in $\mathscr{F}$.*

(2) *$\mathscr{F}_t$ is right-continuous, i.e. $\mathscr{F}_t = \mathscr{F}_{t+} := \bigcap_{s:s>t} \mathscr{F}_s$.*

⟦73⟧ **Some theorems –** The following results are stated without proof. They are the continuous time counterparts of theorems we already saw for the discrete time martingales.

⟦74⟧ **Theorem 4.1 (Doob's Optional Stopping Theorem)** *If $\{M_t\}_{t\ge 0}$ is a continuous martingale w.r.t. a filtration $\mathscr{F}_t$ that satisfies the usual conditions and if τ is a stopping time w.r.t. $\mathscr{F}_t$ then $M_{\tau\wedge t}$ is also a continuous martingale w.r.t. $\mathscr{F}_t$.*

⟦75⟧ **Theorem 4.2 (Maximal inequality)** *Let $M_t^* = \sup_{0\le t\le T} M_t$. If $\{M_t\}_{t\ge 0}$ is a continuous nonnegative submartingale and $\lambda > 0$ then*

$$\lambda^p \mathbb{P}(M_t^* > \lambda) \le \mathbb{E}(M_T^p).$$

If $M_T \in \mathscr{L}^p$ for some $p > 1$ then $||M_t^*||_p \le \frac{p}{p-1} ||M_T||_p$.

⟦76⟧ **Theorem 4.3 (Martingale limit theorem)** *If a continuous martingale $\{M_t\}_{t\ge 0}$ satisfies $\mathbb{E}(|M_t|^p) \le B$ for some $p \ge 1$ and all $t \ge 0$, then there is a random variable M_∞ with $\mathbb{E}(|M_\infty|^p) \le B$ such that $M_t \to M_\infty$ almost surely. If $p > 1$ then $||M_t - M_\infty||_p \to 0$.*

4.5 Classic Brownian Motion Martingales

See exercise 20.

6 THE ITÔ INTEGRAL

6.0 Bounded variation and Lebesgue-Stieltjes integral

⟦77⟧ **Bounded variation –** A real valued function $G:[0,T]\to\mathbb{R}$ is of *bounded variation*, if

$$\sup\sum_{i=1}^{n}|G(t_i)-G(t_{i-1})|<\infty,$$

where the supremum is taken over all partitions of the interval $0=t_0<t_1<\ldots<t_n=T$.

A function of bounded variation is the difference of two nonnegative monotonically increasing functions:

$$G(x)=G_1(x)-G_2(x).$$

⟦78⟧ **Lebesgue-Stieltjes integral –**

It is possible to define measures μ_1,μ_2 on $\mathscr{B}([0,T])$ in a way such that

$$\mu_i([x,y])=G_i(y)-G_i(x),\quad i=1,2,$$

for intervals $[x,y]\subseteq[0,T]$.

The *Lebesgue-Stieltjes integral* of a bounded function $f:[0,T]\to\mathbb{R}$ w.r.t. G is defined as the difference of two well defined Lebesgue-integrals:

$$\int_0^T f(x)\,dG(x)=\int_0^T f(x)\,d\mu_1(x)-\int_0^T f(x)\,d\mu_2(x).$$

If f is continuous then

$$\int_0^T f(x)\,dG(x)=\lim_{\varepsilon\to 0}\sum_{i=1}^{n}f(t_i^*)\left|G(t_i)-G(t_{i-1})\right|<\infty,$$

with $t_i^*\in[t_{i-1},t_i]$ and $\varepsilon=\max\{t_i-t_{i-1}:i=1,2,\ldots,n\}$.

If $G(0)=0$ then one can show by using integration by parts that

$$\int_0^t G(u)\,dG(u)\quad=\quad\frac{1}{2}G(t)^2.$$

We will see that this no longer holds if we replace the Lebesgue-Stieltjes integral by a Itô integral and the function $G(t)$ by Brownian motion B_t.

6.1 Definition of the Itô Integral for functions in $\mathscr{H}^2$

[[79]] **Itô integral not a Lebesgue-Stieltjes integral –** We want to define a stochastic integral

$$\int_0^t f(\omega,s)\,dB_s, \quad t \in [0,T],$$

where B_t is a Brownian motion and $f(\omega,t)$ is some (random) function. If we would interpret the integral as a Lebesgue-Stieltjes integral we would need B_t to be of bounded variation. But B_t is almost surely nowhere of bounded variation.

We are going to define the stochastic integral as a $\mathscr{L}^2$-limit of integrals of simple functions.

[[80]] **Good integrands –** We also need 'good integrands' $f(\omega,t)$ to be able to define a suitable stochastic integral.

[[81]] **Definition (The class $\mathscr{H}^2$)** *A function $f : \omega \times [0,T] \to \mathbb{R}$ is in $\mathscr{H}^2$ if*

(1) f is measurable: f is $\mathscr{F}_T \times \mathscr{B}([0,T])$-measurable, i.e.

$$f^{-1}(C) \in \mathscr{F}_T \times \mathscr{B}([0,T]), \quad \forall C \in \mathscr{B}(\mathbb{R}),$$

where $\mathscr{F}_T \times \mathscr{B}([0,T])$ is the smallest σ-field containg all sets of the form $A \times B$, $A \in \mathscr{F}_T$ and $B \in \mathscr{B}([0,T])$.

(2) f is adapted: $f(\cdot,t)$ is $\mathscr{F}_t$-measurable for all $t \in [0,T]$

(3) $\mathbb{E}\big(\int_0^T f^2(\omega,t)\,dt\big) < \infty$, i.e. $f \in \mathscr{L}^2(d\mathbb{P} \times dt)$.

Note that

$$||f||_{\mathscr{L}^2(d\mathbb{P}\times t)} = \Big(\int f^2(\omega,t)\,d(\mathbb{P}(\omega) \times dt)\Big)^{1/2} = \Big(\mathbb{E}\Big(\int_0^T f^2(\omega,t)\,dt\Big)\Big)^{1/2}.$$

The construction of a stochastic integral for $f \in \mathscr{H}^2$ follows ideas that are similar to the construction of the Lebesgue integral. Starting point are simple functions, for which a integral is easily defined.

[[82]] **Simple functions –** Let $0 = t_0 < t_1 < \ldots < t_n = T$. Simple functions in $\mathscr{H}^2$ are functions $f \in \mathscr{H}^2$ such that

$$f(\omega,t) = \sum_{i=0}^{n-1} a_i(\omega)\mathbb{1}_{t_i < t \le t_{i+1}},$$

with a_i being $\mathscr{F}_{t_i}$-measurable and $\mathbb{E}(a_i^2) < \infty$. The class of simple functions in $\mathscr{H}^2$ is denoted by $\mathscr{H}_0^2$.

Note the measurability condition on a_i. Knowledge of $\{B_s : s \in [0,t]\}$ allows us to find $a_i(w)$ for $t_i \le t$. Hence $f(\omega,t)$ is $\mathscr{F}_{t_i}$-measurable and a member of $\mathscr{H}^2$.

⟦83⟧ **Itô integral for simple functions –** For $f \in \mathscr{H}_0^2$ we define

$$I(f)(\omega) = \sum_{i=0}^{n-1} a_i(\omega)\big(B_{t_{i+1}} - B_{t_i}\big)$$

Suppose that $f_n \in \mathscr{H}_0^2$ and $f_n \to f \in \mathscr{H}^2$. If the mapping $I : \mathscr{H}_0^2 \to \mathscr{L}^2(d\mathbb{P})$ would be continuous then we could define $I(f)$ by the limit of $I(f_n)$ as $f_n \to f$.

⟦84⟧ **Theorem 6.1 (Itô 's Isometry on $\mathscr{H}_0^2$)** *For $f \in \mathscr{H}_0^2$ we have*

$$||I(f)||_2 = ||f||_{\mathscr{L}^2(d\mathbb{P}\times t)}.$$

Hence $I : \mathscr{H}_0^2 \to \mathscr{L}^2(d\mathbb{P})$ is an isometry and in particular continuous.

Proof. The proof is straightforward:

$$\begin{aligned} ||I(f)||_2 &= \mathbb{E}(I(F)^2) = \mathbb{E}\Big(\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} a_i(\omega)a_j(\omega)\big(B_{t_{i+1}} - B_{t_i}\big)\big(B_{t_{j+1}} - B_{j_i}\big)\Big) \\ &= \sum_{i=0}^{n-1} \mathbb{E}(a_i^2)\mathbb{E}\Big(\big(B_{t_{i+1}} - B_{t_i}\big)^2\Big) = \sum_{i=0}^{n-1} \mathbb{E}(a_i^2)\big(t_{i+1} - t_i\big), \end{aligned}$$

where we used that $a_i \in \mathscr{F}_{t_i}$ and hence independent from $B_{t_{i+1}} - B_{t_i}$. Similarly

$$||f||_{\mathscr{L}^2(d\mathbb{P}\times t)} = \mathbb{E}\Big(\int_0^T f^2(\omega,t)\,dt\Big) = \int_0^T \mathbb{E}\Big(\sum_{i=0}^{n-1} a_i^2 \mathbb{1}_{t_i < t \le t_{i+1}}\Big)\,dt = \sum_{i=0}^{n-1} \mathbb{E}(a_i^2)\big(t_{i+1} - t_i\big). \quad \square$$

As the next result shows, we are lucky: any $f \in \mathscr{H}^2$ is actually a limit of simple functions.

⟦85⟧ **Approximation –** For any function $f \in \mathscr{H}^2$ there is a sequence of functions in $\mathscr{H}_0^2$ such that

$$||f_n - f||_{\mathscr{L}^2(d\mathbb{P}\times t)} \to 0, \quad n \to \infty.$$

Almost nothing can stop us now to define $I(f)$ as the limit of the $I(f_n)$.

[[86]] Itô integral in $\mathscr{H}^2$ – The limit

$$I(f) = \mathscr{L}^2 - \lim_{n\to\infty} I(f_n)$$

exists and defines the Itô integral for $f \in \mathscr{H}^2$ (for any other sequence f'_n with $f'_n \to f$ the limit is the same). Moreover, the Itô isometry holds also in $\mathscr{H}^2$, i.e $||I(f)||_2 = ||f||_{\mathscr{L}^2(d\mathbb{P}\times t)}$.

Sketch of proof: We know that there is a sequence f_n in $\mathscr{H}^2_0$ such that

$$||f_n - f||_{\mathscr{L}^2(d\mathbb{P}\times t)} \to 0.$$

So f_n is a Cauchy-sequence in $\mathscr{L}^2(d\mathbb{P}\times dt)$ and hence $I(f_n)$ is a Cauchy sequence in $\mathscr{L}^2$, because of the Itô isometry. $\mathscr{L}^2$ is complete and so

$$I(f_n) \to I(f)$$

in $\mathscr{L}^2$. Given another approximating sequence $f'_n \in \mathscr{H}^2_0$ with $||f'_n - f||_{\mathscr{L}^2(d\mathbb{P}\times t)} \to 0$ then

$$||f'_n - f_n||_{\mathscr{L}^2(d\mathbb{P}\times t)} \le ||f'_n - f||_{\mathscr{L}^2(d\mathbb{P}\times t)} + ||f_n - f||_{\mathscr{L}^2(d\mathbb{P}\times t)} \to 0$$

too, and again by the Itô isometry $||I(f'_n) - I(f_n)||_2 \to 0$.
Moreover $||f_n - f||_{\mathscr{L}^2(d\mathbb{P}\times t)} \ge ||f_n||_{\mathscr{L}^2(d\mathbb{P}\times t)} - ||f||_{\mathscr{L}^2(d\mathbb{P}\times t)}$, so

$$||f_n||_{\mathscr{L}^2(d\mathbb{P}\times t)} \to ||f||_{\mathscr{L}^2(d\mathbb{P}\times t)}.$$

Similarly $||I(f_n)||_2 \to ||I(f)||_2$, so $||I(f)||_2 = ||f||_{\mathscr{L}^2(d\mathbb{P}\times t)}$ since $||f_n||_{\mathscr{L}^2(d\mathbb{P}\times t)} = ||I(f_n)||_2$. □

[[87]] Integral symbol – From now on we use the familiar integral symbol for the Itô - integral for functions in $\mathscr{H}^2$,

$$\int_0^T f(\omega,t)\,dB_t := I(f).$$

[[88]] Itô integral as a stochastic process – How can we define a process $\int_0^t f(\omega,s)\,dB_s$?

Clearly, for each $t \in [0,T]$

$$J_t = \int_0^T \mathbb{1}_{[0,t]}(s) f(\omega,s)\,dB_s$$

is well-defined as a member of $\mathscr{L}^2$, but we can modify each J_t on some null-set A_t and obtain still a version of the integral. But we don't know whether $\mathbb{P}(\bigcup_{t\in[0,T]} A_t) = 0$ so that the process $\{J_t\}_{t\in[0,T]}$ could be ambigous on a set with positive probability.

[[89]] Definition (Versions) *Two stochastic processes X_t and X'_t on $t \in [0,T]$ are called versions of each other if $X_t = X'_t$ a.s. for each $t \in [0,T]$.*

Among the versions of J_t there is one that is not only continuous, but is much more:

⟦90⟧ **Theorem 6.2 (Itô integral)** *The process J_t has a version which is a continuous martingale w.r.t. standard Brownian filtration. For this version we write*

$$\int_0^t f(\omega,s)\,dB_s, \quad t \in [0,T].$$

Hence stochastic integration is a martingale producing machine. Moreover we gain even more integrals if we utilize the Itôisometry.

It follows from Itô's isometry on $\mathscr{H}^2$ that $\left(\int_0^t f(\omega,s)\,dB_s\right)^2 - \int_0^t f^2(\omega,s)\,ds$ is a martingale.

☞ ☞ (8)

⟦91⟧ **Preliminary summary –** We defined the Itô integral for functions $f \in \mathscr{H}^2$, i.e. functions $f : \omega \times [0,T] \to \mathbb{R}$ for which

(1) f is *measurable*,

(2) f is *adapted*,

(3) $\mathbb{E}\left(\int_0^T f^2(\omega,t)\,dt\right) < \infty$.

Then $\int_0^t f(\omega,s)\,dB_S$ is defined to be the continuous martingale version of the limit $\mathscr{L}^2$-limit of the integrals $\int_0^T f_n(\omega,s)\,dB_S$, where $f_n(\omega,s) \in \mathscr{H}_0^2$ are simple functions such that $f_n(\omega,s) \to \mathbb{1}_{s<t} f(\omega,s)$ in $\mathscr{L}^2(d\mathbb{P} \times dt)$.

⟦92⟧ **Properties of the Itô integral –** Let also $g \in \mathscr{H}^2$ and let $a, b \in \mathbb{R}$. Then

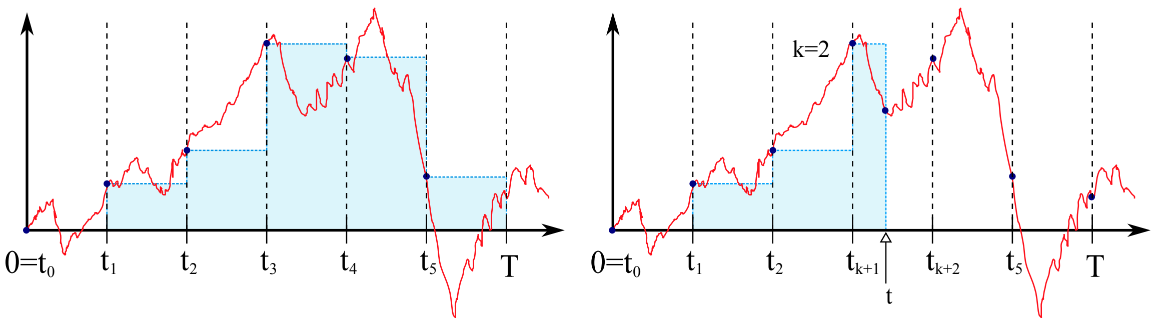
- $\int_0^t (af + bg)\,dB_s = a\int_0^t f\,dB_s + b\int_0^t g\,dB_s$
- $\mathbb{E}(\int_0^t f\,dB_s) = 0$ and $\mathrm{Var}(\int_0^t \mathrm{f dB_s}) = \mathbb{E}(\int_0^t \mathrm{f}^2\,\mathrm{ds})$
- $\int_0^t f\,dB_s$ is a continuous martingale.
- The process $(\int_0^t f\,dB_s)^2 - \int_0^t f^2\,ds$ is a martingale.

6.4 An explicit Calculation

We try to evaluate the stochastic integral $\int_0^t B_s\,dB_s$. We have to approximate $f(\omega,t) = B_t$ by functions in $\mathscr{H}_0^2$. We chose

$$f_n(\omega,t) = \sum_{i=0}^{n-1} B_{t_i} \mathbb{1}_{t_i < t \le t_{i+1}},$$

where $t_i = \frac{i}{n}T$.



Clearly $f_n \in \mathscr{H}_0^2$ and

$$\begin{aligned}
\|f_n - f\|^2_{\mathscr{L}^2(d\mathbb{P}\times t)} &= \mathbb{E}\Big(\int_0^T \big(B_t - \sum_{i=0}^{n-1} B_{t_i} \mathbb{1}_{t_i<t\le t_{i+1}}\big)^2\,dt\Big) = \mathbb{E}\Big(\int_0^T \sum_{i=0}^{n-1} \big(B_t - B_{t_i}\big)^2 \mathbb{1}_{t_i<t\le t_{i+1}}\,dt\Big) \\
&= \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} (t-t_i)\,dt = \frac{1}{2}\sum_{i=0}^{n-1} (t_{i+1}-t_i)^2 = \frac{1}{2}\sum_{i=0}^{n-1} \frac{T^2}{n^2} = \frac{T^2}{2n}.
\end{aligned}$$

As $n \to \infty$ this tends to zero, so $f_n \to f$ in $\mathscr{L}^2(dP \times dt)$. Let $k = \max\{i : t_{i+1} \le t\}$, so that

$$\frac{(k+1)T}{n} = t_{k+1} \le t < t_{k+2} = \frac{(k+2)T}{n}.$$

Then

$$\int_0^t B_s\,dB_s = \int_0^T B_s\,dB_s = \overset{\mathscr{L}^2}{\lim_{n\to\infty}} \int_0^T f_n(\omega,s)\,dB_s = \overset{\mathscr{L}^2}{\lim_{n\to\infty}} \sum_{i=0}^{k} B_{t_i}(B_{t_{i+1}} - B_{t_i}) + B_{t_{k+1}}(B_t - B_{k+1}).$$

For the last term

$$\mathbb{E}\Big(\big(B_{t_{k+1}}(B_t - B_{k+1})\big)^2\Big) = t_{k+1}(t - t_{k+1}) \le tT/n.$$

Moreover, since $2a(b-a) = b^2 - a^2 - (a-b)^2$

$$B_{t_i}(B_{t_{i+1}} - B_{t_i}) = \frac{1}{2}(B^2_{t_{i+1}} - B^2_{t_i}) - \frac{1}{2}(B_{t_{i+1}} - B_{t_i})^2$$

and hence

$$\sum_{i=0}^{k} B_{t_i}(B_{t_{i+1}} - B_{t_i}) = \frac{1}{2}\Big(\sum_{i=0}^{k}\big(B^2_{t_{i+1}} - B^2_{t_i}\big) - \big(B_{t_{i+1}} - B_{t_i}\big)^2\Big)$$
$$= \frac{1}{2}B^2_{k+1} - \frac{1}{2}\sum_{i=0}^{k}\big(B_{t_{i+1}} - B_{t_i}\big)^2.$$

It follows that

$$\int_0^t B_s\,dB_s = \frac{1}{2}B_T^2 - \frac{1}{2}\lim_{n\to\infty}^{\mathscr{L}^2} Y_k$$

where we let $Y_k = \frac{1}{2}\sum_{i=0}^{k}\big(B_{t_{i+1}} - B_{t_i}\big)^2$. Then

$$\mathbb{E}(Y_n) = \sum_{i=0}^{k}(t_{i+1} - t_i) = t_{k+1} = \frac{(k+1)T}{n},$$

so $\mathbb{E}(|Y_n - t|) \le Tk/n \to 0$ as $n \to \infty$. Moreover, since $\mathbb{E}(X^4) = 3\mathrm{Var(X)}^2$ for a normal random variable X with mean 0,

$$\mathrm{Var(Y_n)} = \sum_{\mathrm{i=0}}^{\mathrm{k}}\mathrm{Var((B_{t_{i+1}} - B_{t_i})^2)} = 2\sum_{\mathrm{i=0}}^{\mathrm{k}}\mathrm{(t_{i+1} - t_i)^2} = 2\mathrm{k}\frac{\mathrm{T}^2}{\mathrm{n}^2} \le 2\mathrm{t}\frac{\mathrm{T}}{\mathrm{n}},$$

since $t > k\frac{T}{n}$. So $\mathrm{Var(Y_n)} \to 0$ as $n \to \infty$ and $\mathbb{E}(Y_n) \to t$, so $Y_n \to t$ in $\mathscr{L}^2$.
We proved

⟦93⟧ **Our first Itô integral –** We have

$$\int_0^t B_s\,dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t.$$

7 LOCALIZATION AND ITÔ'S INTEGRAL

7.1 Itô integral for functions in $\mathscr{L}^2_{loc}$

⟦94⟧ **Problem –** We have defined the Itô integral for functions in $\mathscr{H}^2$. We could be happy with this, but the condition
(3) $\mathbb{E}\big(\int_0^T f^2(\omega,t)\,dt\big) < \infty$
for the integrand is a considerable restriction: E.g. if $f(\omega,t) = \exp(B_t^a)$, then $\mathbb{E}(f(\omega,t)^2) = \mathbb{E}(\exp(2B_t^a)) = \infty$ for $a > 2$ (can you show this?).

⟦95⟧ **Solution –** Remember from your probability course that for a random variable X it could happen that $\mathbb{E}(X) = \infty$ even though X is finite with probability one (e.g. take a uniform random variable U in $[0,1]$ and let $X = 1/U$).

The idea is to replace (3) by the weaker

(3*) $\int_0^T f^2(\omega,s),ds < \infty$ almost surely,

and call the extended class of such functions $\mathscr{L}^2_{loc}$. Clearly $\mathscr{H}^2 \subseteq \mathscr{L}^2_{loc}$.

⟦96⟧ **Definition 7.1 (Localizing sequence)** *An increasing sequence ν_n of stopping times is called a $\mathscr{H}^2$ localizing sequence for $f \in \mathscr{L}^2_{loc}$, if*

(1) $f_n(\omega,t) = f(\omega,t)\mathbb{1}_{t<\nu_n}$ is in $\mathscr{H}^2$,

(2) $\nu_n = T$ for some n a.s.

Guess why $f_n(\omega,t)$ should be a member of $\mathscr{H}^2$. The reason is of course that we want to integrate f_n and we already know how to form integrals for functions in $\mathscr{H}^2$.

⟦97⟧ **Definition 7.2 (Local martingale)** *An adapted stochastic process $\{M_t\}$ is a local martingale if there is a sequence τ_n of stopping times, such that*

(1) $\tau_n \to \infty$ as $n \to \infty$ almost surely,

(2) $M_t^{(k)} := M_{t\wedge\tau_k} - M_0$ is a martingale.

Every martingale is a local martingale, just let $\tau_n = n$.

⟦98⟧ **Summary: Construction of the Itô integral for $f \in \mathscr{L}^2_{loc}$ –** Let ν_n be a localizing sequence and define

$$X_{t,n} = \int_0^t \mathbb{1}_{s<\nu_n} f(\omega,s)\,dB_s.$$

(i.e. we chose the continuous martingale defined in the previous section). Then

(1) $X_{t,n} \to X_t$ almost surely as $n \to \infty$ for some continuous process $\{X_t\}_{t\in[0,T]}$.

(2) X_t is independent of the choice of the localizing sequence.

(3) There is a version of X_t which is a local martingale, we write $\int_0^t f(\omega,s)dB_s$ for this version. One can take the localizing sequence $\tau_n = T \wedge \inf\{s : \int_0^s f^2(\omega,t)dt \geq n\}$.

(4) Let $f,g \in \mathscr{L}^2_{loc}$ and ν be a stopping time. If $f(s) = g(s)$ for $s < \nu$ then $\int_0^t f(\omega,s)dB_s = \int_0^t g(\omega,s)dB_s$ with probability one on $\{\omega : t \leq \nu\}$.

⟦99⟧ **notation –** Of course, we write $\int_0^t f(\omega,s)dB_s$ for the local martingale version of the Itô integral for functions in $\mathscr{L}^2_{loc}$.

7.2 Two special cases

⟦100⟧ **Continuous functions of Brownian motion –** We have defined the Itô integral $\int_0^t g(\omega,s)\,dB_s$ for functions in $\mathscr{L}^2_{loc}$, i.e. functions which are measurable, adapted and $\int_0^T g(\omega,s)^2\,ds < \infty$ is a.s. finite. How can we decide measurability?

⟦101⟧ **Proposition (Measurability)** *If $f(\omega,t)$ is adapted and $t \mapsto f(\omega,t)$ is a.s. right- (or left-) continuous in t then f is measurable.*

For example if we let $g(\omega,t) = f(B_t(\omega))$ where B_t is a Brownian motion and f is continuous, then clearly g is adapted and continuous in t and hence measurable. It follows that $g \in \mathscr{L}^2_{loc}$. Since B_t is continuous a.s., $f(B_t)^2$ is continuous and bounded on $[0,T]$, so

$$\int_0^T f(B_s)^2\,ds < \infty$$

almost surely. We may wonder how to calculate the stochastic integral of f w.r.t B_t.

⟦102⟧ **Theorem 7.1 (Continuous functions of Brownian motion)** *If $f:\mathbb{R}\to\mathbb{R}$ is continuous then (letting $t_i = iT/n$)*

$$\int_0^T f(B_s)\,dB_s = \lim_{n\to\infty}^{\mathrm{p}} \sum_{i=1}^n f(B_{t_{i-1}})\big(B_{t_i} - B_{t_{i-1}}\big),$$

where $\lim^{\mathrm{p}}$ *denotes limit in probability.*

The proof is a demonstration on how localization can be used. But before, there is one question to be answered.

Why is the limit on the right a limit in probability, not a stronger limit almost surely?

The construction of Itô's integral for $f \in \mathscr{L}^2_{loc}$ promised a almost sure limit of integrals of functions in $\mathscr{H}^2$. But here we ask for a limit of integrals of simple functions that are not in $\mathscr{H}^2$, so we lose the $\mathscr{L}^2$ convergence.

Proof. We show in an exercise that

$$\tau_M = T \wedge \min\{t : |B_t| \geq M\}$$

is a localizing sequence for $f(B_s(\omega)) \in \mathscr{L}^2_{loc}[0,T]$.

Main idea: there is a continuous function $f_M(x)$ with compact support (i.e. $\mathrm{supp}(f) = \mathrm{closure}\{x : f(x) \neq 0\}$ is bounded, in other words: $f_M(x)$ vanishes for large $|X|$) and

$$f(x) = f_M(x), \quad |x| \leq M.$$

Then $f_M \in \mathscr{H}^2$ and we can construct the Itô integral of $f_M(B_s)$ by using the simple functions

$$\phi_n(\omega,s) = \sum_{i=1}^n f_M(B_{t_{i-1}})\mathbb{1}_{t_{i-1}<s\leq t_i}.$$

One can show that ϕ_n is an approximating sequence to f_M. Then

$$\int_0^t f_M(B_s)\,dB_s \quad = \quad \overset{\mathscr{L}^2}{\lim_{n\to\infty}} \int_0^t \phi_n(\omega,s)\,dB_s = \overset{\mathscr{L}^2}{\lim_{n\to\infty}} \sum_{i=1}^n f_M(B_{t_{i-1}})\big(B_{t_i}-B_{t_{i-1}}\big).$$

Let

$$A_n(\varepsilon) = \left\{\omega : \Big|\sum_{i=1}^n f(B_{t_{i-1}})\big(B_{t_i}-B_{t_{i-1}}\big) - \int_0^T f(B_s)\,dB_s\Big|\right\}.$$

Then

$$\mathbb{P}(A_n(\varepsilon)) \le \mathbb{P}(\tau_M < T) + \mathbb{P}(A_n(\varepsilon), \tau_M = T).$$

The first term tends to zero as $M \to \infty$. On $\{\omega : \tau_M = T\}$ we have a.s. that $\int_0^t f(B_s)\,dB_s = \int_0^t f_M(B_s)\,dB_s$ and

$$\sum_{i=1}^n f_M(B_{t_{i-1}})\big(B_{t_i}-B_{t_{i-1}}\big) = \sum_{i=1}^n f(B_{t_{i-1}})\big(B_{t_i}-B_{t_{i-1}}\big).$$

Hence by Chebyshev's inequality,

$$\mathbb{P}(A_n(\varepsilon), \tau_M = T) \le \frac{1}{\varepsilon^2}\mathbb{E}\Big(\big(\sum_{i=1}^n f_M(B_{t_{i-1}})\big(B_{t_i}-B_{t_{i-1}}\big) - \int_0^T f_M(B_s)\,dB_s\big)^2\Big),$$

which tends to zero.

□

⟦103⟧ **Non-random integrands –** What happens if the integrand $f(\omega,t)$ is non-random, i.e. $f(\omega,t) = f(t)$, independent of ω?

⟦104⟧ **Proposition 7.6 (Gaussian integrals)** *If $f:[0,T]\to\mathbb{R}$ is continuous then*

$$X_t = \int_0^t f(s)\,dB_s$$

is a mean zero Gaussian process with independent increments and $\mathrm{Cov}(X_t, X_s) = \int_0^{t\wedge s} f^2(u)\,du$. *Moreover*

$$X_t = \overset{p}{\lim_{n\to\infty}} \sum_{i=1}^n f(t_i^*)\big(B_{t_i}-B_{t_{i-1}}\big)$$

if $t_i = iT/n$ and $t_i^* \in [t_{i-1}, t_i]$.

We will encounter a nice application in Exercise 25 (see also Exercise 24).

7.4 Local martingales and honest ones

Recall Definition 7.2 of a local martingale.

An adapted stochastic process $\{M_t\}$ is a *local martingale* if there is a sequence τ_n of stopping times, such that

(1) $\tau_n \to \infty$ as $n \to \infty$ almost surely,

(2) $M_t^{(k)} := M_{t\wedge\tau_k} - M_0$ is a martingale.

We have seen that the Itô integral of a function in $\mathscr{L}^2_{loc}$ is a local martingale.

⟦105⟧ **Local martingales and martingales –** Let M_t be a local martingale.

(1) If τ is a stopping time, then $X_{\tau\wedge t}$ is a local martingale.

(2) If M_t is continuous and $|M_s| < B < \infty$ for all $t \geq 0$ then M_t is a martingale.

(3) If $M_t \geq 0$ and $\mathbb{E}(|X_0|) < \infty$, then M_t is a submartingale. If additionally $\mathbb{E}(X_T) = \mathbb{E}(X_0)$ then $\{M_t\}_{t\in[0,T]}$ is a martingale.

⟦106⟧ **Proposition 7.8 (Application of local martingales)** *If X_t is a continuous local martingale with $X_0 = 0$ and if $\tau < \infty$ a.s. for the stopping time*

$$\tau = \inf\{t : X_t \in \{A, -B\}\},$$

then $\mathbb{E}(X_\tau) = 0$ and $\mathbb{P}(X_\tau = A) = \frac{B}{A+B}$.

Proof. Let τ_n be any localizing sequence for X_t. Then

$$Y_t^{(k)} = X_{t\wedge\tau_k}$$

is a martingale for every k, so by optional stopping $Y_{t\wedge\tau}^{(k)}$ is a martingale and

$$\mathbb{E}(Y_{t\wedge\tau}^{(k)}) = 0.$$

Since τ is almost surely finite we have that $Y_{t\wedge\tau}^{(k)} \to Y_\tau^{(k)}$ as $t \to \infty$. Since $Y_{t\wedge\tau}^{(k)} \leq \max\{A, B\}$ we can apply dominated convergence and obtain

$$\mathbb{E}(Y_\tau^{(k)}) = 0.$$

But $Y_\tau^{(k)} = X_{\tau\wedge\tau_k} \to X_\tau$ as $k \to \infty$ (since $\tau_k \to \infty$). Again $Y_\tau^{(k)} \leq \max\{A, B\}$ and using dominated convergence we get

$$\mathbb{E}(X_\tau) = 0 = A\mathbb{P}(X_\tau = A) - B\mathbb{P}(X_\tau = B)$$

from which $\mathbb{P}(X_\tau = A) = \frac{B}{A+B}$ follows. □

Summary and Outlook

(1) We started with the desire to give a SDE

$$X_t = X_0 + \int_0^t \mu(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dB_s$$

some meaning. The troublemaker here is the stochastic integral term on the r.h.s.

(2) We defined stochastic integrals w.r.t Brownian motion for suitable integrands. Suitable integrands are such that $f(\omega, t)$ is left- or right-continuous and adapted to the standard filtration and either

- $\mathbb{E}(\int_0^T f(\omega,s)^2\, ds) < \infty$ (then $f \in \mathscr{H}^2$) and the stochastic integral is a continuous martingale, defined as a $\mathscr{L}^2$ limit of integrals of simple functions,

or

- $\int_0^T f(\omega,s)^2\, ds < \infty$ a.s. (then $f \in \mathscr{L}^2_{loc}$) and the stochastic integral is a continuous local martingale, defined as a almost sure limit of integrals functions in $\mathscr{H}^2$,

(3) We still can't evaluate stochastic integrals without first finding appropriate approximations and going through the whole procedure of proving convergence.

In this way we found our first integral

$$\int_0^t B_s\, dB_s = \frac{1}{2} B_t^2 - \frac{1}{2} t.$$

This is contrary to the situation where A_t is of bounded variation, where

$$\int_0^t A_s\, dA_s = \frac{1}{2} A_t^2.$$

It would be nice to have a formula similar to the fundamental theorem of calculus:

$$\int_0^t f'(A_s)\, dA_s = f(A_t) - f(A_0) \tag{7.1}$$

(if f is continuously differentiable). Indeed, such a formula exists and is – we are not surprised at all – different from (7.1). This formula will be revealed in the second part of the course...

A MORE EXPLANATIONS AND EXAMPLES

☞ $(\mathbf{1})_{p.6}$ Nonmeasurable sets

p.6

Look at the following terrifying construction. For each $x \in \mathbb{R}$ let

$$E_x = \{y \in \mathbb{R} : x - y \in \mathbb{Q}\},$$

so that $y \in E_x$ iff $x - y$ is a rational number (e.g. $\sqrt{2} + 17 \in E_{\sqrt{2}}$). Then $\mathbb{R} = \bigcup_x E_x$ and $E_x \cap E_y = \emptyset$ for $x \neq y$. Now we choose from each set E_x exactly one element in $[0,1]$ and collect these elements in a new set $E \subset [0,1]$ (we need the so called axiom of choice here, a legitimation that we can indeed chose one element from each of the uncountably many E_x). We have

$$\bigcup_{y \in \mathbb{Q}} (y + E) = \mathbb{R}$$

where $(y + E) = \{z : z = y + u, u \in E\}$. If E would be Lebesgue-measurable then

$$\infty = \lambda(\mathbb{R}) = \lambda\Big(\bigcup_{y \in \mathbb{Q}} (y + E)\Big) = \sum_{y \in \mathbb{Q}} \lambda(y + E).$$

But $\lambda(y + E) = \lambda(E)$ (λ is translation invariant: the mass of A is equal to the mass of the shifted set $A + a$), so $\lambda(E) > 0$ (otherwise the above sum would be zero).

On the other hand we also have

$$\bigcup_{q \in [0,1] \cap \mathbb{Q}} (q + E) \subset [0,2)$$

and hence

$$\sum_{q \in [0,1] \cap \mathbb{Q}} \lambda(q + E) \leq 2 < \infty,$$

and it follows that $\lambda(E) = 0$. This is a contradiction and hence E is not Lebesgue-measurable! Since any Borel set (a set from $\mathscr{B}(\mathbb{R})$) is Lebesgue-measurable, E is also an example of a set that is not Borel-measurable.

Btw. there are Borel sets $B \in \mathscr{B}(\mathbb{R}^2)$, such that the projection

$$B_1 = \{x \in \mathbb{R} : (x, y) \in B, \text{some y} \in \mathbb{R}\}$$

is not Borel measurable (not in $\mathscr{B}(\mathbb{R})$). Some nice members in $\mathscr{B}(\mathbb{R}^2)$ have very ugly shadows...

☞ $(\mathbf{2})_{p.9}$ Riemann integral

Indeed, if $f(x) = \mathbb{1}_{\mathbb{Q}}(x)$, then $f(x)$ is one if x is a rational number and $f(x) = 0$ elsewhere. It follows that the upper sums in the definition of the Riemann integral are all 1, while the lower sums are all 0.It follows that the Riemann integral cannot exist. There is no problem with the Lebesgue integral, since $\mathbb{Q}$ is a measurable set and f is a simple function and thus integrable (yielding the value zero).

☞ (3)$_{p.9}$ Two transformation rules

Here are two more formulas, that show how to transform integrals.

⟦**107**⟧ **Theorem (Rule 1)** *Let* $(\omega, \mathscr{F}, \mu)$ *be a measure space and let* $(\omega', \mathscr{F}')$ *be measurable space. Let* $f : \omega \to \omega'$ *and* $g : \omega' \to \mathbb{R}$ *be measurable. Then (if one of the two sides exists)*

$$\int g(f(x))\, d\mu(x) = \int g(y)\, d\nu(y),$$

where $\nu(B) = \mu(f^{-1}(B))$ *for* $b \in \mathscr{F}'$.

⟦**108**⟧ **Theorem (Rule 2)** *Let* $(\omega, \mathscr{F}, \mu)$ *be a measure space. For any measurable functions* $f : \omega \to \mathbb{R}^+$ *and* $g : \omega \to \mathbb{R}$

$$\int f(x) g(x)\, d\mu(x) = \int g(y)\, d\nu(y)$$

where $\nu(A) = \int_A f(x)\, d\mu(x)$.

☞ (4)$_{p.13}$ The Borel-Cantelli Lemma

We really need independence in (2). Let X_n be an independent sequence of random variables, such that $\mathbb{P}(X_n = n) = 1/n$ and $\mathbb{P}(X_n = 0) = 1 - 1/n$ (see exercises). We showed that $X_n \to 0$ in probability and X_n does not converge in $\mathscr{L}^1$. Also, by using Borel-Cantelli, we showed that $X_n = n$ infenitely often with probability one and hence X_n does not converge to zero almost surely.

But if we define on $\omega = [0,1]$ the random sequence $Y_n(\omega) = n\mathbb{1}_{\omega < 1/n}$ then $\mathbb{P}(Y_n = n) = 1/n$ and $\mathbb{P}(Y_n = 0) = 1 - 1/n$ and $X_n \to 0$ almost surely. The reason: the Y_n are not independent, we can not apply Borel-Cantelli here.

☞ (5)$_{p.27}$ Conditional expectation

Here are two classic proofs for the existence of the conditional expectation $\mathbb{E}(X|\mathscr{G})$.

(1) First suppose that $X \in \mathscr{L}^2$, then one can use the projection theorem for $\mathscr{L}^2$ to see that the projection X^π onto the subspace $\mathscr{L}^1(\mathscr{G})$ of $\mathscr{G}$-measurable integrable functions has the required properties of a conditional expectation. For general $X \in \mathscr{L}^1$ one uses an approximation procedure to extend the definition of the conditional expectation.

A very interesting property is the following: the conditional expectation $W = \mathbb{E}(X|\mathscr{G})$ has the smallest expectation $\mathbb{E}((X - W)^2)$ among the random variables that are $\mathscr{G}$-measurable.

(2) More standard is the approach via the Radon-Nikodym theorem, that roughly says that if for two measures $\mu(A) = 0 \Rightarrow \nu(A) = 0$ holds, then $\nu(A) = \int_A f\, d\mu$, for some measurable function f (the Radon-Nikodym-derivative).

Let $\mu(A) = \mathbb{E}(X\mathbb{1}_A)$ and denote ν the restriction of μ to $\mathscr{G}$. $\mathbb{P}(A) = 0$ implies $\nu(A) = 0$ for $A \in \mathscr{G}$ and hence by the Radon-Nikodym theorem there is a $\mathscr{G}$-measurable function Y, such that $\mathbb{E}(X\mathbb{1}_A) = \nu(A) = \int_A Y\,d\mathbb{P} = \mathbb{E}(Y\mathbb{1}_A), \forall A \in \mathscr{G}$.

☞ (6)$_{p.27}$ Uniform integrability I

Note that $\mathbb{E}(|Z|\mathbb{1}_{|Z|>x}) \to 0$ as $x \to \infty$ for any $\mathscr{L}^1$-random variable Z. This follows from dominated convergence, since $\mathbb{E}(|Z|\mathbb{1}_{|Z|>x}) \le \mathbb{E}(|Z|) < \infty$ and $|Z|\mathbb{1}_{|Z|>x} \to 0$ almost surely. The difference with uniform integrability is that zero is a uniform limit for all members in $\mathscr{C}$.

☞ (7)$_{p.27}$ Uniform integrability II

The following very interesting Lemma can be used in the proof of the following Lemma 4.6. Although we don't prove 4.6, it is interesting to note that any $\mathscr{L}^1$ random variable is also in a $\mathscr{L}^p$-like space for some $p > 1$, in the sense that $\mathbb{E}(\phi(|Z|)) < \infty$ for some function $\phi(x)$ that increases faster than x.

⟦109⟧ **Lemma 4.5** *If $\mathbb{E}(|Z|) < \infty$ then there is a convex ϕ, such that $\phi(x)/x \to \infty$ as $x \to \infty$ and*

$$\mathbb{E}(\phi(|Z|)) < \infty.$$

⟦110⟧ **Lemma 4.6** *If Z is a $\mathscr{F}$-measurable random variable with $\mathbb{E}(|Z|) < \infty$ then the collection*

$$\{\mathbb{E}(Z|\mathscr{G}) : \mathscr{G} \text{ sub } \sigma\text{-field of } \mathscr{F}\}$$

is uniformly integrable.

☞ (8)$_{p.34}$ Approximation of functions in $\mathscr{H}^2$ (receipe)

⟦111⟧ **Theorem 6.5 (Approximation)** *Let the operator A_n act on functions $f \in \mathscr{H}^2$ in the following way:*

$$A_n(f) = \sum_{i=1}^{2^n-1} \left(\frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} f(\omega, ds)\,ds \right) \mathbb{1}_{t_i < t \le t_{i+1}},$$

where $t_i = iT/2^n$ for $0 \le i \le 2^n - 1$. Then

(1) A_n is a bounded linear operator from $\mathscr{H}^2$ to $\mathscr{H}^2_0$,

(2) Contraction properties: $||A_n f||_\infty \le ||f||_\infty$ and $||A_n f||_{\mathscr{L}^2(d\mathbb{P}\times t)} \le ||f||_{\mathscr{L}^2(d\mathbb{P}\times t)}$,

(3) $||A_n f - f||_{\mathscr{L}^2(d\mathbb{P}\times t)} \to 0$ as $n \to \infty$.

Then $\int_0^T f(s)\,dB_s = \lim_{n\to\infty} \sum_{i=1}^{2^n-1} \left(\frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} f(\omega, ds)\,ds \right) \left(B_{t_{i+1}} - B_{t_i}\right)$

B EXERCISES

Measure theory and probability

Exercise 1: Suppose that $\mathscr{F}_1$ and $\mathscr{F}_2$ are two σ-fields in ω with $\mathscr{F}_1 \subset \mathscr{F}_2$ (i.e. $A \in \mathscr{F}_1 \Rightarrow A \in \mathscr{F}_2$, $\mathscr{F}_1$ is a subfield of $\mathscr{F}_2$). Let $f : \omega \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ be some function.

(1) Is it true that if f is measurable w.r.t. $\mathscr{F}_1$ then f is also measurable w.r.t. $\mathscr{F}_2$?

(2) Is the converse true?

Exercise 2: Let $\omega = [0,1]$, $\mathscr{F}_2 = \sigma(\{[0,\frac{1}{2}],[0,\frac{1}{4}]\})$.

(1) Find all elements of $\mathscr{F}_2$.

(2) Give an example of a σ-field $\mathscr{F}_1$ with $\mathscr{F}_1 \subset \mathscr{F}_2$, $\mathscr{F}_1 \neq \mathscr{F}_2$ and a function $f : [0,1] \to \mathbb{R}$ that is measurable only w.r.t. one of the two σ-fields.

(3) Find a $g \neq f$ such that g is $\mathscr{F}_1$-measurable and $\int_A f(x)\,dx = \int_A g(x)\,dx$, $\forall A \in \mathscr{F}_1$.

Exercise 3: Show that the following functions $f : (\mathbb{R}, \mathscr{B}(\mathbb{R})) \to \mathbb{R}$ are measurable and decide which are Lebesgue integrable (w.r.t to the Lebesgue measure λ). Which of the integrals $\int_{-\infty}^{\infty} f(x)\,dx$ exist in the Riemann sense?

$f(x) = \mathbb{1}_{[-1,1]}(x)$, $f(x) = \mathbb{1}_{[-3,\infty)}(x) \times e^{-x}$, $f(x) = 2x * e^{-x^2}$, $f(x) = 1/(1+x^2)$,
$f(x) = x/(1+x^2)$, $f(x) = \mathbb{1}_{(1,\infty)}(x) \times 1/x$, $f(x) = \sin(x)/x$, $f(x) = \mathbb{1}_{\mathbb{Q}}(x)$.

Exercise 4: Find the expectation $\mathbb{E}(X)$ for the following dicrete distributions of X:

(1) $\mathbb{P}(X = k) = p(1-p)^{k-1}$, $k = 1,2,\ldots$, (geometric distribution).

(2) $\mathbb{P}(X = k) = 1/n$, $k = 1,2,\ldots,n$, (discrete uniform distribution).

(3) $\mathbb{P}(X = k) = e^{-\lambda}\lambda^k/k!$, $k = 0,1,2,\ldots$, (Poisson distribution).

Find $\mathbb{E}(X)$ for the random variable X, with the following density functions:

(4) $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$, density of the (exponential distribution).

(5) $f(x) = 1/(b-a)$, $x \in [a,b]$, density of the (uniform distribution on $[a,b]$).

(6) $f(x) = e^{-(x-\mu)^2/(2\sigma^2)}/\sqrt{2\pi\sigma^2}$, density of the (normal distribution).

Exercise 5: Let X_n be an independent sequence of random variables, such that $\mathbb{P}(X_n = n) = 1/n$ and $\mathbb{P}(X_n = 0) = 1 - 1/n$.

(1) Show that $X_n \to 0$ in probability.

(2) What is the probability that $X_n = n$ infinitely often?

(3) Does the sequence converge almost surely?

(4) What about convergence in distribution?

(5) Find the limit $\lim_{n\to\infty} \mathbb{E}(X_n)$.

Exercise 6: Let X,Y be random variables with $X,Y \in \mathscr{L}^2$ ($\mathbb{E}(|X|^2) < \infty$ and $\mathbb{E}(|Y|^2) < \infty$).

(1) Is it true that $X+Y \in \mathscr{L}^2$?

(2) Is XY in $\mathscr{L}^2$? For which p is it in $\mathscr{L}^p$?

(3) Let $X = \mathbb{1}_A$ and $Y = \mathbb{1}_B$ and show that $\mathbb{P}(A\cap B)^2 \le \mathbb{P}(A)\mathbb{P}(B)$.

Exercise 7: Show using Fubini's theorem: if $X \ge 0$ a.s. and $\mathbb{E}(X) < \infty$ then

$$\mathbb{E}(X) = \int_0^\infty \mathbb{P}(X > u)\,du.$$

Can you find a similar formula for the higher moments $\mathbb{E}(X^n)$, $n \ge 2$?

Exercise 8: Let X and Y be independent random variables with distribution functions $F(x) = \mathbb{P}(X \le x)$ and $G(x) = \mathbb{P}(X \le x)$. Find the distribution function of the sum $Z = X+Y$. Hint: use the fact that $\mathbb{P}(A) = \mathbb{E}(\mathbb{1}_A)$.

Exercise 9: Let X_n be a sequence of random variables and $C \in \mathbb{R}$ a constant. Show that if $X_n \to C$ in distribution, then $X_n \to C$ also in probability.

Discrete Time Martingales

Exercise 10: Consider the coin tossing example and let $S = \{k \in \mathbb{Z} : \mathbb{P}(M_n = k \text{ infinitely often}) > 0\}$. Show that $S = \emptyset$ or $S = \mathbb{Z}$ and conclude that $\mathbb{P}(\tau < \infty) = 1$, where τ is the first time the process reaches the set A or $-B$.

Exercise 11: Given a martingale $\{M_n\}_{n\in\mathbb{N}}$. Show that $\mathbb{E}(M_n) = \mathbb{E}(M_1)$ for all $n \ge 1$.

Exercise 12: Let $\{X_k\}_{k\in\mathbb{N}}$ be i.i.d.

(1) Suppose that $\mathbb{E}(|X_1|) < \infty$. Show that $M_n = X_1 + \ldots + X_n - n\mathbb{E}(X_1)$ is a martingale (w.r.t. $\{X_k\}_{k\in\mathbb{N}}$).

(2) Suppose that $X_1, X_2, \ldots$ are non-negative with $\mathbb{E}(X_1) = 1$. Show that $M_n = X_1 X_2 \cdots X_n$ is a martingale (w.r.t. $\{X_k\}_{k\in\mathbb{N}}$).

Exercise 13: Let $\{M_n\}$ be a non-negative martingale with $M_1 = 0$. Show that $M_n = 0$ for all n a.s.

Exercise 14: Show: if $\phi : \mathbb{R} \to \mathbb{R}$ is a convex function and M_n is a martingale w.r.t. $X_1, X_2, \ldots$ then $\phi(M_n)$ is a submartingale w.r.t. $X_1, X_2, \ldots$, provided that $\mathbb{E}(|\phi(M_n)|) < \infty$. Use Jensen's inequality $\mathbb{E}(\phi(X)|\mathscr{F}) \ge \phi\big(\mathbb{E}(X|\mathscr{F})\big)$.

Exercise 15: If $\{M_n\}$ is a martingale w.r.t $X_1, X_2, \ldots$, show that $M'_k = (M_{k+m} - M_m)^2$ is a submartingale for all m w.r.t. $\mathscr{F}'_k = X_1, \ldots, X_{m+k}$.

Exercise 16: Let $A_1, A_2, \ldots$ be a sequence of independent events and let $S_n = \sum_{i=1}^n \mathbb{1}_{A_i}$ and $\tau_k = \min\{n : S_n = k\}, \quad k \geq 1$. Suppose that $a_n = \mathbb{E}(S_n) \to \infty$ as $n \to \infty$.

(1) Show that $\mathbb{P}(\tau_k < \infty) = 1$.

(2) Use optional stopping to show that $\mathbb{E}(S_{\tau_k}) = k$.

Brownian Motion

Exercise 17: We defined the functions β_n during the lecture and showed that $X_t = \sum_{n=0}^\infty \beta_n(t) Z_n$ defines a Brownian motion on $[0,1]$. Draw (as good as you can) a typical sample path of the process $Y_t = \sum_{n=1}^\infty \beta_n(t) Z_n$ for $t \in [0,1]$.

(1) What is the distribution of Y_t?

(2) Calculate the covariance $\mathrm{Cov(Y_t, Y_s,)}$.

(3) Show that $Y_t = B_t - tB_1$.

(4) Make a figure showing a typical sample path of $W_t = \sum_{n=4}^\infty \beta_n(t) Z_n$ for $t \in [0,1]$.

Continuous time martingales

Exercise 18: Let $\mathscr{H} \subseteq \mathscr{G}$ be sub-σ-fields of $\mathscr{F}$. Prove that:

(1) $\mathbb{E}(\mathbb{E}(X|\mathscr{G})|\mathscr{H}) = \mathbb{E}(X|\mathscr{H})$.
Hint: Let $Z = \mathbb{E}(X|\mathscr{G})$. Then we have to show that $U = \mathbb{E}(Z|\mathscr{H})$ is a version of the conditional expectation $\mathbb{E}(X|\mathscr{H})$, i.e.

  (i) U is $\mathscr{H}$-measurable,
  (ii) $\mathbb{E}(U \mathbb{1}_A) = \mathbb{E}(X \mathbb{1}_A)$ for all $A \in \mathscr{H}$.

(2) $\mathbb{E}(X|\mathscr{G}) = X$ and $\mathbb{E}(XZ|\mathscr{G}) = X\mathbb{E}(Z|\mathscr{G})$ if X is $\mathscr{G}$-measurable. Argue as in the above exercise.

Exercise 19: Show that the following processes are martingales w.r.t. the standard Brownian filtration:

(1) B_t,

(2) $B_t^2 - t$,

(3) $\exp\left(\alpha B_t - \frac{\alpha^2}{2} t\right)$.

Exercise 20: Let $-B < 0 < A$ and define the random variable $\tau = \inf\{t : B_t \in \{-B, A\}\}$. Obviously τ represents the first time the Brownian motion leaves the open interval $(-B, A)$.

(1) Show that τ is a stopping time w.r.t. standard Brownian filtration.

(2) Show that $\mathbb{P}(\tau < \infty) = 1$ (Hint: show $\mathbb{P}(\tau > n+1) \le \mathbb{P}(|B_{i+1} - B_i| < A+B, \forall i = 1,2,\ldots,n)$).

(3) Employ the optional stopping theorem to show that $\mathbb{E}(B_\tau) = 0$. Use this to deduce that

$$\mathbb{P}(B_\tau = A) = \frac{B}{A+B}, \quad \mathbb{P}(B_\tau = B) = \frac{A}{A+B}.$$

(Compare with equation (2.1) on page 18).

(4) Use the martingale $B_t^2 - t$ from Exercise 19 to show that $\mathbb{E}(B_\tau^2) = \mathbb{E}(\tau)$ and $\mathbb{E}(\tau) = AB$.

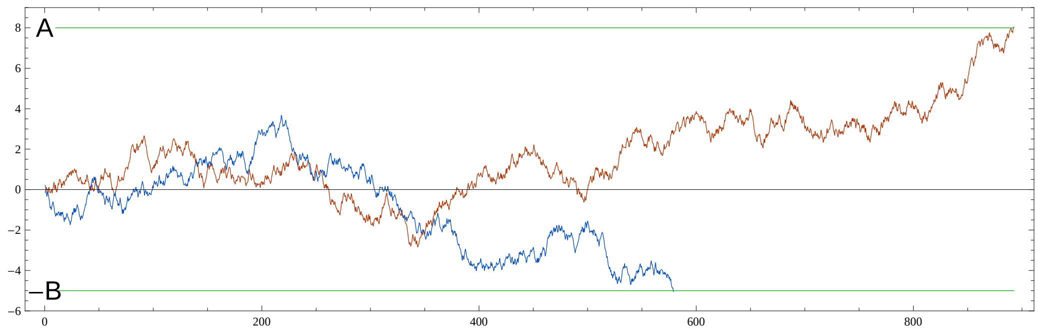


Figure 4: Two samples of B_t for $0 \le t \le \tau$ ($A = 8$, $B = 5$)

(5) Let $\tau_A = \inf\{t : B_t = A\}$. Show that $\mathbb{P}(\tau_A < \infty) = 1$ by using the results from (3).

(6) Show that $M_{\tau_A} = 1$ where $M_t = \exp\left(\alpha B_t - \frac{\alpha^2}{2} t\right)$ is the martingale from Exercise 19(3).

(7) Calculate the Laplace transform $\mathbb{E}(e^{-s\tau_A}) = e^{-A\sqrt{2s}}$, $s \ge 0$ of the hitting time τ_A?

(8) Is $\mathbb{E}(\tau_A)$ finite?

Itô integral

Exercise 21: Calculate the mean and variance of the (non-stochastic) integrals.

(1) $\int_0^t B_s\,ds$, Hint: to calculate the variance write $\mathbb{E}\Big(\big(\int_0^t B_s\,ds\big)^2\Big) = \mathbb{E}\Big(\int_0^t \int_0^t B_s B_w\,dw\,ds\Big)$.

(2) $\int_0^t B_s^2\,ds$, Hint: use the same trick as in (1). Apply the martingale $B_t^2 - t$ to calculate the expectation $\mathbb{E}(B_s^2 B_w^2) = \mathbb{E}(\mathbb{E}(B_s^2 B_w^2 | \mathscr{F}_w))$.

Exercise 22: Calculate the variance of the following stochastic integrals.

(1) $\int_0^t \sqrt{|B_s|}\,dBs$ (you can use that for a normal random variable Z with mean 0 and variance s we have $\mathbb{E}(|Z|) = \sqrt{\frac{2}{\pi}\sqrt{s}}$),

(2) $\int_0^t (B_s + s)^2\,dBs$.

Hint: use the Ito isometry: $\mathbb{E}\big(\big(\int_0^t f(\omega,s)\,dB_s\big)^2\big) = \mathbb{E}\big(\int_0^t f^2(\omega,s))\,ds\big)$.

Exercise 23: This exercises is needed in the proof of Theorem ⟦102⟧. Show that

$$\nu_n = T \wedge \min\{t : |B_t| \ge n\}$$

is a localizing sequence for $f(B_s(\omega)) \in \mathcal{L}^2_{loc}[0,T]$. Here f is a continuous function.

Recall the definition: An increasing sequence ν_n of stopping times is a localizing sequence for $f \in \mathcal{L}^2_{loc}[0,T]$, if

(1) $f_n(\omega,t) = f(\omega,t)\mathbb{1}_{t<\nu_n}$ is in $\mathcal{H}^2$,

(2) $\nu_n = T$ for some n a.s.

Exercise 24: Construct a Gaussian process X_t on $[0,1]$ with independent increments and variance

$$\mathrm{Var(X_t) = t/(1+t)}.$$

Recall ⟦104⟧: If $f : [0,T] \to \mathbb{R}$ is continuous then $Z_t = \int_0^t f(s)\,dB_s$ is a mean zero Gaussian process with independent increments and $\mathrm{Cov(Z_t, Z_s)} = \int_0^{\mathrm{t}\wedge\mathrm{s}} \mathrm{f^2(u)\,du}$.

Exercise 25: Let $f : [0,\infty) \to (0,\infty)$ be continuous and such that $\int_0^t f^2(s)\,ds \to \infty$ as $t \to \infty$. Let $\xi(t)$ be a solution of the equation $\int_0^{\xi(t)} f^2(s)\,ds = t$. Show that $X_t = \int_0^{\xi(t)} f(s)\,dB_s$ is a standard Brownian motion.

C SOLUTIONS

Solution to 1: (1) If f is measurable w.r.t. $\mathcal{F}_1$ then $f-1(B) = \{x \in \omega : f(x) \in B\}$ is a element of $\mathcal{F}_1$. Since $\mathcal{F}_1 \subseteq \mathcal{F}_2$ it follows that $f^{-1}(B)$ is also in $\mathcal{F}_2$, which makes f also measurable w.r.t. $\mathcal{F}_2$.

(2) We guess, that this is not the case. Let $\omega = [0,1]$, $\mathcal{F}_2 = \sigma(\{[0,\frac{1}{2}],[0,\frac{1}{4}]\})$ and $\mathcal{F}_1 = \sigma(\{[0,\frac{1}{2}]\})$, i.e.

$$\mathcal{F}_1 = \{\emptyset, [0,\tfrac{1}{2}], (0,\tfrac{1}{2}], [0,1]\}$$

and

$$\mathcal{F}_2 = \{\emptyset, [0,\tfrac{1}{2}], (\tfrac{1}{2},1], [0,1], [0,\tfrac{1}{4}], (\tfrac{1}{4},1], (\tfrac{1}{4},\tfrac{1}{2}], [0,\tfrac{1}{4}] \cup (\tfrac{1}{2},1]\}$$

(the smallest sigma fields that contain $\{[0,\frac{1}{2}]\}$ and $\{[0,\frac{1}{2}],[0,\frac{1}{4}]\}$ respectively, see figure below).

Now let $f(x) = \mathbb{1}_{[0,\frac{1}{4}]}(x)$. Then $f^{-1}(\{1\}) = [0,\frac{1}{4}]) \notin \mathcal{F}_1$, so f is not measurable w.r.t. $\mathcal{F}_1$.

Solution to 2: (1),(2) See exercise 1.

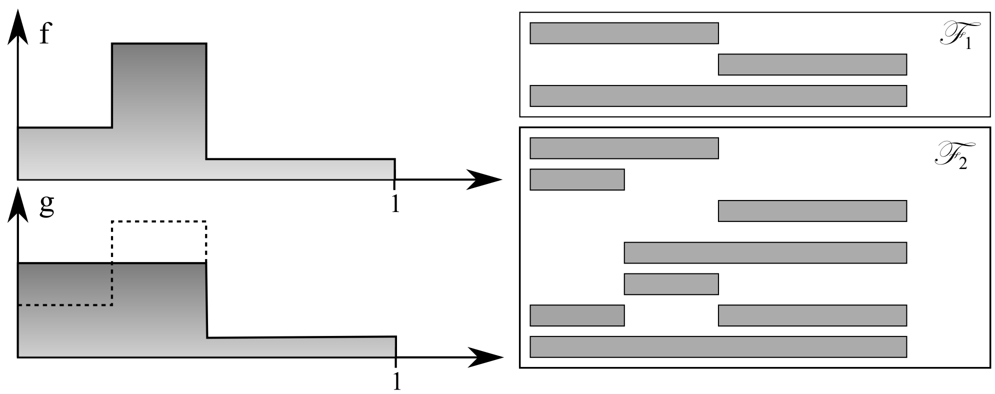
(3) Let

$$g(x) = \frac{1}{2}\mathbb{1}_{[0,\frac{1}{2}]}(x).$$

Then g is certainly $\mathscr{F}_1$ measurable and

$$\int_{[0,\frac{1}{2}]} f(x)\,dx = \frac{1}{4} = \int_{[0,\frac{1}{2}]} g(x)\,dx.$$

The following figure shows another (different) example of a function f that is to 'fine' to be measurable w.r.t. $\mathscr{F}_1$. On the interval $[0,\frac{1}{2}]$ the function g is equal to the average of f. We will later see, that this kind of averaging property is useful to define conditional expectation.



Solution to 3: The functions are all measurable.

function	Leb. integrable?	Riemann integrable?	remark
$\mathbb{1}_{[-1,1]}(x)$	✓	✓	simple function
$\mathbb{1}_{[-3,\infty)}(x) \times e^{-x}$	✓	✓	–
$2x * e^{-x^2}$	✓	✓	–
$1/(1+x^2)$	✓	✓	–
$x/(1+x^2)$	✗	✗	$\approx 1/x$ for large x
$\mathbb{1}_{(1,\infty)}(x) \times 1/x$	✗	✗	$= 1/x$ for large x
$\sin(x)/x$	✗	✓ (improper)	$\int f+ = \infty$, $\int f^- = \infty$
$f(x) = \mathbb{1}_{\mathbb{Q}}(x)$	✓	✗	$\mathbb{Q}$ is countable $\Rightarrow$ measurable

Solution to 5: (1) W.l.o.g. we assume that $\varepsilon < 1$. Then $\mathbb{P}(|X_n| > \varepsilon) = \mathbb{P}(X_n > \varepsilon) = 1/n$, which tends to zero as $n \to \infty$. Hence $X_n \to 0$ in probability.

(2) Using Borel-Cantelli it follows that $\mathbb{P}(X_n = n \text{ i.o.}) = 1$, since

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n = n) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

(3) No, since it follows from (2) that the probability that $X_n = n$ infinitely often is one, so X_n does not converge with probability one.

(4) It follows from (1) that $X_n \to 0$ in distribution. You can also argue as follows: For $x \neq 0$ we have

$$\mathbb{P}(X_n \le x) \to \mathbb{1}_{x \le 0}.$$

(5) Since $\mathbb{E}(X_n) = 1/n \times \mathbb{P}(X_n = 1/n) + 0 \times \mathbb{P}(X_n = 0)$ we have $\mathbb{E}(X_n) \to 1$ as $n \to \infty$. The sequence X_n does not converge to zero in $\mathscr{L}^1$.

Solution to 6: (1) Short answer: Sure, because $\mathscr{L}^1$ is a linear space! Or use Minkowski's inequality to see that $||x+y||_2 \le ||x||_2 + ||y||_2 < \infty$.

(2) For $p = 1$ the product XY is in $\mathscr{L}^p$ because of Cauchy-Schwarz's inequality. It is not clear whether this is true for $p = 2$ (all we know is, that $\mathscr{L}^2 \subseteq \mathscr{L}^1$).

Let $\omega = \mathbb{N} = 1,2,3,\dots$, $\mathbb{P}(\{k\}) = \frac{6}{\pi^2 k^2}$ (note that $\mathbb{P}(\omega) = 1$ because $\sum_{k=1}^{\infty} 1/k^2 = \pi^2/6$). Define the random variables $X(\omega) = Y(\omega) = \sqrt[4]{\omega}$. Then

$$\int X^2 \, d\mathbb{P} = \sum_{k=1}^{\infty} \sqrt{k} \frac{6}{\pi^2 k^2} < \infty$$

but

$$\int (XY)^2 \, d\mathbb{P} = \sum_{k=1}^{\infty} k \frac{6}{\pi^2 k^2} = \infty,$$

so $XY \not\in \in \mathscr{L}^2$.

Solution to 7: Main idea: Use Fubini's theorem and the following transformation rule:

$$\mathbb{E}(g(X)) = \int g(X(\omega))\, dP(\omega) = \int_{-\infty}^{\infty} g(x)\, dF(x).$$

This is useful, since you can calculate expectations without using the integral w.r.t. to the probability measure $\mathbb{P}$ itself, but instead using the probability distribution function $F(x) = \mathbb{P}(X \le x)$. We then obtain

$$\mathbb{P}(X > u) = \mathbb{E}(\mathbb{1}_{X>u}) = \int \mathbb{1}_{(X(\omega)>u)}(\omega)\, dP(\omega) = \int_0^{\infty} \mathbb{1}_{(u,\infty)}(x)\, dF(x),$$

where is the distribution function of X. Then

$$\int_0^{\infty} \mathbb{P}(X > u)\, du \quad = \quad \int_0^{\infty} \int_0^{\infty} \mathbb{1}_{(u,\infty)}(x)\, dF(x)\, du.$$

It follows by Fubini's theorem that we can interchange the order of integration:

$$\int_0^{\infty}\int_0^{\infty} \mathbb{1}_{(u,\infty)}(x)\, dF(x)\, du = \int_0^{\infty}\int_0^{\infty} \mathbb{1}_{(u,\infty)}(x)\, du\, dF(x) = \int_0^{\infty}\int_0^{x} du\, dF(x) = \int_0^{\infty} x\, dF(x).$$

Solution to 8: We want to determine the distribution function of the sum $Z = X + Y$ of two independent random variables X and Y. We have $\mathbb{P}(Z \le z) = \mathbb{E}(\mathbb{1}_{Z \le z})$, the right side being just an abbreviation of $\mathbb{E}(\mathbb{1}_{\{\omega::Z(\omega)\le z\}})$. Then (using again the transformation of exercise 7):

$$\begin{aligned}\mathbb{E}(\mathbb{1}_{Z\le z}) &= \int \mathbb{1}_{Z(\omega)\le z}\, dP(\omega) = \int \mathbb{1}_{X(\omega)+Y(\omega)\le z}\, dP(\omega) \\ &= \int \mathbb{1}_{X+Y\le z}\, dP = \int_{\mathbb{R}^2} \mathbb{1}_{x+y\le z}\, dH(x,y).\end{aligned}$$

Here H is the joint distribution function of X and Y. By independence we have $H(x,y) = F(x)\times G(y)$ with F,G the distribution functions of X and Y. Hence

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{1}_{x+y\le z}\, dH(x,y) &= \int_{-\infty}^{\infty} \mathbb{1}_{x\le z-y}\, dF(x)\, dG(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} dF(x)\, dG(y) \\ &= \int_{-\infty}^{\infty} F(z-y)\, dG(y).\end{aligned}$$

This is the so called convolution of F and G. Two remarks: i) of course it follows similarly that

$$\mathbb{P}(Z\le z) = \int_{-\infty}^{\infty} G(z-x)\, dF(x).$$

ii) If X and Y are nonnegative random variables, then

$$\mathbb{P}(Z\le z) = \int_0^z F(z-y)\, dG(y).$$

Solution to 9: Suppose $X_n \to C$ in distribution, i.e. $\mathbb{P}(X_n < C) \to 0$ and $\mathbb{P}(X_n > C) \to 1$ (we don't know about convergence of $\mathbb{P}(X_n = C)$, but we don't need it). Then

$$\mathbb{P}(|X_n - C| > \varepsilon) = \mathbb{P}(X_n > C+\varepsilon) + \mathbb{P}(X_n < C-\varepsilon) \to 1+0 = 1.$$

Solution to 10: Suppose $S \ne \emptyset$ and $S \ne \mathbb{Z}$. Then $S = \{a, a+1, \ldots, b\}$ with $a,b\in\mathbb{Z}$. But

$$\mathbb{P}(M_k \in \{a,b\} \text{ infinitely often}) > 0.$$

Since $\mathbb{P}(M_{k+1} = b+1 | M_k = b) = 1/2 > 0$ we have $b+1 \in S$ which is a contradiction. Hence $S=\emptyset$ or $S=\mathbb{Z}$ $\mathbb{P}(\tau<\infty)=1$ and it follows that $\mathbb{P}(\tau<\infty)=1$.

Solution to 11: $\mathbb{E}(M_n) = \mathbb{E}(\mathbb{E}(M_n|\mathscr{F}_n)) = \mathbb{E}(M_{n-1})$. It follows by induction that $\mathbb{E}(M_n)$ is constant.

Solution to 12: (1) Clearly $X_n \in \mathscr{F}_n$ and $\mathbb{E}(|M_n|) < \infty$. Then, writing as usual $S_n = X_1 + \ldots + X_n$,

$$\begin{aligned}\mathbb{E}(M_n|\mathscr{F}_{n-1}) &= \mathbb{E}(S_n|\mathscr{F}_{n-1}) - n\mathbb{E}(X_1) = \mathbb{E}(S_{n-1}|\mathscr{F}_{n-1}) + \mathbb{E}(X_n|\mathscr{F}_{n-1}) - n\mathbb{E}(X_1) \\ &= S_{n-1} + \mathbb{E}(X_n) - n\mathbb{E}(X_1) = S_{n-1} - (n-1)\mathbb{E}(X_1) = M_{n-1}.\end{aligned}$$

(2) Obviously $M_n \in \mathscr{F}_n$ and $\mathbb{E}(|M_n|) = \prod_{k=1}^n X_i < \infty$. We have

$$\mathbb{E}(M_n|\mathscr{F}_{n-1}) = \mathbb{E}(\prod_{k=1}^{n-1} X_i|\mathscr{F}_{n-1})\mathbb{E}(X_n|\mathscr{F}_{n-1}) = \prod_{k=1}^{n-1} X_i \mathbb{E}(X_n) = \prod_{k=1}^{n-1} X_i = M_{n-1}.$$

Solution to 13: We use Doob's inequality $\lambda\mathbb{P}(M_n^* \geq \lambda) \leq \mathbb{E}(M_n) = \mathbb{E}(M_1) = 0$ and hence $\mathbb{P}(M_n^* \geq \lambda) = 0$ for all $\lambda > 0$, so $M_n^* = 0$ a.s. and then $M_n = 0$ a.s.

Solution to 14: (1) $\mathbb{E}(|\phi(M_n)|) < \infty$ is given in the exercise,
(2) $\phi(M_n) = f(X_1,\dots,X_n)$ is clear,
(3) The martingale property follows from Jensen's inequality:

$$\mathbb{E}(\phi(M_n)|\mathscr{F}_{n-1}) \geq \phi\big(\mathbb{E}(M_n|\mathscr{F}_{n-1})\big) = \phi(M_{n-1}).$$

Solution to 15: First note that taking the convex function $\phi(x) = x^2$ it follows that M_{m+k} is a martingale and M^2_{m+k} is a submartingale (w.r.t. $\mathscr{F}'_k = X_1,\dots,X_{m+k}$, see exercise 13). Then we have

$$\begin{aligned}
\mathbb{E}(M'_n|\mathscr{F}'_{n-1}) &= \mathbb{E}((M_{n+m} - M_m)^2|\mathscr{F}_{m+n-1}) \\
&= \mathbb{E}(M^2_{n+m}|\mathscr{F}_{m+n-1}) + 2\mathbb{E}(M_{n+m}M_m|\mathscr{F}_{m+n-1}) + \mathbb{E}(M^2_m|\mathscr{F}_{m+n-1}) \\
&\leq M^2_{n+m-1} + 2M_mM_{n+m-1} + M^2_m = (M_{n+m-1} - M_m)^2.
\end{aligned}$$

Solution to 16: (1) Since $\mathbb{E}(S_n) = \sum_{i=1}^n \mathbb{P}(A_i)$ it follows from $\sum_{i=1}^\infty \mathbb{P}(A_i) = \infty$ and Borel-Cantelli that $\mathbb{P}(\mathbb{1}_{A_i} = 1$ infinitely often$) = 1$ and hence $\mathbb{P}(S_i$ reaches $k) = \mathbb{P}(\tau_k < \infty) = 1$.

(2) The process $M_n = S_n - a_n$ is a martingale w.r.t. $\mathscr{F}_n = \{\mathbb{1}_{A_1},\dots,\mathbb{1}_{A_n}\}$. Indeed, (1) $\mathbb{E}(|M_n|) \leq n < \infty$, (2) $M_n = f(X_1,\dots,X_n)$ and (3)

$$\mathbb{E}(M_n|\mathscr{F}_{n-1}) = \mathbb{E}(S_{n-1} + \mathbb{1}_{A_n} - a_n|\mathscr{F}_{n-1}) = S_{n-1} + \mathbb{P}(A_n) - \sum_{i=1}^n \mathbb{P}(A_i) = M_{n-1}.$$

Obviously τ_k is a stopping time w.r.t. $\mathscr{F}_n$ and by optional stopping $M_{n\wedge\tau_k}$ is a martingale and since $\mathbb{E}(M_1) = 0$, it follows that $\mathbb{E}(M_{n\wedge\tau_k}) = 0$. Clearly $|M_{n\wedge\tau_k}| \leq 2k$ and $M_{n\wedge\tau_k} \to M_{\tau_k}$ and hence by dominated convergence $\mathbb{E}(M_{\tau_k}) = 0$. Since $\mathbb{P}(\tau_k < \infty) = 1$ we have $M_{\tau_k} = k - a_{\tau_k}$, so that indeed $a_{\tau_k} = \mathbb{E}(S_{\tau_k}) = k$.

Solution to 17: (1) Y_t is a sum of normal random variables and has a normal distribution with mean 0 and variance

$$\mathrm{Var}(X_t) = \sum_{n=1}^\infty \beta_n^2(t) = \mathrm{Var}(X_t) - \beta_0^2(t) = t - t^2.$$

Actually, this is a bit cheated since we cannot just take the mean and the variance of an infinite sum to be the infinite sum of the means and variances. If you want a more rigorous proof follows along the lines of the proof for the Brownian motion representation (finite dimensional distributions via characteristic functions).

Y_t is a *Brownian bridge*, i.e. it behaves in principle like a Brownian motion, but has $Y_1 = 1$ (so the process is fixed at $t = 1$: we see that the variance has its maximal value at $t = 1/2$ and then goes back to 0).

(2) $\mathrm{Cov}(Y_t, Y_s) = \mathbb{E}(\sum_{n=1}^\infty \beta_n(t)Z_n \sum_{n=1}^\infty \beta_n(s)Z_n) = \sum_{n=1}^\infty \beta_n(t)\beta_n(s) = s \wedge t - ts$.
(3) We have $Y_t = B_t - tZ_0$ and $B_1 = \sum_{n=0}^\infty \beta_n(1)Z_n = tZ_0$.
(4) The process $W_t = \sum_{n=4}^\infty \beta_n(t)Z_n$ is a combination of Brownian bridges, such that $W_t = 0$ for $t \in \{0, 0.25, 0.5, 0.75, 1\}$.

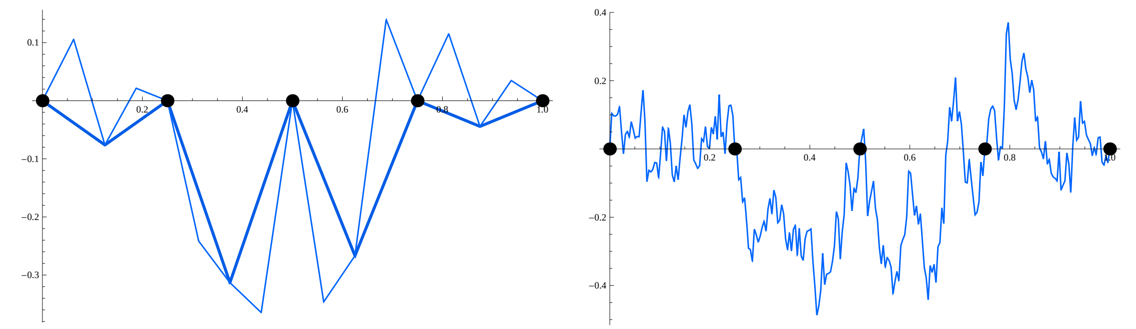


Figure 5: Exercise 17 (4): $W_{m,t} = \sum_{n=4}^{m} \beta_n(t) Z_n$ *for* $m = 7$, $m = 15$ *(left) and* $m = 255$ *(right).*

Solution to 18: (1) U is $\mathscr{H}$ measurable by definition. Also $\mathbb{E}(Z\mathbb{1}_A) = \mathbb{E}(X\mathbb{1}_A)$ for all $A \in \mathscr{G}$ and $\mathbb{E}(U\mathbb{1}_A) = \mathbb{E}(Z\mathbb{1}_A)$ for all $A \in \mathscr{H}$ by definition. Hence

$$\mathbb{E}(U\mathbb{1}_A) = \mathbb{E}(Z\mathbb{1}_A) = \mathbb{E}(X\mathbb{1}_A), \forall A \in \mathscr{H}.$$

So U is aversion of the conditional expectation $\mathbb{E}(X|\mathscr{H})$.

(2) It is clear that X is $\mathscr{G}$-measurable and that $\mathbb{E}(X\mathbb{1}_A) = \mathbb{E}(X\mathbb{1}_A)$ for all $A \in \mathscr{G}$, so X is indeed a version of the conditional expectation of X w.r.t. $\mathscr{G}$.

The second question is not so easy to prove. If $X = \mathbb{1}_B$ with $B \in \mathscr{G}$ then

$$\mathbb{E}(X\mathbb{E}(Z|\mathscr{G})\mathbb{1}_A) = \mathbb{E}(\mathbb{E}(Z|\mathscr{G})\mathbb{1}_{A\cap B}) = \mathbb{E}(Z\mathbb{1}_{A\cap B}) = \mathbb{E}(XZ\mathbb{1}_A),$$

for all $A \in \mathscr{G}$. If Z is a simple random variable (c.f. (1.1)) $Z = \sum_{i=1}^{n} c_i \mathbb{1}_{A_i}$ with $A_i \in \mathscr{G}$, then similarly (use the same reasoning as above) $\mathbb{E}(X\mathbb{E}(Z|\mathscr{G})\mathbb{1}_A) = \mathbb{E}(XZ\mathbb{1}_A)$. If X is $\mathscr{G}$-measurable and nonnegative, then there is a sequence of simple nonnegative r.v. X_n of the above form with $X_n \uparrow X$ and $\mathbb{E}(X_n\mathbb{E}(Z|\mathscr{G})\mathbb{1}_A) = \mathbb{E}(X_n Z\mathbb{1}_A)$.

We have $X_n\mathbb{E}(Z|\mathscr{G})\mathbb{1}_A \uparrow X\mathbb{E}(Z|\mathscr{G})\mathbb{1}_A$ and $X_n Z\mathbb{1}_A \uparrow XZ\mathbb{1}_A$, so by monotone convergence

$$\mathbb{E}(X\mathbb{E}(Z|\mathscr{G})\mathbb{1}_A) = \mathbb{E}(XZ\mathbb{1}_A),$$

showing that $X\mathbb{E}(Z|\mathscr{G})$ is version of the conditional expectation $\mathbb{E}(XZ\mathbb{1}_A)$ (we haven't shown measurability, bit this is clear since X and $\mathbb{E}(Z|\mathscr{G})$ are both $\mathscr{G}$-measurable).

Solution to 19: (1) This can be shown similarly to the Poisson process example ⟦68⟧.
(2) $\mathbb{E}(|B_t^2 - t|) \le \mathbb{E}(|B_t^2|) + t^2 = t + t^2 < \infty$ and for $s \le t$,

$$\begin{aligned} \mathbb{E}(B_t^2 - t|\mathscr{F}_s) &= \mathbb{E}((B_t - B_s)^2|\mathscr{F}_s) + 2\mathbb{E}(B_t B_s|\mathscr{F}_s) - \mathbb{E}(B_s^2|\mathscr{F}_s) - t \\ &= \mathbb{E}((B_t - B_s)^2) - B_s^2 - t = t - s - B_s^2 - t = B_s^2 - s. \end{aligned}$$

(3) Follows similarly if we use $\mathbb{E}(e^{\alpha B_t}) = e^{\frac{\alpha^2}{2}t}$.

Solution to 20:

(1) $\{\tau > t\} = \{B_s \in (-B, A), 0 \le s \le t\} \in \mathscr{F}_t$, so τ is a stopping time.

(2) $\mathbb{P}(\tau > n+1) \le \mathbb{P}(|B_{i+1} - B_i| < A + B, \forall i = 1, 2, \ldots, n)$ is clear. Letting $p = \mathbb{P}(|B_{i+1} - B_i| < A + B)$ we obtain $\mathbb{P}(\tau > n+1) \le p^n$. It follows that $\mathbb{P}(\tau = \infty) = \lim_{n\to\infty} \mathbb{P}(\tau > n+1) = 0$.

(3) $B_{\tau\wedge t}$ is a martingale, so $\mathbb{E}(B_{\tau\wedge t}) = \mathbb{E}(B_0) = 0$. As $t \to \infty$, $B_{\tau\wedge t} \to B_\tau$ almost surely. Moreover $B_{\tau\wedge t} \le A + B$, so by dominated convergence $\mathbb{E}(B_\tau) = 0$.

But at the same time $\mathbb{E}(B_\tau) = A\mathbb{P}(B_\tau = A) - B\mathbb{P}(B_\tau = -B)$ and $\mathbb{P}(B_\tau = A) = 1 - \mathbb{P}(B_\tau = -B)$, leading to the formulas.

(4) We have $\mathbb{E}(B^2_{t\wedge\tau}) = \mathbb{E}(t\wedge\tau)$. $t\wedge\tau \to \tau$ almost surely and $t\wedge\tau \le \tau$. So $\mathbb{E}(t\wedge\tau) \to \mathbb{E}(\tau)$. Also $B^2_{t\wedge\tau} \le (A+B)^2$ and $B^2_{t\wedge\tau} \to B^2_\tau$, yielding $\mathbb{E}(B^2_{t\wedge\tau}) \to \mathbb{E}(B^2_\tau)$ and hence

$$\mathbb{E}(\tau) = \mathbb{E}(B^2_\tau) = A^2\mathbb{P}(B_\tau = A) + B^2(1 - \mathbb{P}(B_\tau = A)) = AB.$$

(5) $\mathbb{P}(\tau_A < \infty) \ge \mathbb{P}(B_\tau = A)$, which tends to 1 as $B \to \infty$.

(6) We have $\mathbb{E}(M_{\tau_A\wedge t}) = \mathbb{E}(M_0) = 1$. Again, $M_{\tau_A\wedge t} \le e^{\alpha a}$ and $\tau_A \wedge t \to \tau_A$, hence $\mathbb{E}(M_{\tau_A}) = 1$.

(7) It follows that $\mathbb{E}(e^{\alpha a - \alpha^2/2\tau_A}) = 1$, so, letting $\alpha^2/2 = s$, $\mathbb{E}(e^{-s\tau_A}) = e^{\sqrt{2s}a}$.

(8) No, since $\mathbb{E}(\tau_A) = \frac{d}{ds}e^{\sqrt{2s}a}\Big|_{s=0} = \infty$.

Solution to 21: (1) Clearly $\mathbb{E}\Big(\int_0^t B_s\,ds\Big) = \int_0^t \mathbb{E}(B_s)\,ds = 0$, and

$$\begin{aligned}\mathbb{E}\Big(\Big(\int_0^t B_s\,ds\Big)^2\Big) &= \mathbb{E}\Big(\int_0^t\int_0^t B_sB_w\,dw\,ds\Big) = \int_0^t\int_0^t \mathbb{E}(B_sB_w)\,dw\,ds \\ &= \int_0^t\int_0^s w\,dw\,ds + \int_0^t\int_s^t s\,dw\,ds = \frac{1}{2}\int_0^t s^2\,ds + \int_0^t (st - s^2)\,ds = \frac{1}{3}t^3.\end{aligned}$$

(2) We have $\mathbb{E}\Big(\int_0^t B_s^2\,ds\Big) = \int_0^t \mathbb{E}(B_s^2)ds = \int_0^t s\,ds = \frac{t^2}{2}$. Moreover,

$$\mathbb{E}\Big(\Big(\int_0^t B_s^2\,ds\Big)^2\Big) = \int_0^t\int_0^t \mathbb{E}(B_s^2B_w^2)\,dw\,ds = 2\int_0^t\int_0^s \mathbb{E}(B_s^2B_w^2)\,dw\,ds$$

(this is not completely obvious. Check it!). It follows form the martingale property of $B_t^2 - t$, that $\mathbb{E}(B_s^2 - s|\mathscr{F}_w) = B_w^2 - w$ for $w \le s$ and hence

$$\mathbb{E}(B_s^2B_w^2|\mathscr{F}_w) = B_w^2\mathbb{E}(B_s^2|\mathscr{F}_w) = B_w^2(B_w^2 + s - w).$$

Then $\mathbb{E}(B_s^2B_w^2) = \mathbb{E}(B_w^2(B_w^2 + s - w)) = 3w^2 + (s-w)w = 2w^2 + sw$. Alternatively you may calculate

$$\begin{aligned}\mathbb{E}(B_s^2B_w^2) &= \mathbb{E}((B_s - B_w)^2)\mathbb{E}(B_w^2) + 2\mathbb{E}((B_s - B_w)B_w^3) + \mathbb{E}(B_w^4) \\ &= (s-w)w + 2\mathbb{E}((B_s - B_w))\mathbb{E}(B_w^3) + 3w^2 = (s-w)w + 3w^2 = 2w^2 + sw.\end{aligned}$$

It follows that

$$\mathbb{E}\Big(\Big(\int_0^t B_s^2\,ds\Big)^2\Big) = 2\int_0^t\int_0^s \big(2w^2 + sw\big)\,dw\,ds = 4\int_0^t \frac{1}{3}s^3\,ds + \int_0^t s^3\,ds = \frac{1}{3}t^4 + \frac{1}{4}t^4 = \frac{7}{12}t^4.$$

Solution to 22: Note that both stochastic integrals define martingales and hence their expectation is zero (and hence the variance is equal to the second moment).

(1) We have

$$\mathbb{E}\Big(\big(\int_0^t \sqrt{|B_s|}\,dB_s\big)^2\Big) = \mathbb{E}\Big(\int_0^t |B_s|\,ds\Big) = \int_0^t \mathbb{E}(|B_s|)\,ds.$$

For a normal random variable Z with mean 0 and variance s, $\mathbb{E}(|Z|) = \sqrt{\frac{2}{\pi}}\sqrt{s}$. Hence

$$\mathbb{E}\Big(\int_0^t |B_s|\,ds\Big) = \sqrt{\frac{2}{\pi}}\int_0^t \sqrt{s}\,ds = \frac{2}{3}\sqrt{\frac{2}{\pi}}t^{3/2}.$$

(2) Here

$$\mathbb{E}\Big(\big(\int_0^t (B_s+s)^2\,dBs\big)^2\Big) = \mathbb{E}\Big(\int_0^t (B_s+s)^4\,ds\Big) = \int_0^t \mathbb{E}((B_s+s)^4)\,ds.$$

We end up with

$$\mathbb{E}((B_s+s)^4) = \mathbb{E}(B_s^4) + 4s\mathbb{E}(B_s^3) + 6s^2\mathbb{E}(B_s^2) + 4s^3\mathbb{E}(B_s) + s^4 = 3s^2 + 6s^3 + s^4.$$

Solution to 23: ν_n is increasing a.s. Moreover $f_n(\omega,t) = f(B_t(\omega))\mathbb{1}_{t<\nu_n}$ is adapted and measurable since $f_n(\omega,t)$ is a right continuous process.

Also (and this is the idea of localization)

$$\mathbb{E}\Big(\int_0^T (f(B_t(\omega))\mathbb{1}_{t<\nu_n})^2\,dt\Big) \le \mathbb{E}\Big(\int_0^T (\sup_{|w|\le n} f(w))^2\,dt\Big) < \infty,$$

the sup being existent since f is continuous. Finally B_t is a.s. continuous on $[0,t]$ and hence bounded by some (random) constant C (with probability one). If $n > C$ then $\nu_n = T \wedge \min\{t : |B_t| \ge n\} = T$.

Solution to 24: Choose $f(u) = 1/(1+u)$. Then

$$\mathrm{Var(X_t)} = \int_0^t \mathrm{f^2(u)\,du} = \int_0^t \frac{1}{(1+u)^2}\,\mathrm{du} = \frac{t}{1+t}.$$

Hence

$$\int_0^t \frac{1}{1+u}\,dB_s(u).$$

Solution to 25: X_0 is clear. Clearly $X_t = Z_{\xi(t)}$ is continuous, because $t \mapsto \xi(t)$ is continuous. Also $\xi(t)$ is strictly increasing, so X is a time change of Z_t.

Using Proposition ⟦104⟧ it follows that Z_t is a mean zero Gaussian process with independent increments and $\mathrm{Cov(Z_t,Z_s)} = \int_0^{t\wedge s} \mathrm{f^2(u)\,du}$. Hence for $s < t$

$$\mathrm{Cov(X_t,X_s)} = \int_0^{\xi(s)} \mathrm{f^2(u)\,du = s}.$$

So $\mathrm{Cov(X_t,X_s) = t \wedge s}$. It follows that X_t is a standard Brownian motion.

D LITERATURE

Measure theory

[🕮1] HALMOS: *Measure theory*, Van Nostrand (1950)

Probability theory

[🕮2] FELLER: *Introduction to probability I & II*, Wiley (1969)

The classic book about probability theory.

[🕮3] KALLENBERG: *Foundation of Modern Probability*, Springer (2002)

You will find almost everything about probability theory in this book. It is written on a more advanced level.

Stochastic processes & applied probability

[🕮4] KARLIN & TAYLOR: *A First Course in Stochastic Processes*, 1975 (Academic Press)

[🕮5] KARLIN & TAYLOR: *A Second Course in Stochastic Processes*, 1981 (Academic Press)

[🕮6] ROSS: *Introduction to Probability Models*, Elsevier/Academic Press (2006)

More elementary approach to probability models. No measure theory involved.

Stochastic integration

[🕮7] STEELE: *Stochastic calculus and financial applications*, Springer (2001)

Is used for this lecture.

[🕮8] PROTTER: *Stochastic Integration and Differential Equations*, Springer (2004)

The standard reference to stochastic integration.