Stochastic Analysis

Achref Bachouch and Nicolas Perkowski

August 1, 2016

Introduction

Stochastic analysis is the study of continous time stochastic processes. In Stochastics II we have encountered discrete time processes and have seen how they can be constructed and used to model different phenomena that evolve in discrete time steps and undergo random influences. Since physical time is continuous, it is the natural next question how to extend this to model continuous time phenomena. Naturally, this is more complex because now it does not suffice any more to describe how the system transitions "from one step to the next". We will only consider stochastic processes with values in the Euclidean space $\mathbb{R}^d$, but many of the tools we develop are useful also in more complex situations. And in fact the Euclidean case is already complicated and interesting enough in its own right.

To motivate the tools and results that we develop in the lecture, let us look at some examples.

Example 0.1 We all know the pictures of stock price trajectories and that they look very irregular, bouncing up and down constantly. We will see that a reasonable first model for the evolution of stock prices is a (time-changed) Brownian motion. Recall that a *Brownian motion* is a continuous time stochastic process $(B_t)_{t\geqslant 0}$ with continuous trajectories, such that $B_t \sim \mathcal{N}(0, t)$ for all $t \geqslant 0$, where $\mathcal{N}(0, t)$ is the normal distribution with mean 0 and variance t, and such that $B_{t+s} - B_t$ is independent of $(B_r)_{0\leqslant r\leqslant t}$. We will see later in the lecture how to construct such a process and that the description above characterizes it uniquely. And we will study some basic path properties of the Brownian motion to see that it indeed behaves quite wildly (PLOT). For example it has no isolated zeros, meaning that if $B_t = 0$ for some t, then in any small interval $[t-\varepsilon, t+\varepsilon]$ there are infinitely many s with $B_s = 0$. We will also see that B is nowhere differentiable and behaves roughly speaking like

$$B_{t+\mathrm{d}t} - B_t \simeq \sqrt{\mathrm{d}t}.$$

Of course, this is not a mathematical statement and part of the work will be to find a suitable mathematical statement that we can actually prove.

Example 0.2 From Stochastics II discrete time processes are familiar. An example for a discrete time evolution is the difference equation

$$X_{n+1} = X_n + b(X_n) + f(X_n, Y_n),$$

where Y_n is random influence, "noise".

A concrete example would be the Malthusian population growth model, where

$$X_{n+1} = X_n + bX_n + X_n Y_n,$$

with b describing the deterministic growth rate and $(Y_n)_{n\in\mathbb{N}}$ being a centered family of independent and identically distributed (i.i.d.) random variables that models randomly occurring deviations from the deterministic growth rate.

If we assume that the random influence Y_n is small, we can apply Taylor's theorem to get

$$X_{n+1} \simeq X_n + b(X_n) + f(X_n, 0) + \partial_y f(X_n, 0) Y_n.$$

If we assume again that the $(Y_n)_{n\in\mathbb{N}}$ are a centered family of (i.i.d.) random variables with finite variance, then we know from Stochastics II that $S_n = \sum_{k=1}^n Y_k$ can be rescaled so that it converges to a Brownian motion. When performing this scaling limit we say that S is evolving in discrete time steps, but the steps are essentially infinitely small so that we can approximate them by continuous time steps.

It then seems reasonable to expect (and indeed it can be shown under suitable assumptions) that if X is evolving in very small time steps, then it can be rescaled in such a way that it converges to a process $(Z_t)_{t\geqslant 0}$ satisfying for $t \geqslant 0$ and $h > 0$

$$Z_{t+h} = Z_t + (b(Z_t) + f(Z_t, 0))h + \partial_y f(Z_t, 0)(B_{t+h} - B_t),$$

where B is a Brownian motion. Bringing Z_t to the left hand side, dividing by h and letting $h \to 0$, we get formally

$$\partial_t Z_t = b(Z_t) + f(Z_t, 0) + \partial_y f(Z_t, 0)\partial_t B_t.$$

But the problem is that B is not differentiable, so it is not clear how to interpret this equation! To make sense of such "stochastic differential equations" will be one of the main goals of the lecture.

Example 0.3 Another situation where stochastic differential equations appear is the following: Applied scientists are often able to derive an ordinary differential equation (ODE)

$$\dot{X}_t = b(X_t)$$

to describe the time evolution of a given system. However, in reality the system will not be isolated from its environment, so the environment will influence the system. Now we have two choices. Either we increase the dimension of our system by attempting to also model the entire environment, which ultimately leads to an infinite-dimensional system and is actually unfeasible because nature is just too complex. Or we try to model the influence of the environment as "random". Under suitable assumptions we should be able to invoke the central limit theorem, so that these random influences should be centered Gaussians. In many situations it is also reasonable to assume that they are stationary in time, and independent for different times. So formally we end up with an equation

$$\dot{X}_t = b(X_t) + \xi_t,$$

where $(\xi_t)_{t\geqslant 0}$ is an i.i.d. family of centered Gaussian variables. It turns out that this equation does not make sense, because it is not possible to construct "a version" of ξ that has measurable trajectories and then it is not clear how to interpret the equation. The solution to this problem is to formally assume that $\xi(t)$ has infinite variance for fixed times. We will see how to make this rigorous and how to model an ODE forced by a "white noise" (which turns out to be the time derivative of Brownian motion).

Example 0.4 A more concrete version of this example is a particle in a double well potential. Assume that we take $x : [0, \infty) \to \mathbb{R}$ and $b(x) = U'(x)$ for $U(x) = \frac{1}{4}x^4 - \frac{1}{2}x^2$. One can easily verify that there are two stable fixpoints for the dynamics, $\{-1, 1\}$ (PLOT, imagine a ball rolling down, except damping at the bottom), and one unstable fixpoint $\{0\}$. So if we start in $x < 0$ we will converge to -1 for $t \to \infty$, and if we start in $x > 0$ we will converge to 1. Such a simple system can be already used to model the qualitative behavior of earth's climate: Assume -1 represents an ice age and $+1$ a warm period. These two states are relatively stable for the climate, after all we are not constantly switching between ice ages and warm periods. But from time to time there are transitions, and in the deterministic model we wrote down we will never see them. But if we add a very small random forcing of white noise type, as described above, then the forcing can "kick" the solution over the hill into the domain of attraction of the other stable fixpoint. It is then for example interesting to calculate how long this should take.

Example 0.5 Observe that if B is a Brownian motion, then for $t > 0$ the random variable B_t has the density

$$p(t, x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right).$$

Now it is a simple exercise to verify that p solves the *heat equation*:

$$\partial_t p(t, x) = \frac{1}{2}\partial^2_{xx} p(t, x)$$

for all $t > 0$ and $x \in \mathbb{R}$. As a consequence, we get that if φ is a nice function, then

$$u(t, x) = \mathbb{E}[\varphi(x + B_t)]$$

solves

$$\partial_t u(t, x) = \partial_t \int_{\mathbb{R}} \varphi(x + y) p(t, y) \mathrm{d}y = \int_{\mathbb{R}} \varphi(x + y) \frac{1}{2} \partial^2_{yy} p(t, y) \mathrm{d}y \tag{1}$$

$$= \int_{\mathbb{R}} \frac{1}{2} \partial^2_{yy} \varphi(x + y) p(t, y) \mathrm{d}y = \int_{\mathbb{R}} \frac{1}{2} \partial^2_{xx} \varphi(x + y) p(t, y) \mathrm{d}y = \frac{1}{2} \partial^2_{xx} u(t, x), \tag{2}$$

where we applied integration by parts to shift ∂^2_{yy} from p to φ. Moreover, obviously $u(0, x) = \varphi(x)$, so that we found a solution to the equation

$$\partial_t u = \partial^2_{xx} u, \qquad u(0) = \varphi.$$

This suggests a link between stochastic processes and partial differential equations, and in fact this link is quite deep and powerful. For example if for $x \in \mathbb{R}$ the process X^x solves the stochastic differential equation

$$\partial_t X^x_t = b(X^x_t) + \sigma(X^x_t) \partial_t B_t, \qquad X_0 = x,$$

then $u(t,x) = \mathbb{E}[\varphi(X_t^x)]$ solves the partial differential equation (PDE)

$$\partial_t u(t,x) = b(x)\partial_x u(t,x) + \frac{1}{2}\sigma^2(x)\partial_{xx}^2 u(t,x), \qquad u(0) = \varphi,$$

and conversely the PDE can be used to characterize the law of X^x.

Hopefully these examples show that there are many interesting questions to be asked and problems to be studied. We will now start to develop the basic tools and methods of stochastic analysis. The notes are based mainly on Le Gall's notes and on those of Jacod.

Notation and conventions

$\mathbb{N} = \{1, 2, \ldots\}$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, $\mathbb{R}_+ = [0, \infty)$. Inner product on $\mathbb{R}^d$ is $x \cdot y = \sum_{j=1}^d x_j y_j$. Transpose is denoted with A^T. The Borel sigma algebra of a topological space E is $\mathcal{B}(E)$. If we do not specify it, we always assume an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as given. $a \lesssim b$ means there exists some $C > 0$, independent of the variables under consideration, such that $a \leqslant Cb$.

1 Gaussian processes

1.1 Quick recap on Gaussian random variables

Definition 1.1 *A random variable X is called standard Gaussian (or standard normal) if X has the density*

$$p_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

with respect to Lebesgue measure.

Definition 1.2 *Let $m \in \mathbb{R}$ and $\sigma \geqslant 0$. A random variable Y has the Gaussian distribution $\mathcal{N}(m, \sigma^2)$ if there exists a standard normal variable X such that*

$$Y = m + \sigma X. \tag{3}$$

Equivalently, $Y \sim \mathcal{N}(m, \sigma^2)$ if

$$\mathbb{E}[e^{iuY}] = e^{ium - \sigma^2 u^2/2}, \qquad u \in \mathbb{R}. \tag{4}$$

A random variable Y is (centered) Gaussian if it has distribution $\mathcal{N}(m, \sigma^2)$ for some $m \in \mathbb{R}$ $(m = 0)$, $\sigma \geqslant 0$.

Remark 1.3

i. Let $\sigma > 0$ and $m \in \mathbb{R}$. Then $Y \sim \mathcal{N}(m, \sigma^2)$ if and only if Y has the density

$$p_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-m)^2}{2\sigma^2}\right).$$

For $\sigma = 0$ we have $Y \sim \mathcal{N}(m, \sigma^2)$ if and only if $Y = m$ almost surely.

ii. If $Y \sim \mathcal{N}(m, \sigma^2)$ and $\tilde{Y} \sim \mathcal{N}(\tilde{m}, \tilde{\sigma}^2)$ are independent Gaussian random variables, then (4) immediately yields that $Y + \tilde{Y} \sim \mathcal{N}(m + \tilde{m}, \sigma^2 + \tilde{\sigma}^2)$.

Lemma 1.4 *Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of Gaussian random variables such that $X_n \sim \mathcal{N}(m_n, \sigma_n^2)$. Then X_n converges in distribution if and only if there exist $m \in \mathbb{R}$ and $\sigma \geqslant 0$ such that $m_n \to m$ and $\sigma_n \to \sigma$. In that case the limit in distribution X of the (X_n) satisfies $X \sim \mathcal{N}(m, \sigma^2)$.*

Definition 1.5 *Let $d \in \mathbb{N}$ and let X be a random variable with values in $\mathbb{R}^d$. Then we say that X is Gaussian if for every $u \in \mathbb{R}^d$ the random variable*

$$u \cdot X = \sum_{j=1}^{d} u_j X_j$$

is Gaussian.

Lemma 1.6 *If X is Gaussian, then there exists $m \in \mathbb{R}^d$ and a symmetric positive semi-definite matrix $C \in \mathbb{R}^{d \times d}$ such that for all $u \in \mathbb{R}^d$*

$$\mathbb{E}[u \cdot X] = u \cdot m, \qquad \operatorname{var}(u \cdot X) = u^T C u.$$

In particular, the characteristic function of X is given by

$$\mathbb{E}[e^{iu \cdot X}] = e^{iu \cdot m - (u^T C u)/2}, \qquad u \in \mathbb{R}^d,$$

and m and C determine the law of X uniquely. We therefore also write $X \sim \mathcal{N}(m, C)$. If $m = 0$ we say that X is centered.

Corollary 1.7 *Let $m \in \mathbb{R}^d$, let $C \in \mathbb{R}^{d \times d}$ be a symmetric and positive semi-definite matrix, and let $X \sim \mathcal{N}(m, C)$. Then X has a density p_X with respect to the d-dimensional Lebesgue measure if and only if C is invertible. In that case*

$$p_X(x) = \frac{1}{(2\pi)^{d/2} (\det(C))^{1/2}} \exp\left(-\frac{1}{2} (x - m)^T C^{-1} (x - m) \right). \tag{5}$$

Corollary 1.8 *Let $m \in \mathbb{R}^d$, let $C \in \mathbb{R}^{d \times d}$ be a symmetric and positive semi-definite matrix, and let $X \sim \mathcal{N}(m, C)$. Then the coordinates $(X_1, \dots, X_d)$ are independent if and only if C is a diagonal matrix.*

Remark 1.9 If X_1 and X_2 are Gaussian random variables with $\operatorname{cov}(X_1, X_2) = 0$, then it is not necessarily true that X_1 and X_2 are independent. For Corollary 1.8 we need that (X_1, X_2) is a two-dimensional Gaussian vector, which implies that X_1 and X_2 are Gaussian random variables but is stronger than that.

Corollary 1.10 *Let $m \in \mathbb{R}^d$, let $C \in \mathbb{R}^{d \times d}$ be a symmetric and positive semi-definite matrix, and let $X \sim \mathcal{N}(m, C)$. Let $Y \sim \mathcal{N}(0, I)$ where I is the identity matrix in $\mathbb{R}^{d \times d}$. Then X has the same distribution as*

$$m + \sqrt{C} Y,$$

where $\sqrt{C}$ is a symmetric square root of C, that is a symmetric and positive semi-definite matrix such that $\sqrt{C}\sqrt{C} = C$. In particular, for every $m \in \mathbb{R}^d$ and every symmetric and positive semi-definite $C \in \mathbb{R}^{d \times d}$ there exists a random variable with distribution $\mathcal{N}(m, C)$.

1.2 Stochastic processes

Definition 1.11 *Let $(E, \mathcal{E})$ be a measurable space and let $\mathbb{T}$ be an index set. A stochastic process $(X_t)_{t\in\mathbb{T}}$ (indexed by $\mathbb{T}$) with values in E is a family of random variables with values in E. If we do not specify $(E, \mathcal{E})$, then we usually mean $E = \mathbb{R}$ and $\mathcal{E} = \mathcal{B}(\mathbb{R})$. The maps*

$$\mathbb{T} \ni t \mapsto X_t(\omega) \in E, \qquad \omega \in \Omega,$$

are called the trajectories of X.

Definition 1.12 *Let $(E, \mathcal{E})$ be a measurable space and let $\mathbb{T}$ be an index set. The Kolmogorov sigma algebra on $E^{\mathbb{T}}$ is defined as the smallest sigma algebra with respect to which the maps*

$$E^{\mathbb{T}} \ni x \mapsto x(t) \in E, \qquad t \in \mathbb{T},$$

are measurable. We denote it with $\mathcal{E}^{\otimes\mathbb{T}}$.

Lemma 1.13 *Let $(E, \mathcal{E})$ be a measurable space and let $\mathbb{T}$ be an index set. Given $I \in \mathbb{T}$ let us write $\mathcal{G}(I)$ for the sigma algebra generated by $(x(t))_{t\in I}$, that is the smallest sigma algebra with respect to which the maps*

$$E^{\mathbb{T}} \ni x \mapsto x(t) \in E, \qquad t \in I,$$

are measurable. In particular $\mathcal{G}(\mathbb{T}) = \mathcal{E}^{\otimes\mathbb{T}}$. Then

$$\mathcal{G}(\mathbb{T}) = \bigvee_{I\subset\mathbb{T}} \mathcal{G}(I) = \bigvee_{I\subset\mathbb{T}:|I|<\infty} \mathcal{G}(I) = \bigvee_{t\in\mathbb{T}} \mathcal{G}(\{t\}). \tag{6}$$

Moreover, $\mathcal{G}^0 = \bigcup_{I\subset\mathbb{T}:|I|<\infty} \mathcal{G}(I)$ is an algebra (but in general not a sigma algebra).

Definition 1.14 *The law of X is the probability measure* $\mathrm{law}(X) = \mathbb{P} \circ X^{-1}$ *on $(E^{\mathbb{T}}, \mathcal{E}^{\otimes\mathbb{T}})$.*

Definition 1.15 *The family of finite-dimensional distributions of a stochastic process X indexed by $\mathbb{T}$ is the family of probability measures $(\mu_I : I \subset \mathbb{T}, |I| < \infty)$, where we write for $I \subset T$*

$$\mu_I = \mathbb{P} \circ (X_t)^{-1}_{t\in I},$$

so that μ_I is a probability measure on $(E^I, \mathcal{E}^{\otimes I})$. If μ is a probability measure on $(E^{\mathbb{T}}, \mathcal{E}^{\otimes\mathbb{T}})$, then we define the finite dimensional distributions of μ as those of the canonical process on $E^{\mathbb{T}}$.

Lemma 1.16 *If μ and $\tilde{\mu}$ are two probability measures on $(E^{\mathbb{T}}, \mathcal{E}^{\otimes\mathbb{T}})$ that have the same finite dimensional distributions, that is $\mu_I = \tilde{\mu}_I$ for all $I \subset \mathbb{T}$ with $|I| < \infty$, then $\mu = \tilde{\mu}$.*

Theorem 1.17 *(Kolmogorov's extension theorem)*

Let E be a Polish space (complete separable metric space) and let $\mathcal{E}$ be the Borel sigma algebra of E. Let $\mathbb{T}$ be an index set and assume that for every $I \subset \mathbb{T}$ with $|I| < \infty$ we are given a probability measure μ_I on $(E^I, \mathcal{E}^{\otimes I})$. Then there exists a stochastic process X with finite dimensional distributions $(\mu_I : I \subset \mathbb{T}, |I| < \infty)$ if and only if the (μ_I) satisfy the Kolmogorov consistency relation

$$\mu_{I\cup\{t\}}(A \times E) = \mu_I(A), \qquad A \in \mathcal{E}^{\otimes I},$$

for all $I \subset \mathbb{T}$ finite and all $t \in \mathbb{T}$.

1.3 Gaussian proceses

Definition 1.18 *Let $\mathbb{T}$ be an index set. A real-valued stochastic process $X = (X_t)_{t\in\mathbb{T}}$ is called a (centered) Gaussian process if for every finite subset $I \subset \mathbb{T}$ and for every $(\alpha_t)_{t\in I} \in \mathbb{R}^I$ the random variable $\sum_{t\in I} \alpha_t X_t$ is (centered) Gaussian.*

Theorem 1.19 *Let $\mathbb{T}$ be an index set, let $m : \mathbb{T} \to \mathbb{R}$, and let Γ be a symmetric and positive semi-definite function. Then there exists a Gaussian process X with mean m and covariance Γ, and the law of X is uniquely determined by m and Γ.*

Example 1.20 Let $(E, \mathcal{E}, \mu)$ be a sigma-finite measure space. Then there exists a (unique-in-law) centered Gaussian process $(\xi(f) : f \in L^2(E, \mu))$ such that

$$\mathrm{cov}(\xi(f), \xi(g)) = \mathbb{E}[\xi(f)\xi(g)] = \int_E f(x)g(x)\mu(\mathrm{d}x) =: \langle f, g\rangle_{L^2(E,\mu)}, \qquad f, g \in L^2(E, \mu).$$

The process ξ is called *Gaussian measure with intensity* μ. We have for $f, g \in L^2(E, \mu)$

$$\begin{aligned}
\mathbb{E}[|\xi(f+g) - (\xi(f) + \xi(g))|^2] &= \mathbb{E}[\xi(f+g)^2] + \mathbb{E}[\xi(f)^2] + \mathbb{E}[\xi(g)^2] && (7)\\
&-2\mathbb{E}[\xi(f+g)\xi(f)] - 2\mathbb{E}[\xi(f+g)\xi(g)] + 2\mathbb{E}[\xi(f)\xi(g)] && (8)\\
&= \langle f+g, f+g\rangle_{L^2(E,\mu)} + \langle f, f\rangle_{L^2(E,\mu)} + \langle g, g\rangle_{L^2(E,\mu)} && (9)\\
&-2\langle f+g, f\rangle_{L^2(E,\mu)} - 2\langle f+g, g\rangle_{L^2(E,\mu)} + 2\langle f, g\rangle_{L^2(E,\mu)} && (10)\\
&= 2\langle f, f\rangle_{L^2(E,\mu)} + 2\langle g, g\rangle_{L^2(E,\mu)} + 2\langle f, g\rangle_{L^2(E,\mu)} && (11)\\
&-\langle f, f\rangle_{L^2(E,\mu)} - 2\langle g, g\rangle_{L^2(E,\mu)} - 2\langle f, g\rangle_{L^2(E,\mu)} = 0, && (12)
\end{aligned}$$

so that $\xi(f+g) = \xi(f) + \xi(g)$ almost surely and therefore ξ is a linear map from $L^2(E, \mu)$ to $L^2(\Omega, \mathbb{P})$. In combination with $\mathbb{E}[\xi(f)^2] = \|f\|^2_{L^2(E,\mu)}$, we deduce that ξ is an isometry from $L^2(E, \mu)$ to $L^2(\Omega, \mathbb{P})$.

Remark 1.21 Let ξ be a Gaussian measure with intensity μ and define the set function $G(A) = \xi(\mathbb{I}_A)$ for all $A \in \mathcal{E}$ with $\mu(A) < \infty$. Then we have $G(\emptyset) = 0$ almost surely. If $(A_n)_{n\in\mathbb{N}}$ is a disjoint family of sets in $\mathcal{E}$ with $\mu(\cup_n A_n) < \infty$, then $\sum_n \mathbb{I}_{A_n}$ converges to $\mathbb{I}_{\cup_n A_n}$ in $L^2(E, \mu)$. So since ξ is an isometry, we obtain

$$G(\cup_n A_n) = \xi(\mathbb{I}_{\cup_n A_n}) = \sum_n \xi(\mathbb{I}_{A_n}) = \sum_n G(A_n),$$

where the equality holds almost surely and the right hand side converges in $L^2(\Omega, \mathbb{P})$. This suggests that G might be a (signed) measure. However, this is nearly always false because we have to be careful with the position of "almost surely": It is in general not possible to isolate one null set N such that for $\omega \in \Omega \setminus \{0\}$ and for every sequence $(A_n)_{n\in\mathbb{N}}$ of disjoint sets in $\mathcal{E}$ the indentity

$$G(\cup_n A_n)(\omega) = \sum_n G(A_n)(\omega) \tag{13}$$

holds. Instead, the null set where (13) fails depends on the sequence (A_n) and of course in general there are uncountably many sequences of disjoint sets.

Example 1.22 The Gaussian measure on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ with Lebesgue measure as density is called *white noise*. If ξ is a white noise, then

$$B_t = \xi(\mathbb{I}_{[0,t]}), \qquad t \geqslant 0,$$

is a centered Gaussian process with covariance

$$\mathrm{cov}(B_s, B_t) = \mathbb{E}[B_s B_t] = \mathbb{E}[\xi(\mathbb{I}_{[0,t]})\xi(\mathbb{I}_{[0,s]})] = \int_{\mathbb{R}_+} \mathbb{I}_{[0,t]}(x)\mathbb{I}_{[0,s]}(x)\mathrm{d}x = s \wedge t,$$

where we write $s \wedge t = \min\{s, t\}$. The centered Gaussian process with this covariance is called *pre-Brownian motion.*

Remark 1.23 Physicists usually say that the white noise is the centered Gaussian process with covariance $\mathbb{E}[\xi(x)\xi(y)] = \delta(x - y)$ for $x, y \in \mathbb{R}$, where δ is the Dirac delta in 0. This does not make sense rigorously, but let us see how to relate it to our definition: If ξ has this covariance, then we formally have

$$\mathbb{E}[\xi(f)\xi(g)] = \mathbb{E}\left[\int_{\mathbb{R}_+} f(x)\xi(x)\mathrm{d}x \int_{\mathbb{R}_+} g(y)\xi(y)\mathrm{d}y\right] = \int_{\mathbb{R}_+\times\mathbb{R}_+} f(x)g(y)\mathbb{E}[\xi(x)\xi(y)]\mathrm{d}x\mathrm{d}y \tag{14}$$

$$= \int_{\mathbb{R}_+\times\mathbb{R}_+} f(x)g(y)\delta(x-y)\mathrm{d}x\mathrm{d}y = \int_{\mathbb{R}_+} f(x)g(x)\mathrm{d}x, \tag{15}$$

so ξ is a white noise. So formally we obtain that a white noise is a family $(\xi(x))_{x\in\mathbb{R}_+}$ of independent centered Gaussian variables, such that $\mathrm{var}(\xi(x)) = \delta(0) = \infty$ for all $x \in \mathbb{R}_+$. Of course all of these manipulations are only allowed for physicists and not for mathematicians, and in particular we cannot evaluate a white noise in a single point but only test it against $L^2(\mathbb{R}_+, \mathrm{d}x)$ functions!

Lemma 1.24 *Let $(B_t)_{t\geqslant 0}$ be a real valued stochastic process. Then B is a pre-Brownian motion if and only if the following conditions are satisfied:*

i. $B_0 = 0$ *almost surely;*

ii. for all $0 \leqslant s < t$ *the random variable* $B_t - B_s$ *is independent of the variables* $(B_r)_{0\leqslant r\leqslant s}$*;*

iii. for all $0 \leqslant s < t$ *we have* $B_t - B_s \sim \mathcal{N}(0, t-s)$.

2 The Brownian motion

2.1 Continuity of stochastic processes

We say that a stochastic process $X = (X_t)_{t\geqslant 0}$ with values in $\mathbb{R}^d$ is *continuous* if all of its trajectories are continuous.

Definition 2.1 *A stochastic process $(B_t)_{t\geqslant 0}$ is called a Brownian motion if it is a continuous pre-Brownian motion. Such a process is also referred to as Wiener process.*

Example 2.2 Let X be a continuous stochastic process with values in $\mathbb{R}$, and let T be a random variable which is uniformly distributed on $[0,1]$. Then

$$\tilde{X}_t = X_t + \mathbb{I}_{\{T\}}(t), \qquad t \geqslant 0,$$

is discontinuous for all ω and satisfies $\mathbb{P}(\tilde{X}_t = X_t) = 1$ for all $t \geqslant 0$. In particular, $\tilde{X}$ and X have the same finite-dimensional distributions and thus the same law by Lemma 1.16.

Definition 2.3 *Let $X = (X_t)_{t\in\mathbb{T}}$ and $\tilde{X} = (\tilde{X}_t)_{t\in\mathbb{T}}$ be two stochastic processes with values in E.*

i. We say that $\tilde{X}$ is a modification of X if $\mathbb{P}(X_t = \tilde{X}_t) = 1$ for all $t \in \mathbb{T}$;

ii. if outside of a null set we have $X_t = \tilde{X}_t$ for all $t \in \mathbb{T}$, then X and $\tilde{X}$ are called indistinguishable.

Definition 2.4 *Define the Haar mother function $\chi : [0,1] \to \mathbb{R}$ as*

$$\chi(t) = \mathbb{I}_{[0,1/2)}(t) - \mathbb{I}_{[1/2,1)}(t).$$

For $n \in \mathbb{N}_0$ and $0 \leqslant k < 2^n$ we set

$$\chi_{n,k}(t) = 2^{n/2}\chi(2^n t - k) = 2^{n/2}(\mathbb{I}_{[2^{-n}k,2^{-n-1}(2k+1))}(t) - \mathbb{I}_{[2^{-n-1}(2k+1),2^{-n}(k+1))}(t)).$$

We also define $\chi_{-1,0}(t) = \mathbb{I}_{[0,1)}(t)$. The functions $(\chi_{n,k} : n \geqslant -1, 0 \leqslant k < 2^n)$ are called the Haar wavelets. By definition, all $\chi_{n,k}$ satisfy $\int_0^1 |\chi_{n,k}(t)|^2 \mathrm{d}t$ and it is not hard to see that $\int_0^1 \chi_{n,k}(t)\chi_{m,\ell}(t)\mathrm{d}t = \delta_{n,m}\delta_{k,\ell}$. As an exercise you may use that the indicator functions $(\mathbb{I}_{[k2^{-n},(k+1)2^{-n})}, n \in \mathbb{N}_0, 0 \leqslant k < 2^n)$ are dense in $L^2([0,1],\mathrm{d}t)$ to show that the Haar wavelets form an orthonormal basis of $L^2([0,1],\mathrm{d}t)$.

The integrated Haar functions will play an important role in what follows, so let us introduce the notation

$$\varphi_{n,k}(t) = \int_0^t \chi_{n,k}(s)\mathrm{d}s,$$

and we call $(\varphi_{n,k} : n \geqslant -1, 0 \leqslant k < 2^n)$ the *Schauder functions.* Observe that for $n \geqslant 0$ and $0 \leqslant k < 2^n$

$$\langle f', \chi_{n,k}\rangle_{L^2([0,1],\mathrm{d}t)} = 2^{n/2}(2f(2^{-n-1}(2k+1)) - f(2^{-n}k) - f(2^{-n}(k+1))),$$

which makes sense for any function f and does not require differentiability.

Definition 2.5 *For $\alpha \in (0,1)$ and $T > 0$ the space of α-Hölder continuous functions is defined as*

$$C^\alpha([0,T],\mathbb{R}) = \{f : [0,T] \to \mathbb{R}, \|f\|_\alpha < \infty\}, \qquad \text{where } \|f\|_\alpha = \frac{|f(t) - f(s)|}{|t-s|^\alpha}.$$

Lemma 2.6 *(Ciesielski)*
Let $\alpha \in (0,1)$ and let $(f_{n,k} : n \geqslant -1, 0 \leqslant k < 2^n)$ be real numbers such that

$$\sup_{n\geqslant -1} \max_{0\leqslant k<2^n} 2^{n(\alpha-1/2)}|f_{n,k}| = C < \infty. \tag{16}$$

Then the series $\sum_{n,k} f_{n,k}\varphi_{n,k}$ *converges uniformly and absolutely, and the limit* f *is* α*-Hölder continuous and satisfies* $\|f\|_\alpha \lesssim C$.

Conversely, if $f \in C^\alpha([0,1],\mathbb{R})$ *and* $f_{n,k} := \langle f', \chi_{n,k}\rangle_{L^2([0,1],\mathrm{d}t)}$, *then*

$$\sup_{n\geqslant -1} \max_{0\leqslant k<2^n} 2^{n(\alpha-1/2)} |\langle f', \chi_{n,k}\rangle_{L^2([0,1],\mathrm{d}t)}| \lesssim \|f\|_\alpha.$$

Remark 2.7 Since Ciesielski's result, the relationship between many function spaces and (wavelet) bases has been explored in countless works. For an overview see Triebel.

Lemma 2.8 *Let* $\alpha \in \mathbb{R}$ *and* $p \in [1,\infty]$ *and let* $(f_{n,k} : n \geqslant -1, 0 \leqslant k < 2^n)$ *be real numbers. Then we define*

$$\|(f_{n,k})\|_{B_p^\alpha} := \left(\sum_n 2^{np(\alpha-1/2-1/p)} \left(\sum_k |f_{n,k}|^p\right)\right)^{1/p} = \|(2^{n(\alpha-1/2-1/p)}\|(f_{n,k})_k\|_{\ell^p})_n\|_{\ell^p} \quad (17)$$

with the usual interpretation as supremum norm if $p = \infty$. *Then we get for all* $1 \leqslant p \leqslant q \leqslant \infty$

$$\|(f_{n,k})\|_{B_q^{\alpha-(1/p-1/q)}} \leqslant \|(f_{n,k})\|_{B_p^\alpha}.$$

Theorem 2.9 *(Kolmogorov's continuity criterion)*

Let $T > 0$ *and let* $(\tilde{X}_t)_{t\in[0,T]}$ *be a real valued stochastic process such that there exists* $p \in [1,\infty)$, $\alpha \in (1/p, 1]$ *and* $M \geqslant 0$ *with*

$$\mathbb{E}[|\tilde{X}_t - \tilde{X}_s|^p]^{1/p} \leqslant M|t-s|^\alpha. \quad (18)$$

Then there exists a continuous modification X *of* $\tilde{X}$ *such that for every* $\beta \in (0, \alpha - 1/p)$ *there exists a constant* $C = C(\alpha,\beta,T) \geqslant 0$ *with*

$$\mathbb{E}[\|X\|_\beta^p]^{1/p} \leqslant CM.$$

Remark 2.10 Actually we never used that $\tilde{X}$ was real valued, Kolmogorov's continuity criterion also works if $\tilde{X}$ takes values in a Banach space. But to simplify the presentation we restrict ourselves to real-valued processes here.

Corollary 2.11 *The Brownian motion* B *exists and if* $T > 0$, *then* $(B_t)_{t\in[0,T]}$ *is almost surely in* $C^\alpha([0,T],\mathbb{R})$ *whenever* $\alpha < 1/2$ *and we even have*

$$\mathbb{E}[\|B\|^p_{C^\alpha([0,T],\mathbb{R})}] < \infty, \qquad p \in [1,\infty).$$

Remark 2.12 Paul Lévy's original construction of the Brownian motion was exactly the same as ours, except written in a slightly different language. Translated into our terminology, he considered a sequence $(Z_{n,k})$ of independent standard normal variables and defined $B_t = \sum_{n,k} Z_{n,k}\varphi_{n,k}(t)$. Note that the $(Z_{n,k})$ have to be independent because if B is a Brownian motion, then $(\langle B', \chi_{n,k}\rangle_{L^2([0,1],\mathrm{d}t)})_{n,k}$ is a centered Gaussian process with

$$\mathbb{E}[\langle B', \chi_{n,k}\rangle_{L^2([0,1],\mathrm{d}t)} \langle B', \chi_{m,\ell}\rangle_{L^2([0,1],\mathrm{d}t)}] = \mathbb{E}[\xi(\chi_{n,k})\xi(\chi_{m,\ell})] = \delta_{n,m}\delta_{k,\ell},$$

where ξ is a white noise, we used the construction $B_t = \xi(\mathbb{I}_{[0,t]})$ of a pre-Brownian motion, and also that $(\chi_{n,k})$ is an orthonormal basis of $L^2([0,1],\mathrm{d}t)$

Remark 2.13 The Brownian motion is only Hölder-continuous on compact intervals, but not on $\mathbb{R}_+$: The variables $(B_{n+1} - B_n)_{n\in\mathbb{N}_0}$ are independent standard Gaussians, and therefore their supremum is almost surely infinite: For any $C \in \mathbb{R}$

$$\mathbb{P}(\sup_{n\in\mathbb{N}_0} (B_{n+1} - B_n) \leqslant C) = \mathbb{P}\left(\bigcap_{n\in\mathbb{N}_0} \{B_{n+1} - B_n \leqslant C\}\right) = \prod_{n\in\mathbb{N}_0} \mathbb{P}(B_{n+1} - B_n \leqslant C),$$

and since $\mathbb{P}(B_{n+1} - B_n \leqslant C) = \mathbb{P}(B_1 - B_0 \leqslant C) < 1$, the infinite product on the right hand side is 0. In particular we have almost surely

$$\sup_{0\leqslant s<t<\infty} \frac{|B_t - B_s|}{|t-s|^\alpha} \geqslant \sup_{n\in\mathbb{N}_0} \frac{|B_{n+1} - B_n|}{1^\alpha} = \infty,$$

independently of $\alpha \in \mathbb{R}$.

2.2 Path properties of the Brownian motion

Proposition 2.14

i. $(-B_t)_{t\geqslant 0}$ *is a Brownian motion;*

ii. *more generally* $(\lambda^{-1} B_{\lambda^2 t})_{t\geqslant 0}$ *for* $\lambda \in \mathbb{R} \setminus \{0\}$ *is a Brownian motion;*

iii. $(B_{t+s} - B_s)_{t\geqslant 0}$ *for* $s \geqslant 0$ *is a Brownian motion, and is independent of* $(B_r)_{r\in[0,s]}$ *(Markov property);*

iv. $(tB_{1/t})_{t\geqslant 0}$ *where we set* $0B_{1/0} := 0$ *is indistinguishable from a Brownian motion.*

Theorem 2.15 *With probability* 1 *there exists no* $t \geqslant 0$ *at which* B *is differentiable.*

Exercise 2.16 Adapt the proof of Theorem 2.15 to show that if $\alpha > 1/2$, then with probability 1 there exists no $t \geqslant 0$ with

$$\limsup_{s\to t} \frac{|B_s - B_t|}{|s-t|^\alpha} < \infty.$$

Why does the same argument not work for $\alpha = 1/2$?

Exercise 2.17 Indeed it is not true that the Brownian motion is nowhere 1/2-Hölder continuous: there are so-called "slow points" where the Brownian motion shows an exceptional behavior. This is beyond the scope of our lecture. But show that if $t \geqslant 0$ is fixed, then almost surely

$$\limsup_{s\to t} \frac{|B_s - B_t|}{|s-t|^{1/2}} = \infty.$$

Conclude that almost surely there exists no interval $[S, T]$ with $0 \leqslant S < T$ on which the Brownian motion is 1/2-Hölder continuous.

Hint: The argumentation in Remark 2.13 might be helpful.

Corollary 2.18 *With probability* 1 *we have for any* $\alpha > 1/2$

$$0 = \limsup_{t\to\infty} \frac{|B_t|}{t^\alpha} < \limsup_{t\to\infty} \frac{|B_t|}{t^{1/2}} = \infty.$$

Proposition 2.19 *(Quadratic variation)*

Let $t > 0$ and let $0 = t_0^n < t_1^n < \ldots < t_{k_n}^n = t$ be a sequence of partitions of $[0,t]$ with $\max_{0 \leqslant i < k_n} |t_{i+1}^n - t_i^n|$ converging to zero as $n \to \infty$. Then

$$\lim_{n \to \infty} \sum_{i=0}^{k_n - 1} (B_{t_{i+1}^n} - B_{t_i^n})^2 = t,$$

where the convergence takes place in $L^2(\Omega, \mathbb{P})$.

3 Filtrations and stopping times

3.1 Filtrations and stopping times

Definition 3.1 *A filtration is an increasing family $\mathbb{F} = (\mathcal{F}_t)_{t \geqslant 0}$ of sub sigma algebras of $\mathcal{F}$, i.e. such that $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ whenever $0 \leqslant s \leqslant t$. In that case we write*

$$\mathcal{F}_\infty = \bigvee_{t \geqslant 0} \mathcal{F}_t \qquad \text{and} \qquad \mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s, \qquad t \geqslant 0.$$

A filtration is called right-continuous if $\mathcal{F}_{t+} = \mathcal{F}_t$ for all $t \geqslant 0$. If $\mathbb{F}$ is a filtration we write $\mathbb{F}^+ = (\mathcal{F}_t^+)_{t \geqslant 0}$ for the smallest right-continuous filtration containing $\mathbb{F}$, given by

$$\mathcal{F}_t^+ = \bigcap_{s > t} \mathcal{F}_s = \mathcal{F}_{t+}, \qquad t \geqslant 0.$$

Note that $(\mathbb{F}^+)^+ = \mathbb{F}^+$ for every filtration.

Example 3.2 An important example is $\mathcal{F}_t = \sigma(X_s : s \leqslant t)$, where X is a stochastic process. In that case we write $(\mathcal{F}_t)_{t \geqslant 0} = \mathbb{F}^X$ and we call $\mathbb{F}^X$ the *canonical filtration* of X. We also write $\mathbb{F}^{X+} = (\mathbb{F}^X)^+$. In general we have $\mathbb{F}^{X+} \neq \mathbb{F}^X$, even if X is continuous and real valued: For example we have

$$A := \{\omega : X(\omega) \text{ has a right derivative in } t\} \in \mathcal{F}_t^{X+},$$

but in general $A \notin \mathcal{F}_t^X$.

Definition 3.3 *A stochastic process $(X_t)_{t \geqslant 0}$ is called adapted to a given filtration $\mathbb{F}$ if X_t is $\mathcal{F}_t$-measurable for all $t \geqslant 0$.*

Definition 3.4 *Let $\mathbb{F}$ be a filtration. An $\mathbb{F}$-stopping time (or simply stopping time if there is no ambiguity about the filtration) is a map $T : \Omega \to [0, \infty]$ such that*

$$\{T \leqslant t\} \in \mathcal{F}_t$$

for all $t \geqslant 0$. If T is a $\mathbb{F}$-stopping time, then we write

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leqslant t\} \in \mathcal{F}_t \text{ for all } t \geqslant 0\} \tag{19}$$

for the sigma algebra of events determined until T.

Exercise 3.5 Let $\mathbb{F}$ be a filtration and T a stopping time.

i. Show that $\mathcal{F}_T$ defined in (19) is indeed a sigma algebra.

ii. Show that if $S(\omega) = t$ for all ω, where $t \in [0, \infty]$ is fixed, then S is a stopping time and $\mathcal{F}_t = \mathcal{F}_S$ where $\mathcal{F}_S$ is defined in (19). So our definitions are conistent.

iii. Show that $T + t$ is a stopping time whenever $t \in [0, \infty]$. We write $\mathcal{F}_{T+} = \bigcap_{t>0} \mathcal{F}_{T+t}$. Is the same true for $T - t$?

iv. Show that S is a $\mathbb{F}^+$-stopping time if and only if $\{S < t\} \in \mathcal{F}_t$ for all $t > 0$.

v. Show that

$$\mathcal{F}_T^+ = \{A \in \mathcal{F} : A \cap \{T < t\} \in \mathcal{F}_t \text{ for all } t > 0\}.$$

vi. Show that T is $\mathcal{F}_T$ measurable.

vii. Show that $\mathcal{F}_{T+} = \mathcal{F}_T^+$, where $\mathcal{F}_{T+} = \bigcap_{t>0} \mathcal{F}_{T+t}$.

viii. Show that if S is a stopping time with $S(\omega) \leqslant T(\omega)$ for all $\omega \in \Omega$, then $\mathcal{F}_S \subseteq \mathcal{F}_T$.

ix. Show that if S is a stopping time, then $S \vee T$ and $S \wedge T$ are stopping times and $\mathcal{F}_{S \wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$.
Hint: Use iv. and that $A \cap \{S \wedge T \leqslant t\} = (A \cap \{S \leqslant t\}) \cup (A \cap \{T \leqslant t\})$.

x. Show that if $(T_n)_{n \in \mathbb{N}}$ is a sequence of $\mathbb{F}$-stopping times, then $\sup_n T_n$ is a $\mathbb{F}$-stopping time and $\inf_n T_n$ is a $\mathbb{F}^+$-stopping time. Moreover, $\mathcal{F}_{(\inf_n T_n)+} = \bigcap_n \mathcal{F}_{T_n+}$.

xi. Show that if S is a $\mathbb{F}^+$ stopping time, then there exists a sequence of $\mathbb{F}$-stopping times $(S_n)_{n \in \mathbb{N}}$ with $\lim_{n \to \infty} S_n(\omega) = S(\omega)$ for all $\omega \in \Omega$, such that every S_n only takes finitely many values, we have $S_{n+1}(\omega) \leqslant S_n(\omega)$ for all $n \in \mathbb{N}$ and $\omega \in \Omega$, and $S_n(\omega) > S(\omega)$ for all $\omega \in \Omega$ with $S(\omega) < \infty$ and all $n \in \mathbb{N}$.
Hint: Take for example $S_n = (k+1)2^{-n}$ on the set $\{S \in [k2^{-n}, (k+1)2^{-n})\}$, $k = 0, \dots, n2^n - 1$, and $S_n = \infty$ on the set $\{S \geqslant n\}$.

Definition 3.6 *Let $X = (X_t)_{t \geqslant 0}$ be a stochastic process taking values in a measurable space $(E, \mathcal{E})$. For $A \in \mathcal{E}$ we define the entrance time of X into A as*

$$T_A(\omega) = \inf\{t \geqslant 0 : X_t(\omega) \in A\},$$

with $\inf \emptyset = \infty$.

Proposition 3.7 *Let E be a metric space and $\mathcal{E}$ be its Borel sigma algebra, and let X be adapted to the filtration $\mathbb{F}$.*

i. If $A \subset E$ is open and X is right-continuous or left-continuous, then T_A is an $\mathbb{F}^+$-stopping time.

ii. If $A \subset E$ is closed and X is continuous, then T_A is an $\mathbb{F}$-stopping time.

iii. If $E = \mathbb{R}$, $A = [K, \infty)$, and X is right-continuous and increasing, then T_A is an $\mathbb{F}$-stopping time.

Definition 3.8 *A stochastic process $X = (X_t)_{t \geqslant 0}$ with values in E is called measurable if the map*

$$\Omega \times \mathbb{R}_+ \ni (\omega, t) \mapsto X_t(\omega) \in \mathcal{E}$$

is $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+) - \mathcal{E}$ measurable.

It is called progressively measurable if for all $t \geqslant 0$ the map

$$\Omega \times [0,t] \ni (\omega, s) \mapsto X_s(\omega) \in \mathcal{E}$$

is $\mathcal{F}_t \otimes \mathcal{B}([0,t]) - \mathcal{E}$ measurable. The progressive sigma algebra consists of all $A \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)$ for which the process $\mathbb{I}_A$ is progressively measurable.

Lemma 3.9 *Let $(X_t)_{t\geqslant 0}$ be a stochastic process with values in E.*

i. If X is progressively measurable, then it is adapted and the map $t \mapsto X_t(\omega)$ is $\mathcal{B}(\mathbb{R}_+) - \mathcal{E}$ measurable for all $\omega \in \Omega$.

ii. If E is a topological space and X is right-continuous and adapted, then X is progressively measurable.

Remark 3.10 If X is left-continuous and adapted, then a simple adaptation of the arguments gives the same result as in ii.

Corollary 3.11 *If T is a stopping time and Y with values in $\mathbb{R}^d$ is $\mathcal{F}_T$-measurable, then the following processes are progressively measurable:*

$$(Y\mathbb{I}_{\{T\leqslant t\}})_{t\geqslant 0}, \qquad (Y\mathbb{I}_{\{T<t\}})_{t\geqslant 0}, \qquad (Y\mathbb{I}_{\{T=t\}})_{t\geqslant 0}.$$

Lemma 3.12 *Let X be a progressively measurable process and let T be a stopping time. Then $\omega \mapsto X_T(\omega)$ is $\mathcal{F}_T - \mathcal{E}$ measurable.*

Corollary 3.13 *If X is progressively measurable and T is a stopping time, then the stopped process*

$$X_t^T := X_{t\wedge T} = X_t\mathbb{I}_{\{t\leqslant T\}} + X_T\mathbb{I}_{\{t>T\}}$$

is progressively measurable.

Definition 3.14 *A filtration $\mathbb{F}$ satisfies the usual conditions if $\mathbb{F}$ is right-continuous and for all $t \geqslant 0$ the sigma algebra $\mathcal{F}_t$ contains all $\mathbb{P}$-negligible sets (with respect to $\mathcal{F}$).*

Given a filtration $\mathbb{F}$, we construct an enlargement $\mathbb{F}^{+,\mathbb{P}}$ of $\mathbb{F}$ that satisfies the usual conditions as follows: Write $\mathcal{N}^{\mathbb{P}}$ for the $\mathbb{P}$-negligible sets (with respect to $\mathcal{F}$), and set

$$\mathcal{F}_t^{+,\mathbb{P}} = \sigma(\mathcal{F}_t^+, \mathcal{N}^{\mathbb{P}}), \qquad t \geqslant 0.$$

One can show that $C \subset \Omega$ is in $\mathcal{F}_t^{+,\mathbb{P}}$ if and only if there exists $A \in \mathcal{F}_t^+$ such that $A\Delta B = (A \setminus B) \cup (B \setminus A) \in \mathcal{N}^{\mathbb{P}}$. One can also show that $\mathbb{F}^{+,\mathbb{P}}$ is the smallest right-continuous sigma algebra containing $\mathbb{F}^{\mathbb{P}}$, where $\mathcal{F}_t^{\mathbb{P}} = \sigma(\mathcal{F}_t, \mathcal{N}^{\mathbb{P}})$ for $t \geqslant 0$.

Theorem 3.15 *("Debut theorem")*

Assume that $\mathbb{F}$ satisfies the usual conditions, let X be a measurable process with values in a measurable space $(E, \mathcal{E})$, and let $A \in \mathcal{E}$. Then the entrance time T_A is a stopping time.

Proposition 3.16 *Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$ and let T be a $\mathbb{F}^{+,\mathbb{P}}$-stopping time. Then there exists a $\mathbb{F}^+$-stopping time $T^{\mathbb{P}}$ such that the set $\{T \neq T^{\mathbb{P}}\}$ is $\mathbb{P}$-negligible.*

3.2 Applications to Brownian motion

Definition 3.17 *A d-dimensional Brownian motion is a stochastic process* $B = (B^1, \dots, B^d)$ *consisting of independent (1-dimensional) Brownian motions* B^j, $j = 1, \dots, d$.

Definition 3.18 *Let* $\mathbb{F}$ *be a filtration. A (d-dimensional)* $\mathbb{F}$*-Brownian motion is a continuous stochastic process that is adapted to* $\mathbb{F}$*, and such that for all* $s, t \geqslant 0$ *the vector* $B_{t+s} - B_t$ *is independent of* $\mathcal{F}_t$ *and has law* $\mathcal{N}(0, (t-s)I)$ *(where* I *denotes the identity matrix on* $\mathbb{R}^d$*).*

Theorem 3.19 *(Strong Markov property)*

Let $\mathbb{F}$ *be a filtration and let* B *be a d-dimensional* $\mathbb{F}$*-Brownian motion. Then for any finite stopping time* T *the process* $B^{(T)} = (B_{T+t} - B_T)_{t \geqslant 0}$ *is a d-dimensional Brownian motion and independent of* $\mathcal{F}_{T+}$.

Remark 3.20 Let T be a not necessarily finite stopping time with $\mathbb{P}(T < \infty) > 0$. Then the proof of Theorem 3.19 still shows that

$$\mathbb{E}[\mathbb{I}_{A \cap \{T < \infty\}} G(B^{(T)})] = \mathbb{P}(A \cap \{T < \infty\}) \mathbb{E}[G(B)]$$

for all $A \in \mathcal{F}_{T+}$. Dividing both sides by $\mathbb{P}(T < \infty)$, we deduce that

$$\mathbb{E}[\mathbb{I}_A G(B^{(T)}) | T < \infty] = \mathbb{P}(A | T < \infty) \mathbb{E}[G(B)],$$

which shows that under the conditional probability measure $\mathbb{P}(\cdot | T < \infty)$ the process $(B_{T+t} - B_T)_{t \geqslant 0}$ (defined as 0 on the set $T = \infty$) is a Brownian motion independent of $\mathcal{F}_{T+}$.

Corollary 3.21 *(Blumenthal's 0-1 law)*

Let B *be a (d-dimensional) Brownian motion and let* $A \in \mathcal{F}^B_{0+}$*. Then* $\mathbb{P}(A) = 0$.

Corollary 3.22 *Let* B *be a one-dimensional Brownian motion. Then with probability* 1 *we have for all* $\varepsilon > 0$

$$\sup_{s \in [0,\varepsilon]} B_s > 0, \qquad \inf_{s \in [0,\varepsilon]} B_s < 0.$$

Moreover, if for $a \in \mathbb{R}$ *we set* $T_a = \inf\{t \geqslant 0 : B_t = a\}$*, then* $T_a < \infty$ *for all* $a \in \mathbb{R}$ *with probability* 1*, so that in particular*

$$\liminf_{t \to \infty} B_t = -\infty < \infty = \limsup_{t \to \infty} B_t.$$

Remark 3.23 It is a priori not obvious whether $\sup_{s \in [0,\varepsilon]} B_s$ is measurable. To see this recall that B is continuous, and therefore $\sup_{s \in [0,\varepsilon]} B_s = \sup_{s \in [0,\varepsilon] \cap \mathbb{Q}} B_s$. In the following we will often implicitly use this kind of argument when dealing with continuous (or right- or left-continuous) processes.

Theorem 3.24 *Let* B *be a one-dimensional Brownian motion. Then almost surely the set*

$$\text{Zeros} = \{t \geqslant 0 : B_t = 0\}$$

is closed and has no isolated points (such a set is called perfect). Moreover, Zeros *has Lebesgue measure zero and it is unbounded (not contained in* $[0, n]$ *for any* $n \in \mathbb{N}$*).*

Exercise 3.25 Show that the complement of Zeros consist of countably many open intervals.

Exercise 3.26 Let B be a one-dimensional Brownian motion and write $S_t = \sup_{s\in[0,t]} B_s$. Show that for all $a \geqslant 0$ and $b \leqslant a$ we have

$$\mathbb{P}(S_t \geqslant a, B_t \leqslant b) = \mathbb{P}(B_t \geqslant 2a - b).$$

Deduce that

$$\mathbb{P}(S_t \geqslant a) = 2\mathbb{P}(B_t \geqslant a) = \mathbb{P}(|B_t| \geqslant a),$$

so in particular S_t has the same distribution as $|B_t|$.

Hint: Consider the stopping time $T_a = \inf\{t \geqslant 0 : B_t = a\}$ and apply the strong Markov property.

4 Continuous time martingales

4.1 Path regularity

Throughout this section we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathbb{F}$. A stochastic process X is called *integrable* if $\mathbb{E}[|X_t|] < \infty$ for all $t \geqslant 0$. If $p > 0$, then we call X p-integrable if $\mathbb{E}[|X_t|^p] < \infty$ for all $t \geqslant 0$.

Definition 4.1 *An adapted real-valued and integrable process $X = (X_t)_{t\geqslant 0}$ is called a*

i. martingale if $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ for all $0 \leqslant s \leqslant t$;

ii. supermartingale if $\mathbb{E}[X_t|\mathcal{F}_s] \leqslant X_s$ for all $0 \leqslant s \leqslant t$;

iii. submartingale if $\mathbb{E}[X_t|\mathcal{F}_s] \geqslant X_s$ for all $0 \leqslant s \leqslant t$.

If X is a submartingale, then obviously $-X$ is a supermartingale. For that reason we state some of the following results only for submartingales, although they also hold for supermartingales.

Example 4.2 Let B be a d-dimensional $\mathbb{F}$-Brownian motion. Then

i. All components B^j are martingales for $j = 1, \dots, d$:

$$\mathbb{E}[B_t^j|\mathcal{F}_s] = \mathbb{E}[B_t^j - B_s^j|\mathcal{F}_s] + B_s^j = \mathbb{E}[B_t^j - B_s^j] + B_s^j = 0.$$

ii. $X_t = (B_t^j)^2 - t$, $t \geqslant 0$, is a martingale:

$$\mathbb{E}[X_t|\mathcal{F}_s] = \mathbb{E}[(B_t^j - B_s^j)^2 + 2(B_t^j - B_s^j)B_s^j + (B_s^j)^2|\mathcal{F}_s] - t = (t-s) + (B_s^j)^2 - t = X_s.$$

iii. $Y_t = B_t^i B_t^j$, $t \geqslant 0$, is a martingale if $i \neq j$:

$$\mathbb{E}[Y_t|\mathcal{F}_s] = \mathbb{E}[(B_t^i - B_s^i)(B_t^j - B_s^j) + (B_t^i - B_s^i)B_s^j + B_s^i(B_t^j - B_s^j) + B_s^i B_s^j|\mathcal{F}_s] = Y_s.$$

iv. $Z_t = e^{\lambda B_t^j - \lambda^2 t/2}$, $t \geqslant 0$, is a martingale:

$$\mathbb{E}[Z_t|\mathcal{F}_s] = \mathbb{E}[e^{\lambda(B_t^j - B_s^j)}|\mathcal{F}_s]e^{\lambda B_s^j - \lambda^2 t/2} \tag{20}$$

$$= \mathbb{E}[e^{\lambda(B_t^j - B_s^j)}]e^{\lambda B_s^j - \lambda^2 t/2}e^{\lambda^2(t-s)/2}e^{\lambda B_s^j - \lambda^2 t/2} = Z_s, \tag{21}$$

by the formula $\mathbb{E}[e^{\lambda U}] = e^{\lambda^2\sigma^2/2}$ for the Laplace transform of a $\mathcal{N}(0, \sigma^2)$ variable U.

Example 4.3 Let ξ be a white noise and $f \in L^2(\mathbb{R}_+)$ and define $X_t = \xi(\mathbb{I}_{[0,t]} f)$, $t \geqslant 0$. Then X is a martingale in the filtration $\mathcal{F}_t = \sigma(X_s : s \leqslant t)$: By orthogonality and Gaussianity we get that $\xi(\mathbb{I}_{[s,t]} f)$ is independent of $\mathcal{F}_s$, and therefore

$$\mathbb{E}[X_t|\mathcal{F}_s] = \mathbb{E}[\xi(\mathbb{I}_{[0,s]} f) + \xi(\mathbb{I}_{[s,t]} f)|\mathcal{F}_s] = X_s + \mathbb{E}[\xi(\mathbb{I}_{[s,t]} f)] = X_s.$$

Example 4.4 Let N be a process with independent increments, such that $N_0 = 0$ and for all $0 \leqslant s \leqslant t$ the increment $N_t - N_s$ is Poisson distributed with parameter $\lambda(t-s)$ for a fixed $\lambda > 0$. Then $(N_t - \lambda t)_{t \geqslant 0}$ is a martingale in the filtration $(\mathcal{F}_t)_{t \geqslant 0}$ generated by N:

$$\mathbb{E}[N_t - \lambda t|\mathcal{F}_s] = \mathbb{E}[N_t - N_s|\mathcal{F}_s] + N_s - \lambda t = \lambda(t-s) + N_s - \lambda t = N_s - \lambda s.$$

The process N is called *Poisson process*. One can show that it has a modification which is an increasing right-continuous step function with jumps of size 1.

Remark 4.5 If X is a martingale and f is convex and such that $\mathbb{E}[|f(X_t)|] < \infty$ for all $t \geqslant 0$, then $f(X)$ is a submartingale. If X is a submartingale and f is convex and increasing and such that $\mathbb{E}[|f(X_t)|] < \infty$ for all $t \geqslant 0$, then $f(X)$ is a submartingale. Both of these statements follow by a simple application of Jensen's inequality. In particular, $|X|^p$ is a submartingale if X is a p-integrable martingale and $p \geqslant 1$, and X^+ is a submartingale if X is a submartingale.

Definition 4.6 *Let I be an index set and $f : I \to \mathbb{R}$. For $a < b$, the number of downcrossings of f across the interval $[a,b]$ in I is the supremum over all n for which there exist times $s_k, t_k \in I$, $k = 1, \ldots, n$, such that $s_1 < t_1 < s_2 < t_2 < \ldots < s_n < t_n$ with $f(s_k) \geqslant b$ and $f(t_k) \leqslant a$ for all $k = 1, \ldots, n$. We denote it with*

$$D([a,b]; I; f).$$

Lemma 4.7 *(Doob's downcrossing lemma)*

Let X be a submartingale defined on a finite index set $I \subset \mathbb{R}_+$ and let $T = \max\{t : t \in I\}$. Then for all $a < b$

$$\mathbb{E}[D([a,b]; I; X)] \leqslant \frac{1}{b-a} \mathbb{E}[(X_T - b)^+].$$

Let us introduce the notation

$$\lim_{s \downarrow\downarrow t} f(s) = \lim_{\substack{s \to t \\ s > t}} f(s), \qquad \lim_{s \uparrow\uparrow t} f(s) = \lim_{\substack{s \to t \\ s < t}} f(s)$$

and similarly $\limsup_{s \downarrow\downarrow t} f(s)$, $\liminf_{s \uparrow\uparrow t} f(s)$, etc.

Definition 4.8 *A function $f : \mathbb{R}_+ \to \mathbb{R}$ is called càdlàg if it is right-continuous and at every $t > 0$ the limit $\lim_{s \uparrow\uparrow t} f(s)$ exists (but it is not necessarily equal to $f(t)$).*

The acronym càdlàg comes from French and stands for "continue à droite, limite à gauche", that is "continuous from the right, limits from the left".

Exercise 4.9 Let $f : \mathbb{Q}_+ \to \mathbb{R}$, where $\mathbb{Q}_+ = \mathbb{Q} \cap \mathbb{R}_+$, be such that for all $k \in \mathbb{N}$ and for all $a, b \in \mathbb{Q}$ with $a < b$ we have

$$D([a,b]; [0,k] \cap \mathbb{Q}_+; f) < \infty \qquad \text{and} \qquad \sup_{s \in [0,k] \cap \mathbb{Q}_+} |f(s)| < \infty.$$

Then the limit from the right

$$f(t+) = \lim_{s \downarrow\downarrow t, s \in \mathbb{Q}_+} f(s)$$

exists for all $t \geqslant 0$ and the limit from the left

$$f(t-) = \lim_{s \uparrow\uparrow t, s \in \mathbb{Q}_+} f(s)$$

exists for all $t > 0$. Moreover, the function $t \mapsto f(t+)$, $t \geqslant 0$, is càdlàg.

Lemma 4.10 *Let X be a submartingale defined on a finite index set $I \subset \mathbb{R}_+$ and let $T = \max\{t : t \in I\}$. Then for all $\lambda > 0$*

$$\lambda \mathbb{P}(\sup_{t \in I} |X_t| \geqslant \lambda) \leqslant 2\mathbb{E}[X_T^+] - \mathbb{E}[X_0].$$

Lemma 4.11 *Let X be a submartingale. Then there exists a null set N such that for all $\omega \in \Omega \setminus N$ the function $(X_r(\omega))_{r \in \mathbb{Q}_+}$ has a right limit $X_{t+}(\omega)$ at every $t \geqslant 0$ and a left limit $X_{t-}(\omega)$ at every $t > 0$, and the function $t \mapsto X_{t+}(\omega)$ is càdlàg.*

Lemma 4.12 *Let X be a submartingale and $(t_n)_{n \in \mathbb{N}}$ a decreasing sequence of positive numbers. Then the family $(X_{t_n})_{n \in \mathbb{N}}$ is uniformly integrable.*

Definition 4.13 *Let $(\mathcal{G}_n)_{n \in -\mathbb{N}_0}$ be a family of sigma algebras with $\mathcal{G}_{n-1} \subset \mathcal{G}_n$ for all $n \in -\mathbb{N}_0 = \{0, -1, -2, \ldots\}$. An integrable process $(Y_n)_{n \in -\mathbb{N}_0}$ which is adapted to $(\mathcal{G}_n)$ is called a*

i. backward martingale if $\mathbb{E}[Y_n | \mathcal{G}_{n-1}] = Y_{n-1}$ for all $n \in -\mathbb{N}_0$;

ii. backward submartingale if $\mathbb{E}[Y_n | \mathcal{G}_{n-1}] \geqslant Y_{n-1}$ for all $n \in -\mathbb{N}_0$.

Lemma 4.14 *Let $(Y_n)_{n \in -\mathbb{N}_0}$ be a backward submartingale with $\lim_{n \to -\infty} \mathbb{E}[Y_n] > -\infty$. Then (Y_n) is uniformly integrable.*

Theorem 4.15 *Let X be a submartingale in a right-continuous filtration $\mathbb{F}$. We define*

$$Y_t(\omega) = \begin{cases} X_{t+}(\omega), & \lim_{s \downarrow\downarrow t, s \in \mathbb{Q}_+} X_s(\omega) \text{ exists}, \\ 0, & \text{else}. \end{cases}$$

Then

i. Y is a submartingale and almost surely càdlàg;

ii. there exists a modification $\tilde{Y}$ of Y which is adapted to $\mathbb{F}^{\mathbb{P}}$ and for which all trajectories are càdlàg; moreover $\tilde{Y}$ is a $\mathbb{F}^{\mathbb{P}}$-submartingale;

iii. for all $t \geqslant 0$ we have $X_t \leqslant Y_t$ almost surely;

iv. Y is a modification of X if and only if the map $t \mapsto \mathbb{E}[X_t]$ is right-continuous (in particular if X is a martingale because then $\mathbb{E}[X_t] = \mathbb{E}[X_0]$ is constant in t).

4.2 The stopping theorem

Proposition 4.16 *(Doob)*

i. If X is a right-continuous submartingale, then we have for all $T \geqslant 0$ and $\lambda > 0$

$$\mathbb{P}(\sup_{t\in[0,T]} X_t \geqslant \lambda) \leqslant \frac{1}{\lambda}\mathbb{E}[X_T^+], \qquad \mathbb{P}(\sup_{t\geqslant 0} X_t \geqslant \lambda) \leqslant \frac{1}{\lambda}\sup_{t\geqslant 0}\mathbb{E}[X_t^+].$$

ii. If X is a right-continuous martingale, then we have for all $T \geqslant 0$ and $\lambda > 0$

$$\mathbb{P}(\sup_{t\in[0,T]} |X_t| \geqslant \lambda) \leqslant \frac{1}{\lambda}\mathbb{E}[|X_T|], \qquad \mathbb{P}(\sup_{t\geqslant 0} |X_t| \geqslant \lambda) \leqslant \frac{1}{\lambda}\sup_{t\geqslant 0}\mathbb{E}[|X_t|],$$

and for all $p > 1$

$$\mathbb{E}[\sup_{t\in[0,T]} |X_t|^p] \leqslant \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_T|^p], \qquad \mathbb{E}[\sup_{t\geqslant 0} |X_t|^p] \leqslant \left(\frac{p}{p-1}\right)^p \sup_{t\geqslant 0}\mathbb{E}[|X_t|^p].$$

Proposition 4.17 *Let X be a right-continuous submartingale with $\sup_{t\geqslant 0}\mathbb{E}[X_t^+] < \infty$. Then there exists a random variable $X_\infty \in L^1$ with $\lim_{t\to\infty} X_t = X_\infty$ almost surely.*

Example 4.18 Without having a condition like $\sup_{t\geqslant 0}\mathbb{E}[X_t^+] < \infty$ convergence can fail: For example the Brownian motion does not converge because

$$-\infty = \liminf_{t\to\infty} B_t < \limsup_{t\to\infty} B_t = \infty.$$

If X is a martingale that converges, then we do not have in general $\mathbb{E}[X_\infty] = \mathbb{E}[X_0]$. Consider for example $X_t = \exp(B_t - t/2)$, which is a positive martingale and therefore almost surely converges. But we know from Corollary **??** that given $\alpha \in (1/2, 1)$, for almost every $\omega \in \Omega$ there exists $C(\omega) > 0$ with $|B_t(\omega)| \leqslant C(\omega)t^\alpha$ for all $t \geqslant 0$, so

$$\lim_{t\to\infty} X_t \leqslant \limsup_{t\to\infty} \exp(C(\omega)t^\alpha - t/2) = 0,$$

and of course $0 \neq 1 = \mathbb{E}[X_t]$ for all $t \geqslant 0$.

Proposition 4.19 *For a right-continuous martingale X the following conditions are equivalent:*

a) The family $(X_t)_{t\geqslant 0}$ is uniformly integrable (we say X is a uniformly integrable martingale);

b) X_t converges almost surely and in L^1 to a limit X_∞ as $t \to \infty$;

c) there exists $Y \in L^1$ with $X_t = \mathbb{E}[Y|\mathcal{F}_t]$ for all $t \geqslant 0$.

In that case we can always take $Y = X_\infty$.

Corollary 4.20 *Let X be a right-continuous martingale and $p > 1$ such that*

$$\sup_{t\geqslant 0}\mathbb{E}[|X_t|^p] < \infty.$$

The we call X bounded in L^p, and X is uniformly integrable and

$$\mathbb{E}[\sup_{t\geqslant 0} |X_t|^p] \leqslant \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_\infty|^p] = \left(\frac{p}{p-1}\right)^p \sup_{t\geqslant 0}\mathbb{E}[|X_t|^p] < \infty.$$

If T is a stopping time and X a right-continuous submartingale with $\sup_{t\geqslant 0}\mathbb{E}[X_t^+]<\infty$ (which is in particular the case if X is a uniformly integrable martingale), then we define

$$X_T(\omega) := \mathbb{I}_{\{T(\omega)<\infty\}}X_{T(\omega)}(\omega) + \mathbb{I}_{\{T(\omega)=\infty\}}X_\infty(\omega).$$

Theorem 4.21 *(Stopping theorem for u.i. martingales)*
Let X be a right-continuous uniformly integrable martingale and let S and T be stopping times with $S\leqslant T$. Then X_S and X_T are in L^1 and we have

$$\mathbb{E}[X_T|\mathcal{F}_S] = X_S.$$

Corollary 4.22 *(Stopping theorem with bounded stopping times)*
Let X be a martingale and let S,T be stopping times with $S\leqslant T\leqslant K$ for some $K\in\mathbb{R}$. Then $X_S, X_T\in L^1$ and

$$\mathbb{E}[X_T|\mathcal{F}_S] = X_S.$$

Corollary 4.23 *Let X be a right-continuous martingale and T be a stopping time. Then the stopped process*

$$X_t^T = X_{t\wedge T}, \qquad t\geqslant 0,$$

is a martingale. If X is uniformly integrable, then X^T is as well and we have

$$X_t^T = X_{t\wedge T} = \mathbb{E}[X_T|\mathcal{F}_t], \qquad t\geqslant 0.$$

Theorem 4.24 *(Stopping theorem for positive supermartingales)*
Let X be a right-continuous positive supermartingale. Then

i. there exists a random variable $X_\infty\in L^1$ with $\lim_{t\to\infty}X_t = X_\infty$ almost surely;

ii. if $S\leqslant T$ are stopping times, then

$$\mathbb{E}[X_T|\mathcal{F}_S]\leqslant X_S;$$

iii. the process $X_t^T = X_{t\wedge T}$, $t\geqslant 0$, is a supermartingale.

Exercise 4.25 Let X be a positive supermartingale. Show that X is a uniformly integrable martingale if and only if $\mathbb{E}[X_\infty]=\mathbb{E}[X_0]$.

Proposition 4.26 *Let B be a Brownian motion and write $T_x = \inf\{t\geqslant 0 : B_t = x\}$ for $x\in\mathbb{R}$. Let $a,b>0$. Then*

$$\mathbb{P}(T_{-a}<T_b) = \frac{b}{a+b}, \qquad \mathbb{P}(T_{-a}>T_b) = \frac{a}{a+b}.$$

5 Continuous semimartingales

Throughout this section we assume that our filtration $\mathbb{F}$ satisfies the usual conditions. By being careful, it is possible to develop the entire theory that follows without this assumption, but doing so would complicate the presentation and as the material is already technical enough as it is, we prefer to work under the usual conditions..

5.1 Processes of finite variation

Definition 5.1 *Let $0 \leqslant s < t < \infty$. A continuous function $a : [s,t] \to \mathbb{R}$ is of bounded variation on $[s,t]$ if there exist two increasing functions $a(+)$ and $a(-)$ such that $a = a(+) - a(-)$. A continuous function $a : \mathbb{R}_+ \to \mathbb{R}$ is of finite variation if $a(0) = 0$ and $a|_{[0,T]}$ is of bounded variation whenever $0 \leqslant T < \infty$.*

If $a : [0,T] \to \mathbb{R}$ with $a(0) = 0$ is of bounded variation, then $a(+)$ and $a(-)$ are of course not unique: we can add any increasing function to both of them. But there is a minimal decomposition that is unique: First note that $a(+)$ and $a(-)$ are the "distribution functions" of two measures, determined via

$$\mu_+([0,t]) = a(+)(t) \qquad \text{and} \qquad \mu_-([0,t]) = a(-)(t).$$

So $\mu = \mu_+ - \mu_-$ is a signed measure on $([0,T], \mathcal{B}([0,T]))$, i.e. a set function $\mu : \mathcal{B}([0,T]) \to \mathbb{R}$ such that $\mu(\emptyset) = 0$ and $\mu(\cup_n A_n) = \sum_n \mu(A_n)$ with absolutely converging series on the right hand side whenever $(A_n)_n$ is a sequence of disjoint sets in $\mathcal{B}([0,T])$. Therefore, there exists a unique Jordan decomposition of μ, denoted by abuse of notation again by $\mu = \mu_+ - \mu_-$, where μ_+ and μ_- are still positive measures, but now they are also mutually singular (there exists $B_+ \in \mathcal{B}([0,T])$ with $\mu_+ = \mu_+(\cdot \cap B_+)$ and $\mu_- = \mu_-(\cdot \cap B_+^c)$).

Definition 5.2 *Let a be of finite variation and let μ be the signed measure associated to it. The total variation of μ is the measure*

$$|\mu| = \mu_+ + \mu_-,$$

where (μ_+, μ_-) is the Jordan decomposition of μ. We also call

$$V(a)(t) = \mu_+([0,t]) + \mu_-([0,t])$$

the total variation of a, and we write $a(+)(t) = \mu_+([0,t])$ and $a(-)(t) = \mu_-([0,t])$. Then $a(+)$ and $a(-)$ are increasing functions with $a = a(+) - a(-)$, and if f and g are increasing functions with $a = f - g$, then $a(+) \leqslant f$ and $a(-) \leqslant g$.

Note that

$$|\mu(B)| = |\mu_+(B) - \mu_-(B)| \leqslant \mu_+(B) + \mu_-(B) = |\mu|(B)$$

for all $B \in \mathcal{B}([0,T])$. Since a is continuous, also $V(a)$ is continuous because $\mu(\{t\}) = \mu_+(\{t\}) - \mu_-(\{t\}) = 0$ for all $t \geqslant 0$, which by the mutual singularity of μ_+ and μ_- yields $\mu_+(\{t\}) = \mu_-(\{t\}) = 0$ for all $t \geqslant 0$. Similarly we get $V(a)(0) = 0$.

Proposition 5.3 *Let a be of finite variation. Then we have for all $t \geqslant 0$*

$$V(a)(t) = \sup \left\{ \sum_{j=0}^{n-1} |a(t_{j+1}) - a(t_j)| : n \in \mathbb{N}, 0 = t_0 < \ldots < t_n = t \right\} = \lim_{n \to \infty} V^n(a)(t),$$

where

$$V^n(a)(t) = \sum_{k:(k+1)/n \leqslant t} |a((k+1)/n) - a(k/n)|.$$

Definition 5.4 *If $h : \mathbb{R}_+ \to \mathbb{R}$ is measurable and satisfies $\int_0^T |h(t)||\mu|(\mathrm{d}t) < \infty$ for all $T \geqslant 0$, then we set*

$$\int_0^t h(s)\mathrm{d}a(s) := \int_{\mathbb{R}_+} \mathbb{I}_{[0,t]}(s)h(s)\mu(\mathrm{d}s) = \int_{\mathbb{R}_+} \mathbb{I}_{[0,t]}(s)h(s)\mu_+(\mathrm{d}s) - \int_{\mathbb{R}_+} \mathbb{I}_{[0,t]}(s)h(s)\mu_-(\mathrm{d}s),$$

$t \geqslant 0$, and

$$\int_0^t h(s)\mathrm{d}V(a)(s) := \int_{\mathbb{R}_+} \mathbb{I}_{[0,t]}(s)h(s)|\mu|(\mathrm{d}s).$$

Both $\int_0^\cdot h(s)\mathrm{d}a(s)$ and $\int_0^\cdot h(s)\mathrm{d}V(a)(s)$ are of finite variation and the associated measures are $h(s)\mu(\mathrm{d}s)$ and $h(s)|\mu|(\mathrm{d}s)$.

Example 5.5 If $a \in C^1(\mathbb{R}_+, \mathbb{R})$, then a is of bounded variation and we have for $t \geqslant 0$

$$\int_0^t h(s)\mathrm{d}a(s) = \int_0^t h(s)a'(s)\mathrm{d}s, \qquad \int_0^t h(s)\mathrm{d}V(a)(s) = \int_0^t h(s)|a'(s)|\mathrm{d}s.$$

Indeed it suffices to note that the measures $a'(s)\mathrm{d}s$ and μ assign the same value $\int_u^v a'(s)\mathrm{d}s$ to any interval $(u, v] \subset \mathbb{R}_+$.

Lemma 5.6 *Let a be of bounded variation on $[0, T]$, let $h : [0, T] \to \mathbb{R}$ be left-continuous and bounded and let $0 = t_0^n < \ldots < t_{N^n}^n = T$ be a sequence of partitions of $[0, T]$ with $\max_{k<N_n} |t_{k+1}^n - t_k^n| \to 0$ for $n \to \infty$. Then*

$$\int_0^T h(t)\mathrm{d}a(t) = \lim_{n\to\infty} \sum_{k=0}^{N^n-1} h(t_j)(a(t_{j+1}) - a(t_j)).$$

Definition 5.7 *A stochastic process $A = (A_t)_{t\geqslant 0}$ is a process of finite variation if it is adapted and the function $A(\omega)$ is of finite variation for all $\omega \in \Omega$. In that case we write $A \in \mathcal{A}$. If furthermore $A(\omega)$ is increasing for all $\omega \in \Omega$, we write $A \in \mathcal{A}^+$.*

Proposition 5.8 *Let $A \in \mathcal{A}$ and let H be a progressively measurable process such that for all $\omega \in \Omega$ and all $t \geqslant 0$*

$$\int_0^t |H_s|(\omega)\mathrm{d}V(A)_s(\omega) < \infty.$$

Then

$$(H \cdot A)_t(\omega) := \int_0^t H_s(\omega)\mathrm{d}A_s(\omega)$$

defines a process of finite variation.

Remark 5.9

i. We will often only have the condition $\int_0^t |H_s|\mathrm{d}V(A)_s < \infty$ satisfied outside a null set. In that case we set

$$\begin{cases} \int_0^t H_s(\omega)\mathrm{d}A_s(\omega), & \int_0^T |H_s|(\omega)\mathrm{d}V(A)_s(\omega) < \infty \text{ for all } T \geqslant 0, \\ 0, & \text{else,} \end{cases}$$

which still defines a finite variation process because we assumed our filtration $\mathbb{F}$ to satisfy the usual conditions, so altering a process on a null set does not change its adaptedness properties.

ii. If almost surely $\int_0^t |H_s| \mathrm{d}V(A)_s < \infty$ and $\int_0^t |G_s H_s| \mathrm{d}V(A)_s < \infty$, then we have

$$G \cdot (H \cdot A) = (GH) \cdot A,$$

which holds because $(H \cdot A)$ is a finite variation process associated to the measure $H_s \mathrm{d}\mu$, where μ is the measure associated to A.

5.2 Local martingales and their quadratic variation

Definition 5.10 *An adapted continuous process M is called a local martingale if there exists an increasing sequence of stopping times (T_n) with $\lim_{n\to\infty} T_n = \infty$ almost surely, such that all of the stopped processes M^{T_n} are martingales. In that case we write $M \in \mathcal{M}_{\mathrm{loc}}$ and call (T_n) a localizing sequence for M.*

Any increasing sequence of stopping times (S_n) with $\lim_{n\to} S_n = \infty$ almost surely is called a localizing sequence. We also write $\mathcal{M}$ for the set of all continuous uniformly integrable martingales.

Remark 5.11

i. We do not require that local martingales are integrable.

ii. Every continuous martingale is a local martingale. It suffices to take $T_n \equiv \infty$.

iii. If $M \in \mathcal{M}_{\mathrm{loc}}$ (respectively $M \in \mathcal{M}$) and T is a stopping time, then $M^T \in \mathcal{M}_{\mathrm{loc}}$ (respectively $M \in \mathcal{M}$): apply the stopping theorem and note that $(M^T)^{T_n} = (M^{T_n})^T$.

iv. If $M \in \mathcal{M}_{\mathrm{loc}}$ is localized by (T_n) and (S_n) is a sequence of stopping times with $S_n \leqslant T_n$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} S_n = \infty$ almost surely, then (S_n) is also a localizing sequence for M (apply the stopping theorem).

v. For every $M \in \mathcal{M}_{\mathrm{loc}}$ there exists a localizing sequence of stopping times (T_n) such that $M^{T_n} \in \mathcal{M}$ for all $n \in \mathbb{N}$. (Take $T_n = S_n \wedge n$ for a localizing sequence (S_n) for M).

Proposition 5.12 *Let $M \in \mathcal{M}_{\mathrm{loc}}$.*

i. If M is positive, then it is a supermartingale.

ii. If there exists $Z \in L^1$ with $M_t \leqslant Z$ for all $t \geqslant 0$, then $M \in \mathcal{M}$.

iii. If $M_0 = 0$, then $T_n = \inf\{t \geqslant 0 : |M_t| \geqslant n\}$, $n \in \mathbb{N}$, is a localizing sequence for M.

Remark 5.13 Given point ii. one might be lead to believe that a local martingale M for which $(M_t)_{t\geqslant 0}$ is uniformly integrable is always a martingale. This is false, as Exercise ... will show!

Lemma 5.14 *Let $M \in \mathcal{A} \cap \mathcal{M}_{\mathrm{loc}}$. Then almost surely $M_t = 0$ for all $t \geqslant 0$.*

Remark 5.15 This lemma shows that any nontrivial *continuous* martingale is almost surely of infinite variation. For discontinuous martingales this is not at all true: Recall for example the compensated Poisson process of Example 4.4.

Theorem 5.16 *Let $M \in \mathcal{M}_{\text{loc}}$. Then there exists an increasing process $\langle M, M\rangle \in \mathcal{A}^+$, unique up to indistinguishability, such that $M^2 - M_0^2 - \langle M, M\rangle \in \mathcal{M}_{\text{loc}}$. We call $\langle M, M\rangle$ the quadratic variation of M.*

Remark 5.17 We will later see that if $(T_k^n : n, k \in \mathbb{N}_0)$ is a family of stopping times with $0 = T_0^n < T_1^n < T_2^n < \dots$ and $\lim_{k\to\infty} T_k^n = \infty$ for all n and such that for all $\varepsilon, K > 0$ we have

$$\lim_{n\to\infty} \mathbb{P}(\exists k : |T_{k+1}^n - T_k^n| > \varepsilon, T_{k+1}^n < K) = 0,$$

then

$$\langle M, M\rangle_t = \lim_{n\to\infty} \sum_{k=0}^{\infty} (M_{T_{k+1}^n \wedge t} - M_{T_k^n \wedge t})^2, \tag{22}$$

where the convergence is in probability, uniformly in t on any compact set. A typical example of stopping times T_k^n that satisfy these assumptions are the deterministic times $T_k^n = k/n$.

Lemma 5.18 *Let $M \in \mathcal{M}$ with $\mathbb{E}[M_\infty^2] < \infty$, let $(T_k)_{k\in\mathbb{N}}$ be an increasing sequence of stopping times with $\lim_{k\to\infty} T_k = \infty$, and let $C > 0$ be such that for every k the random variable H_{T_k} is $\mathcal{F}_{T_k}$-measurable and satisfies $|H_{T_k}| \leqslant C$. Then for*

$$H_t = \sum_{k=0}^{\infty} H_{T_k} \mathbb{I}_{[T_k, T_{k+1})}(t), \qquad t \geqslant 0,$$

we define

$$(H \cdot M)_t := \sum_{k=0}^{\infty} H_{T_k} (M_{T_{k+1}\wedge t} - M_{T_k \wedge t}), \qquad t \geqslant 0,$$

which is a martingale in $\mathcal{M}$ that satisfies

$$\mathbb{E}[\sup_{t\geqslant 0} |(H \cdot M)_t|^2] \leqslant 4C^2 (\mathbb{E}[M_\infty^2] - \mathbb{E}[M_0^2]).$$

Remark 5.19 In the proof we showed that whenever T is a stopping time, then

$$\langle M^T, M^T\rangle = \langle M, M\rangle^T$$

by the uniqueness of the quadratic variation. Moreover, we have

$$\langle M, M\rangle = \langle M - M_0, M - M_0\rangle.$$

Theorem 5.20 *Let $M \in \mathcal{M}_{\text{loc}}$ with $M_0 = 0$.*

i. The following conditions are equivalent:

a) $M \in \mathcal{M}$ and M is L^2-bounded;

b) $\mathbb{E}[\langle M, M\rangle_\infty] < \infty$.

In that case $M^2 - \langle M, M\rangle \in \mathcal{M}$ and in particular $\mathbb{E}[M_\infty^2] = \mathbb{E}[\langle M, M\rangle_\infty]$.

ii. The following conditions are equivalent:

a) M is a martingale and $\mathbb{E}[M_t^2] < \infty$ for all $t \geqslant 0$;

b) $\mathbb{E}[\langle M, M\rangle_t] < \infty$ *for all* $t \geqslant 0$.

Proposition 5.21 *Let* $M \in \mathcal{M}_{\text{loc}}$ *be such that* $\langle M, M\rangle$ *is indistinguishable from* 0. *Then* $M_t = M_0$ *almost surely for all* $t \geqslant 0$.

Definition 5.22 *Let* $M, N \in \mathcal{M}_{\text{loc}}$. *Then we define*

$$\langle M, N\rangle = \frac{1}{4}(\langle M+N, M+N\rangle - \langle M-N, M-N\rangle) \in \mathcal{A}$$

and call $\langle M, N\rangle$ *the quadratic covariation of* M *and* N.

The definition of $\langle M, N\rangle$ is in analogy to the formula

$$< x, y >= \frac{1}{4}(< x+y, x+y > - < x-y, x-y >) = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$$

which links the inner product and the norm on a Hilbert space. Note that once we proved the convergence (22) for all local martingales, it will follow directly that

$$\langle M, N\rangle_t = \lim_{n\to\infty} \sum_{k=0}^{\infty} (M_{T^n_{k+1}\wedge t} - M_{T^n_k \wedge t})(N_{T^n_{k+1}\wedge t} - N_{T^n_k\wedge t}).$$

Proposition 5.23 *Let* $M, N \in \mathcal{M}_{\text{loc}}$.

i. $\langle M, N\rangle$ *is the unique (up to indistinguishability) process in* $\mathcal{A}$ *for such that* $MN - M_0N_0 - \langle M, N\rangle$ *is a local martingale.*

ii. *The map* $(M, N) \mapsto \langle M, N\rangle$ *is bilinear and symmetric.*

iii. *If* T *is a stopping time we have* $\langle M, N\rangle^T = \langle M^T, N^T\rangle = \langle M^T, N\rangle$.

iv. $\langle M, N\rangle = \langle M - M_0, N - N_0\rangle$.

v. *If* M *and* N *are* L^2*-bounded martingales, then* $MN - M_0N_0 - \langle M, N\rangle \in \mathcal{M}$, *and in particular* $\langle M, N\rangle_\infty = \lim_{t\to\infty}\langle M, N\rangle_t$ *exists and satisfies*

$$\mathbb{E}[\langle M, N\rangle_\infty] = \mathbb{E}[M_\infty N_\infty] - \mathbb{E}[M_0N_0].$$

Example 5.24 If B^1 and B^2 are independent Brownian motions we have seen that B^1B^2 is a martingale, so we must have $\langle B^1, B^2\rangle \equiv 0$.

6 Stochastic integration

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space, where $(\mathcal{F}_t)_{t\geq 0}$ is a filtration satisfying the usual conditions, and let B be an $(\mathcal{F}_t)$-Brownian motion. We want to define the integral $\int_0^t \phi_s dB_s$ for a suitable class of processes ϕ. The Brownian motion is not of finite variation, we cannot define it path by path.

6.1 Integration against Brownian motion

Before defining the stochastic integral against Brownian motion, we need to introduce the stochastic integral for elementary processes.

6.1.1 Stochastic integration for elementary processes

Definition 6.1 *An elementary process is a process of the form* $\phi_t := \sum_{i=0}^{n-1} u_i \mathbb{1}_{(t_i, t_{i+1}]}(t)$, $t \geq 0$, *where* $n \in \mathbb{N}$, u_i *is a bounded and* $\mathcal{F}_{t_i}$*-measurable random varibale and* $0 \leq t_0 < t_1 < \ldots < t_n$.
We denote by $\mathcal{E}$ *the set of elementary processes. For* $\phi \in \mathcal{E}$ *and* $0 < t \leq \infty$, *we define the stochastic integral of* ϕ *against* B *by:*

$$\int_0^t \phi_s dB_s := \sum_{i=0}^{n-1} u_i (B_{t_{i+1} \wedge t} - B_{t_i \wedge t}).$$

$\mathcal{H}_c^2$ denotes the set of L^2-bounded continuous martingales M such that $M_0 = 0$. The following proposition gives important properties of the stochastic integral defined above.

Proposition 6.2 *Let* $\phi \in \mathcal{E}$. *Then*

$$M_t := \int_0^t \phi_s dB_s, t \geq 0, \text{ belongs to } \mathcal{H}_c^2 \text{ and for all } t \geq 0, \langle M, M \rangle_t = \int_0^t \phi_s^2 ds.$$

In particular,

$$\mathbb{E}\Big[\int_0^\infty \phi_s dB_s\Big] = 0 \text{ and } \mathbb{E}\Big[(\int_0^\infty \phi_s dB_s)^2\Big] = \mathbb{E}\Big[\int_0^\infty \phi_s^2 ds\Big].$$

6.1.2 Stochastic integration against Brownian motion

We introduce $\Lambda^2 := L^2(\mathbb{R}_+ \times \Omega, Prog, \lambda \otimes \mathbb{P})$ as the set of progressively measurable real valued processes $(\phi_t)_{t \geq 0}$ such that $\|\phi\|_{\Lambda^2} := \Big(\mathbb{E}\Big[\int_0^\infty \phi_s^2 ds\Big]\Big)^{1/2} < \infty$. Note that $(\Lambda^2, \|.\|_{\Lambda^2})$ is a Hilbert space. Since $\mathcal{E} \subset \Lambda^2$, the map

$$I : \phi \longmapsto I(\phi) := \int_0^\infty \phi_s dB_s$$

is a linear isometry on $\mathcal{E}$, then it is uniquely extended to a linear isometry on $\bar{\mathcal{E}}$. In the next propositon, we identify $\bar{\mathcal{E}}$.

Proposition 6.3 $\mathcal{E}$ *is dense in* $(\Lambda^2, \|.\|_{\Lambda^2})$.

Consequently, the linear isometry I is extended to a linear isometry on Λ^2, denoted again I. This is the aim of the following theorem.

Theorem 6.4 *There exists a unique linear map* $I : \Lambda^2 \longmapsto L^2(\Omega, \mathcal{F}, \mathbb{P})$ *such that:*

i) For all $\phi \in \Lambda^2$, $\|I(\phi)\|_{L^2(\Omega)} = \|\phi\|_{\Lambda^2}$.

ii) Let $s \leq t$. *If* $\phi := u\mathbb{1}_{(s,t]}$, *where* u *is an* $\mathcal{F}_s$*-measurable random variable, then*

$$I(\phi) = u(B_t - B_s).$$

In addition, $\mathbb{E}[I(\phi)] = 0$.

We write $I(\phi) = \int_0^\infty \phi_s dB_s$ and we call $I(\phi)$ the stochastic integral of ϕ against B on $\mathbb{R}_+$.

For $t > 0$, we set $\Lambda^2(t) := L^2([0,t] \times \Omega, \text{Prog}, \lambda|_{[0,t]} \otimes \mathbb{P})$. Then, we define $\int_0^t \phi_s dB_s$ as follows:

If $\phi \in \Lambda^2(t)$, the process $(\mathbb{1}_{[0,t](s)}\phi)_s$ belongs to Λ^2, and we define

$$\int_0^t \phi_s dB_s := \int_0^\infty \mathbb{1}_{[0,t]}(s)\phi_s dB_s.$$

In this case, we have the following.

iii) $\mathbb{E}\Big[\int_0^t \phi_s dB_s\Big] = 0$ *and* $\mathbb{E}\Big[\Big(\int_0^t \phi_s dB_s\Big)^2\Big] = \mathbb{E}\Big[\int_0^t \phi_s^2 ds\Big]$.

iv) If $\phi, \psi \in \Lambda^2(t)$, then $\mathbb{E}\Big[\int_0^t \phi_s dB_s \int_0^t \psi_s dB_s\Big] = \mathbb{E}\Big[\int_0^t \phi_s \psi_s ds\Big]$.

For $\phi \in \Lambda^2$ and $t \geq 0$, $M_t := \int_0^t \phi_s dB_s$ is a class of L^2 (since it is defined up to an equivalence). Thus, a natural question is to find a continuous version of M. This is the aim of the next theorem.

Theorem 6.5 *For all $\phi \in \Lambda^2$, there exists $M \in \mathcal{H}_c^2$ such that for all t, $M_t = \int_0^t \phi_s dB_s$ almost surely and $M_t^2 - \int_0^t \phi_s^2 ds$ is a martingale.*

From now on, for $\phi \in \Lambda^2$, $\Big(\int_0^t \phi_s dB_s\Big)_t$ will denote the continuous martingale M such that for all t, $M_t = \int_0^t \phi_s dB_s$ almost surely. M is called the stochastic integral of ϕ. Its main properties are given in the following corollary.

Corollary 6.6 *Let ϕ and $\psi \in \Lambda^2$. For $t \geq 0$, set $M_t := \int_0^t \phi_s dB_s$ and $N_t := \int_0^t \psi_s dB_s$.*

a) $\langle M, N\rangle_t = \int_0^t \phi_s \psi_s ds$ and $M_t N_t - \langle M, N\rangle_t$ is a martingale.

b) If T is a stopping time, then $\mathbb{E}\Big[M_t^T N_t^T\Big] = \mathbb{E}\Big[\int_0^{t\wedge T} \phi_s \psi_s ds\Big]$.

6.2 Stochastic integration against general local martingales

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space, where $(\mathcal{F}_t)_{t\geq 0}$ is a filtration satisfying the usual conditions.

6.2.1 Stochastic integration against L^2-bounded martingales

By Theorem 5.20, if $M \in \mathcal{H}_c^2$, then $\mathbb{E}[\langle M, M\rangle_\infty] < \infty$. Thus, for $M, N \in \mathcal{H}_c^2$, we have $\mathbb{E}[\langle M, M\rangle_\infty] < \infty$ using Theorem 6.7. Consequently, we can define a scalar product on $\mathcal{H}_c^2$ by setting $(M, N)_{\mathcal{H}_c^2} := \mathbb{E}[\langle M, M\rangle_\infty]$. We can easily check that $\|.\|_{\mathcal{H}_c^2}$ is a norm since by using Lemma 5.14 (and identifying indistinguishable processes), we have that $\|M\|_{\mathcal{H}_c^2} = 0$ if and only if $M = 0$.

Now, we give an important inequality which is a generalization of Cauchy-Schwarz inequality to integrals against the quadratic covariation of local martingales.

Theorem 6.7 *(Kunita-Watanabe inequality)*

Let $M, N \in \mathcal{M}_{\text{loc}}$ and let H and K be measurable processes. Then

$$\left|\int_0^\infty H_t K_t \mathrm{d}\langle M,N\rangle_t\right| \leqslant \int_0^\infty |H_t K_t| \mathrm{d}V(\langle M,N\rangle)_t \tag{23}$$

$$\leqslant \left(\int_0^\infty |H_t|^2 \mathrm{d}\langle M,M\rangle_t\right)^{1/2} \left(\int_0^\infty |K_t|^2 \mathrm{d}\langle N,N\rangle_t\right)^{1/2}. \tag{24}$$

Now, we are in position to define the stochastic integrals against L^2-bounded martingales.

Definition 6.8 *For $M \in \mathcal{H}_c^2$, we set $L^2(M) := L^2(\mathbb{R}_+ \times \Omega, \mathrm{Prog}, d\mathbb{P}d\langle M,M\rangle)$, the space of progressively measurable processes H satisfying the condition $\mathbb{E}\left[\int_0^\infty H_s^2 d\langle M,M\rangle_s\right] < \infty$.*

First, note that $L^2(M)$ is a Hilbert space for the canonical scalar product defined be $(H,K)_{L^2(M)} := \mathbb{E}\left[\int_0^\infty H_s K_s d\langle M,M\rangle_s\right]$. Then, we give a density result which will be a key tool in the construction of the stochastic integral against L^2-bounded martingales. We recall that $\mathcal{E}$ denotes the set of elementary processes.

Proposition 6.9 *For all $M \in \mathcal{H}_c^2$, $\mathcal{E}$ is dense in $(L^2(M), \|.\|_{L^2(M)})$.*

Now we define the stochastic integral for an elementary process H against $M \in \mathcal{H}_c^2$.

Definition

Let $M \in \mathcal{H}_c^2$. For all $H \in \mathcal{E}$ i.e. of the form $H := \sum_{i=0}^{p-1} H_i \mathbb{1}_{(t_i,t_{i+1}]}$, we define the stochastic integral of H against M by

$$(H\cdot M)_t := \sum_{i=0}^{p-1} H_i (M_{t_{i+1}\wedge t} - M_{t_i\wedge t}), \qquad t \geqslant 0,$$

Proposition 6.10 *For $M \in \mathcal{H}_c^2$ and $H \in \mathcal{E}$, we have*

i) $H\cdot M \in \mathcal{H}_c^2$.

ii) For all $N \in \mathcal{H}_c^2$,

$$\langle H\cdot M, N\rangle = \int_0^t H_s d\langle M,N\rangle_s = H\cdot\langle M,N\rangle.$$

Theorem 6.11 *Let $M \in \mathcal{H}_c^2$. The map $H \in \mathcal{E} \longmapsto H\cdot M$ is uniquely extended to a linear isometry on $L^2(M)$ with values in $\mathcal{H}_c^2$. In addition,*

i) $H\cdot M$ is characterized by the equality:

$$\text{For all } N \in \mathcal{H}_c^2, \langle H\cdot M, N\rangle = H\cdot\langle M,N\rangle.$$

ii) If T is a stopping time we have

$$(H\mathbb{1}_{[0,T]})\cdot M = (H\cdot M)^T = H\cdot M^T.$$

With the integral notation, we write:

$$\text{For all } t \geq 0, \int_0^t H\mathbb{1}_{[0,T]})dM = \int_0^{t\wedge T} HdM = \int_0^t HdM^T.$$

Proposition 6.12 *Let $M \in \mathcal{H}_c^2$ and $K \in L^2(M)$, $H \in L^2((K \cdot M))$. Then $HK \in L^2(M)$ and*

$$((HK) \cdot M) = (H \cdot (K \cdot M)).$$

We should read this formally as

$$\int_0^\cdot H_s K_s \mathrm{d}M_s = \int_0^\cdot H_s \mathrm{d}\left(\int_0^\cdot K_r \mathrm{d}M_r\right)_s.$$

Remark 6.13 Let $M, N \in \mathcal{H}_c^2$ and $H \in L^2(M)$, $K \in L^2(N)$. Then for all $t \in [0, \infty]$

$$\mathbb{E}[(H \cdot M)_t] = 0, \qquad \mathbb{E}[(H \cdot M)_t (K \cdot N)_t] = \mathbb{E}\left[\int_0^t H_s K_s \mathrm{d}\langle M, N\rangle_s\right],$$

so in particular

$$\mathbb{E}[(H \cdot M)_t^2] = \mathbb{E}\left[\int_0^t H_s^2 \mathrm{d}\langle M, M\rangle_s\right].$$

6.2.2 Stochastic integration against general local martingales

Definition 6.14 *For $M \in \mathcal{M}_{\text{loc}}$ with $M_0 = 0$, $L^2(M)_{\text{loc}}$ denotes the set of progressively measurable processes H such that for all $t \leq 0$, $\int_0^t H_s^2 \mathrm{d}\langle M, M\rangle_s < \infty$ almost surely.*

Theorem 6.15 *(Localization of the stochastic integral)*
Let $M \in \mathcal{M}_{\text{loc}}$ with $M_0 = 0$ and let $H \in L^2_{\text{loc}}(M)$. Then there exists a unique process $H \cdot M = \int_0^\cdot H_s \mathrm{d}M_s \in \mathcal{M}_{\text{loc}}$ such that $(H \cdot M)_0 = 0$ and for all $N \in \mathcal{M}_{\text{loc}}$

$$\langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle.$$

If T is a stopping time we have

$$(H \mathbb{1}_{[0,T]}) \cdot M = (H \cdot M)^T = (H \cdot M^T).$$

If $K \in L^2_{\text{loc}}$ and $H \in L^2_{\text{loc}}(K \cdot M)$, then $HK \in L^2_{\text{loc}}(M)$ and

$$H \cdot (K \cdot M) = (HK) \cdot M.$$

If $M \in \mathcal{H}_c^2$ and $H \in L^2(M)$, then $H \cdot M$ is the same process that we constructed in Theorem 6.11.

Remark 6.16 Let $M \in \mathcal{M}_{\text{loc}}$ and $H \in L^2_{\text{loc}}(M)$. If $\mathbb{E}\left[\int_0^t H_s^2 \mathrm{d}\langle M, M\rangle_s\right] < \infty$ for all $t \geqslant 0$, then $H \cdot M$ is a martingale and

$$\mathbb{E}[(H \cdot M)_t] = 0, \qquad \mathbb{E}[(H \cdot M)_t^2] = \mathbb{E}\left[\int_0^t H_s^2 \mathrm{d}\langle M, M\rangle_s\right], \qquad t \geqslant 0.$$

If even $\mathbb{E}\left[\int_0^\infty H_s^2 \mathrm{d}\langle M, M\rangle_s\right] < \infty$, then $H \cdot M \in \mathcal{H}_c^2$. This follows from Theorem 5.20.

6.3 Itô's formula

6.3.1 Continuous semimartingales

Definition 6.17 *Recall that a continuous semimartingale X is an adapted process*

$$X = X_0 + M + A,$$

where $M \in \mathcal{M}_{\text{loc}}$ with $M_0 = 0$ and $A \in \mathcal{A}$ (so in particular also $A_0 = 0$), and that this decomposition is unique because $\mathcal{A} \cap \mathcal{M}_{\text{loc}} = \{0\}$. If X is a semimartingale and $H \in L^2_{\text{loc}}(M)$ is such that almost surely

$$\int_0^t |H_s| \mathrm{d}V(A)_s < \infty$$

for all $t > 0$, then we define

$$(H \cdot X) := \int_0^\cdot H_s \mathrm{d}X_s := (H \cdot M) + (H \cdot A).$$

In that case we write $H \in \mathbb{L}(X)$.

Remark 6.18 Let $H_s(\omega) = \sum_{i=0}^{p-1} H_i(\omega)\mathbb{I}_{(t_i,t_{i+1}]}(s)$ with $\mathcal{F}_{t_i}$-measurable H_i. Then $H \in \mathbb{L}(X)$ for any continuous semimartingale and

$$(H \cdot X)_t = \sum_{i=0}^{p-1} H_i(X_{t_{i+1}\wedge t} - X_{t_i \wedge t}).$$

Indeed it suffices to show the equality for $X = M \in \mathcal{M}_{\text{loc}}$ because it is obvious for the finite variation part. Then the only difficulty is that $H \notin \mathcal{E}$ because we did not assume the H_i to be bounded. So let $T_n := \inf\{t \geqslant 0 : |H_t| \geqslant n\}$. Then $H\mathbb{I}_{[0,T_n]} \in \mathcal{E}$ and

$$(H \cdot M)^{T_n} = (H\mathbb{I}_{[0,T_n]}) \cdot M = \sum_{i=0}^{p-1} H_i \mathbb{I}_{\{t_i \leqslant T_n\}}(M_{t_{i+1}\wedge t} - M_{t_i\wedge t}).$$

Now simply let $n \to \infty$.

Proposition 6.19 *Let X be a continuous semimartingale and H be adapted and continuous. Then $H \in \mathbb{L}(X)$ and if $t > 0$ and $0 = t_0^n < \ldots < t_{p_n}^n = t$ is a sequence of partitions of $[0,t]$ with $\lim_{n\to\infty} \max\{|t_{k+1}^n - t_k^n|\} = 0$ (we say the mesh size of the partition goes to zero), then*

$$\lim_{n\to\infty} \sup_{s\leqslant t} \left| \sum_{k=0}^{p_n-1} H_{t_k^n}(X_{t_{k+1}^n \wedge s} - X_{t_k^n\wedge s}) - \int_0^s H_r \mathrm{d}X_r \right| = 0,$$

where the convergence is in probability.

Definition 6.20 *Let $X = X_0 + M + A$ and $Y = Y_0 + N + B$ be continuous semimartingales. Then we define*

$$\langle X, Y \rangle := \langle M, N \rangle.$$

Proposition 6.21 *Let X, Y be continuous semimartingales. Let $t > 0$ and $0 = t_0^n < \ldots < t_{p_n}^n = t$ be a sequence of partitions of $[0, t]$ with $\lim_{n\to\infty} \max\{|t_{k+1}^n - t_k^n|\} = 0$. Then*

$$\lim_{n\to\infty} \sup_{s\leqslant t} \left| \sum_{k=0}^{p_n-1} X_{t_k^n \wedge s, t_{k+1}^n \wedge s} Y_{t_k^n \wedge s, t_{k+1}^n \wedge s} - \langle X, Y \rangle_s \right| = 0,$$

where the convergence is in probability. Here we introduced the notation

$$Z_{u,v} := Z_v - Z_u.$$

6.3.2 Itô's formula

Theorem 6.22 *(Itô's formula)*

Let $X = (X^1, \ldots, X^d)$ be a d-dimensional continuous semimartingale (i.e. every X^i is a continuous semimartingale) and let $F \in C^2(\mathbb{R}^d, \mathbb{R})$. Then $F(X)$ is a continuous semimartingale and

$$F(X_t) = F(X_0) + \sum_{i=1}^{d} \int_0^t \partial_i F(X_s) \mathrm{d}X_s^i + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \partial_i \partial_j F(X_s) \mathrm{d}\langle X^i, X^j \rangle_s, \qquad t \geqslant 0. \quad (25)$$

Remark(for tutorial)
If $x \in C^1([0, \infty), \mathbb{R})$, then the fundamental theorem of calculus gives for $f \in C^1(\mathbb{R}, \mathbb{R})$

$$f(x_t) - f(x_0) = \int_0^t \partial_s(f(x_s)) \mathrm{d}s = \int_0^t f'(x_s) \partial_s x_s \mathrm{d}s = \int_0^t f'(x_s) \mathrm{d}x_s.$$

The fundamental theorem of calculus is shown by considering telescope sums. So let us try to mimick the proof say for x replaced by a Brownian motion B. Then formally

$$f(B_t) - f(B_0) = \sum_{\mathrm{d}t} (f(B_{t+\mathrm{d}t}) - f(B_t)) = \sum_{\mathrm{d}t} \{f'(B_t)(B_{t+\mathrm{d}t} - B_t) + O((B_{t+\mathrm{d}t} - B_t)^2)\}.$$

But now very formally $(B_{t+\mathrm{d}t} - B_t)^2 = O\left(\left(\sqrt{\mathrm{d}t}\right)^2\right) = O(\mathrm{d}t)$, so we cannot hope the second contribution to vanish. Instead we need to go one order higher in the Taylor expansion and get

$$f(B_t) - f(B_0) = \sum_{\mathrm{d}t} \left\{ f'(B_t)(B_{t+\mathrm{d}t} - B_t) + \frac{1}{2} f''(B_t)(B_{t+\mathrm{d}t} - B_t)^2 + O((B_{t+\mathrm{d}t} - B_t)^3) \right\}.$$

Now the last term should be of order $(\mathrm{d}t)^{3/2}$ and thus its contribution should vanish, and since $\sum_{\mathrm{d}t} (B_{t+\mathrm{d}t} - B_t)^2 = \langle B \rangle_t = t$, we guess

$$f(B_t) - f(B_0) = \int_0^t f'(B_s) \mathrm{d}B_s + \frac{1}{2} \int_0^t f''(B_s) \mathrm{d}s,$$

which is exactly Itô's formula which involves the additional second order term compared to the smooth setting. Now it only remains to make this formal argumentation rigorous in the general semimartingale setting, which is not much more difficult than what we sketched above.

Corollary 6.23 *(Integration by parts)*
If X and Y are continuous semimartingales, then

$$X_tY_t = X_0Y_0 + \int_0^t X_s \mathrm{d}Y_s + \int_0^t Y_s \mathrm{d}X_s + \langle X, Y\rangle_t, \qquad t \geqslant 0,$$

and in particular

$$X_t^2 = X_0^2 + 2\int_0^t X_s \mathrm{d}X_s + \langle X, X\rangle_t, \qquad t \geqslant 0.$$

If M is a local martingale, we find again that

$$M^2 - M_0^2 - \langle M, M\rangle = 2\int_0^{\cdot} M_s \mathrm{d}M_s$$

is a local martingale.

Remark 6.24 If B is a d-dimensional Brownian motion, then $\langle B^i, B^j\rangle_t = \delta_{i,j}t$, and therefore Itô's formula applied to $F : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ and the d-dimensional semimartingale $(t, B_t)_{t\geqslant 0}$ gives

$$F(t, B_t) = F(t, B_0) + \sum_{i=1}^d \int_0^t \partial_i F(s, B_s)\mathrm{d}B_s^i + \int_0^t \left(\partial_s F(s, X_s) + \frac{1}{2}\Delta F(s, X_s)\right)\mathrm{d}s, \qquad t \geqslant 0,$$

where $\Delta F(x) = \sum_{i=1}^d \partial_{ii}^2 F(x)$ is the Laplacian of F.

In the following we will apply Itô's formula also to complex valued functions, which can be done by treating the real and the imaginary part separately. Similarly, a complex valued continuous stochastic process is a (local) martingale if its real and its imaginary part are both (local) martingales.

Proposition 6.25 *Let $M \in \mathcal{M}_{\mathrm{loc}}$ and $\lambda \in \mathbb{C}$. We set*

$$\mathcal{E}(\lambda M)_t := \exp(\lambda M_t - \lambda^2\langle M, M\rangle_t/2), \qquad t \geqslant 0.$$

Then $\mathcal{E}(\lambda M)$ is a local martingale and solves the stochastic differential equation

$$\mathcal{E}(\lambda M)_t = e^{\lambda M_0} + \lambda\int_0^t \mathcal{E}(\lambda M)_s \mathrm{d}M_s.$$

For $\lambda = 1$ we call $\mathcal{E}(M)$ the stochastic exponential of M.

6.3.3 Some applications of Itô's formula

Theorem 6.26 *(Lévy's characterization of the Brownian motion)*
Let $X = (X^1, \dots, X^d)$ be a d-dimensional adapted continuous process with $X_0^i = 0$ for all i. Then X is a d-dimensional $\mathbb{F}$-Brownian motion if and only if the following two conditions hold:

i. All components X^i are local martingales and *ii.* $\langle X^i, X^j \rangle_t = \delta_{i,j} t$ *for all* $t \geqslant 0$.

In particular, a continuous local martingale M *with* $M_0 = 0$ *is a Brownian motion if and only if* $\langle M, M \rangle_t = t$. *This is false if* M *is allowed to jump!*

Lemma 6.27 *Let* $M \in \mathcal{M}_{\text{loc}}$. *Then* M *and* $\langle M, M \rangle$ *almost surely have the same intervals of constance. That is, for almost all* $\omega \in \Omega$ *we have for all* $0 \leqslant s < t$

$$M_r(\omega) = M_s(\omega) \text{ for all } r \in [s,t] \qquad \Leftrightarrow \qquad \langle M, M \rangle_t(\omega) = \langle M, M \rangle_s(\omega).$$

Theorem 6.28 *(Dambis, Dubins-Schwarz)*
Let $M \in \mathcal{M}_{\text{loc}}$ *with* $M_0 = 0$ *and* $\langle M, M \rangle_\infty = \infty$ *almost surely. Then there exists a Brownian motion* β *such that almost surely*

$$M_t = \beta_{\langle M, M \rangle t}, \qquad t \geqslant 0.$$

In other words, M *is a time-changed Brownian motion.*

Remark 6.29 i. The assumption $\langle M, M \rangle_\infty = \infty$ is not necessary, but without it we might have to enlarge the probability space (think of $|\Omega| = 1$, $\mathcal{F} = \{\emptyset, \Omega\}$, $M = 0$). See Revuz-Yor Theorem V.1.7 for a proof.

ii. The Brownian motion β is in general not adapted to our original filtration, but instead to a "time-changed filtration", as we will see in the proof.

Theorem 6.30 *(Burkholder-Davis-Gundy)*
For every $p > 0$ *there exist* $c_p, C_p > 0$ *such that for all* $M \in \mathcal{M}_{\text{loc}}$ *with* $M_0 = 0$ *and all stopping times* T

$$c_p \mathbb{E}[\langle M, M \rangle_T^{p/2}] \leqslant \mathbb{E}[\sup_{t \leqslant T} |M_t|^p] \leqslant C_p \mathbb{E}[\langle M, M \rangle_T^{p/2}].$$

Now we will establish that if the filtration on Ω is the natural filtration of the Brownian motion, then all martingales can be represented as stochastic integrals against this Brownian motion.

Theorem 6.31 *(Martingale represntation theorem)*
Assume that $(\mathcal{F}_t)_{t \geq 0}$ *is the natural (completed) filtration of the standard Brownian motion* B. *Then, for all random variable* $Z \in L^2(\Omega, \mathcal{F}_\infty)$, *there exists a (unique) process* $h \in L^2(B)$ *(in particular progressively measurable then adapted) such that:*

$$Z = \mathbb{E}[Z] + \int_0^\infty h_s dB_s.$$

Consequently, for any continuous L^2*-bounded martingale (respectively for any* $M \in \mathcal{M}_{\text{loc}}$*), there exists a (unique) process* $h \in L^2(B)$ *(respectively* $h \in L^2_{\text{loc}}(B)$*) and a constant* $C \in \mathbb{R}$ *such that:*

$$M_t = C + \int_0^t h_s dB_s.$$

Lemma 6.32 *Under the assumptions of the previous theorem, the vector space generated by the random variables*

$$\exp i \sum_{j=1}^{n} \lambda_j (B_{t_j} - B_{t_{j+1}}), \text{ for } 0 = t_0 < t_1 < \ldots < t_n \text{ and } \lambda_1, \ldots, \lambda_n \in \mathbb{R},$$

is dense in $L^2_{\mathbb{C}}(\Omega, \mathcal{F}_\infty)$.

6.4 Girsanov's theorem

Itô's formula explores how a semimartingale is transformed when we apply a smooth transformation. Girsanov's theorem study the question how a semimartingale is transformed when we apply a change of the probability measure $\mathbb{P}$.
Throughout this section, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ denotes a filtered probability space, where $(\mathcal{F}_t)_{t\geq 0}$ is a filtration satisfying the usual conditions.

6.4.1 Stochastic logarithm

The martingale property is related to the probability used: if we change $\mathbb{P}$ into Q, a martingale under $\mathbb{P}$ has no reason to be a martingale under Q.
In this section, we study how a semimartingale behaves under a change of measure from $\mathbb{P}$ to Q. We will write X is a $\mathbb{P}$-martingale, or is a $(\mathcal{F}_t, \mathbb{P})$-martingale. We also write $\mathbb{E}_{\mathbb{P}}$ for any expectaction under $\mathbb{P}$.
In the sequel, $Q << \mathbb{P}$ on $\mathcal{F}_\infty$. This implies that for all $t \geq 0$, $Q << \mathbb{P}$ on $\mathcal{F}_t$. We denote by D_t the Radon-Nikodym derivative of Q w.r.t. $\mathbb{P}$ on $\mathcal{F}_t$.

Proposition 6.33 *(Radon-Nikodym derivative process)*

1) *D is an $(\mathcal{F}_t)$-martingale uniformly integrable.*

2) *D has a càdlàg modification. For this version and for any stopping time T,*

$$D_T = \frac{dQ}{d\mathbb{P}}|_{\mathcal{F}_T}.$$

3) *If $Q \sim \mathbb{P}$ on $\mathcal{F}_\infty$, then almost surely for all $t \geq 0$, $D_t > 0$.*

In the sequel, we assume that D has continuous trajectories.

Proposition 6.34 *(Stochastic logarithm) Let $D \in \mathcal{M}_{\text{loc}}$ and positive. Then, there exists a unique $L \in \mathcal{M}_{\text{loc}}$, called the stochastic logarithm of D, such that for all $t \geq 0$,*

$$\text{for all } t \geq 0,\ D_t = \exp\{L_t - \frac{1}{2}\langle L, L\rangle_t\}.$$

Moreover, L satisfies:

$$\text{for all } t \geq 0,\ L_t = \log D_0 + \int_0^t D_s^{-1} dD_s$$

.

Proposition 6.35 *($\mathbb{P}$-martingale and Q-martingale)*

Let $Q \sim \mathbb{P}$ and L be the stochastic logarithm associated to the martingale $D_t = \frac{dQ}{d\mathbb{P}}|_{\mathcal{F}_t}$ that we assume to be continuous. Let X be a continuous adapted process and T be a stopping time such that $(XD)^T$ is a $\mathbb{P}$-martingale. Then, X^T is a Q-martingale.
In particular, if XD is a $\mathbb{P}$-martingale, X is a Q-martingale.

6.4.2 Girsanov's theorem

Theorem 6.36 *(Girsanov's theorem)*

Let $Q \sim \mathbb{P}$ and L be the stochastic logarithm (assumed continuous) associated to the martingale

$$D_t = \frac{dQ}{d\mathbb{P}}|_{\mathcal{F}_t}.$$

If M is a continuous adapted $(\mathcal{F}_t - \mathbb{P})$-local martingale, then the process

$$\tilde{M} := M - \langle M, L\rangle$$

is a continuous $(\mathcal{F}_t - \mathbb{P})$-local martingale.

Remark 6.37 i) Under the assumptions of the previous theorem and denoting $\tilde{M} = G_Q^{\mathbb{P}}(M)$, the map $G_Q^{\mathbb{P}}$ satifies the following.

- $G_Q^{\mathbb{P}}$ send the set of continuous $\mathbb{P}$-local martingales into the set of continuous Q-local martingales.
- $G_{\mathbb{P}}^Q \circ G_Q^{\mathbb{P}} = Id$.
- $G_Q^{\mathbb{P}}$ commutes with the stochastic integral.

ii) A continuous $(\mathcal{F}_t - \mathbb{P})$-local martingale M is a continuous $(\mathcal{F}_t - Q)$-semimartingale, by the decomposition $M = \tilde{M} + \langle M, L\rangle$.

iii) Under the assumptions of the previous theorem , the classes of continuous $(\mathcal{F}_t - \mathbb{P})$-semimartingales and continuous $(\mathcal{F}_t - Q)$-semimartingales coincide.

iv) Let X and Y be two continuous semimartingales (with respect to $\mathbb{P}$ or Q). Then, the quadratic covariation $\langle X, Y\rangle$ remainds unchanged under $\mathbb{P}$ or Q.

v) Let $T > 0$ and $\mathcal{F}_{t\in[0,T]}$ be a given filtration satisfying the usual conditions.
If $Q \sim \mathbb{P}$, we define as previously the martingale $D_{t\in[0,T]}$, and if D has a continous version, we define the martingale $(L_t)_{t\in[0,T]}$. Then, the analogue of the previous theorem (Girsanov) reminds true for $[0, T]$.

Corollary 6.38 *(The Cameron-Martin formula) Let $T > 0$, let B be an $(\mathcal{F}_t - \mathbb{P})$-Brownian motion and $f \in L^2([0, T])$. Then, we have the following.*

1) The random variable

$$D_T = \exp\Big\{\int_0^T f(s)dB_s - \frac{1}{2}\int_0^T f(s)^2 ds\Big\}$$

is a probability density defining a probability measure Q (by $dQ = D_T d\mathbb{P}$).

2) The process

$$B_t^Q := B_t - \int_0^{t\wedge T} f(s)ds$$

is a Q-Brownian motion.
That is , under Q, the $\mathbb{P}$-Brownian motion is written as

$$B_t = B_t^Q + \int_0^{t\wedge T} f(s)ds.$$

Example 6.39 (The supremum of a drifted Brownian)
To study the distribution of the supremum of $\tilde{B}_t := B_t + bt$ on $[0,T]$, it is sufficient to know the distribution of $(B_T, \sup_{t\leq T} B_t)$. Indeed, by the Cameron-Martin formula,

$$\mathbb{P}\big[\sup_{t\leq T}(B_t + bt) \geq x\big] = \mathbb{E}_{\mathbb{P}}\big[\exp\big\{bB_T - \frac{1}{2}b^2T\big\}\mathbb{1}_{\{\sup_{t\leq T}(B_t)\geq x\}}\big].$$

7 Stochastic differential equations (SDEs)

In this chapter, we deal with Brownian Stochastic differential equations (SDEs in short). In next section, we motivate the study of SDEs as a generalization of

7.1 Introduction and definitions

Standard differential equations rule may deterministic phenomenas. To take into account random phenomena, we must take into account "Sochastic differentials", which transforms Differential equations into SDEs.
Differential equations are of the form

$$\dot{x}(t) = a(t, x(t)),$$

where the unknown is the function $t \mapsto x(t)$, satisfying an equation involving $\dot{x}(t)$ and $x(t)$ itself.
SDEs are a generalization of the previous equation differential equation, where the deterministic dynamic a is disturbed by a random term, generally modeled by a Brownian motion B and an intensity of the noise $\sigma(t,x)$:

$$dX_t = a(t, X_t)dt + \sigma(t, X_t)dB_t.$$

Note that the previous equation is symbolic since dB_t has no sens (B is not differentiable). We should write this equation under the form

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

We generalize this definition to the multidimensional case.

Definition 7.1 *(Stochastic diferential equations)*
A stochastic diferential equation (SDE) is an equation of the form:

$$dX_t = a(t, X_t)dt + \sigma(t, X_t)dB_t, \qquad (E(a,\sigma))$$

where the unknown is the $\mathbb{R}^d$-valued process X.

$(E(a,\sigma))$ is also written

$$X_t^i = X_0^i + \int_0^t a_i(s, X_s)ds + \sum_{j=1}^{m} \int_0^t \sigma_{ij}(s, X_s)dB_s^j, 1 \le i \le d,$$

where for $m, d \in \mathbb{N}$:

- *$a(t,x) = (a_i(t,x))_{1\le i\le d}$ is an $\mathbb{R}^d$-valued measurable function defined on $\mathbb{R}_+ \times \mathbb{R}^d$, called the drift.*
- *$\sigma(t,x) = (\sigma_{i,j}(t,x))_{1\le i\le d, 1\le j\le m}$ is an $\mathbb{R}^{d\times m}$-valued measurable function defined on $\mathbb{R}_+ \times \mathbb{R}^d$, called the diffusion coefficient.*
- *$B = (B_1, \dots, B_1)$ is a standard m-dimensional Brownian motion.*

Definition 7.2 *(Solution of an SDE)*
A solution of the SDE $(E(a,\sigma))$ is

- *A filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ satisfying the usual conditions,*
- *An $(\mathcal{F}_t)$-Brownian motion $B = (B_1, \dots, B_1)$ defined on the given probability space.*
- *An $(\mathcal{F}_t)$-adapted continuous $\mathbb{R}^d$-valued process X satisfying equation $(E(a,\sigma))$.*
  If in addition $X_0 = x \in \mathbb{R}^d$, X is said to be solution of the SDE $(E_x(a,\sigma))$

Examples

- The Ornstein-Uhlenbeck process: when $a(t,x) = -ax, a > 0$, and $\sigma(t,x) = \sigma \in \mathbb{R}$.
- The Geometric Brownian motion: when $a(t,x) = ax$ and $\sigma(t,x) = \sigma x$.

7.2 Existence and uniqueness of solutions of SDEs

Definition 7.3 *Consider the equation $(E(a,\sigma))$. We say that we have:*

- *Weak existence if: for all $x \in \mathbb{R}^d$, there exists a solution of $(E_x(a,\sigma))$.*
- *Weak existence and uniqueness if in addition: x being fixed, all the solututions of $(E_x(a,\sigma))$ have the same distribution.*
- *Pathwise uniqueness if: when the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ and the Brownian motion B are fixed, any two solutions X' and X of $(E(a,\sigma))$ such that $X'_0 = X_0$ almost surely are indistinguishable.*

  In addition, we say that a solution X of $(E_x(a,\sigma))$ is a strong solution if X is adapted to the natural filtration of B. We have strong uniqueness for $(E(a,\sigma))$ if for any Brownian motion B, any two strong solutions associated to B are indistinguishable.

Remark 7.4 We can have weak existence and uniqueness without having pathwise uniqueness.

The next theorem gives the link between these notions. For its proof, see Karatzas-Shreve, Proposition 3.20.

Theorem 7.5 *(Yamada-Watanabe)*
Weak existence and pathwise uniqueness imply weak uniqueness. In addition, in this case, for all filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ *and all* $(\mathcal{F}_t)$*-Brownian motion* B*, for any* $x \in \mathbb{R}^d$*, there exists a (unique) strong solution of* $(E_x(a, \sigma))$*.*

In the sequel, we make the following assumptions:
Lipschitz assumptions
The functions a and σ are continuous on $\mathbb{R}_+ \times \mathbb{R}^d$ and Lipschitz in x i.e. there exists a constant $K > 0$ such that: for all $t \geq 0$ and all x and $y \in \mathbb{R}^d$,

$$|a(t, x) - a(t, y)| \leq K|x - y|,$$
$$|\sigma(t, x) - \sigma(t, y)| \leq K|x - y|,$$

and $\int_0^T |a(t, 0)| + |a(t, 0)|dt < \infty$, for all $T > 0$.

Theorem 7.6 *(Cauchy-Lipschitz for SDEs)*
Under the ***Lipschitz assumptions*** *given above, there is pathwise uniqueness for* $(E(a, \sigma))$*. Moreover, for all filtered probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ *and all* $(\mathcal{F}_t)$*-Brownian motion* B*, for any* $x \in \mathbb{R}^d$*, there exists a (unique) strong solution of* $(E_x(a, \sigma))$*.*

Remark 7.7 The continuity assumption on the variable t can be relaxed, since it is useful essentially to upper bound $\sup_{t\leq T} |\sigma(t, x)|$ and $\sup_{t\leq T} |a(t, x)|$ for x fixed:
We can localize the Lipschitz assumption on a and σ and assume only having a constant K depending on the compact set on which t and x are considered. In this case, we need to keep the assumption:

$$|\sigma(t, x)| \leq K(1 + |x|), |\sigma(t, x)| \leq K(1 + |x|), \text{ for all } t \geq 0, x \in \mathbb{R}^d.$$

7.3 Flows on the Wiener space

In this section, we interpreat the solution of the SDE $(E(a, \sigma))$ as a functional on the Wiener space $\big(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^m), \mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^m)), \mathbb{W}\big)$ i.e. space of trajectories of a standard Brownian motion. Here $\mathbb{W}$ denotes the Wiener measure (See J.F. Le Gall lecture notes, section 2.2 for a reminder on the Wiener measure and the Wiener space).
$[x]_t := \{s \mapsto x_s : 0 \leq s \leq t\}$ will denote the trajectory of a function x on $[0, t]$.

Theorem 7.8 *Under the* ***Lipschitz assumptions****, for all* $x \in \mathbb{R}^d$*, there exists a measurable function*

$$\begin{aligned} F_x : \quad \mathcal{C}(\mathbb{R}_+, \mathbb{R}^m) &\to \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d) \\ w &\longmapsto F_x(w) \end{aligned}$$

(when $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^m)$ *is equipped withe the Borel* σ*-algebras completed witth the* $\mathbb{W}$*-null sets)satisfying the following properties:*

i) For all $t \geq 0$, $F_x(w)_t$ coincide $\mathbb{W}(dw)$-almost surely with a measurable function of $[w]_t = (w(r) : 0 \leq r \leq t)$. We denote it by $F_x(t, [B]_t)$.

ii) For all $w \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^m)$, the map: $\mathbb{R}^d \to \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d), x \longmapsto F_x(w)$ is continuous.

*iii) For all $x \in \mathbb{R}^d$, for any choice of the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and any m-dimensional $(\mathcal{F}_t)$-Brownian motion B, the process X defined by $X_t = F_x(B_t)$ is the unique solution of $(E_x(a, \sigma))$.
Moreover, if Z is a $\mathcal{F}_0$-measurable random variable, the process $F_Z(B_t)$ is the unique solution with inital value Z.*

Remark 7.9 Point iii) of the previous theorem shows in particular that there is weak exstence for the SDE $(E(a, \sigma))$: the solutions of $(E_x(a, \sigma))$ are all of the form $F_x(B)$ and then have the same distribution, image of the Wiener measure $\mathbb{W}$ by F_x.

FLow property

We assume always the Lipschitz conditions. Now we consider the general case of an SDE $(E(a, \sigma))$ starting from x at time r i.e. $X_r = x$. Denote by $X_t^{r,x}$ that solution at time $t \geq r$. By the previous theorem, one can write

$$X_t^{r,x} = F(r, x, t, [B_. - B_r]_t),$$

where $(B_. - B_r)_{s+r} := B_{s+r} - B_r, s \geq 0$, which is a Brownian motion by the Markov property of B.

Theorem 7.10 *(Flow property)*

*Under the **Lipschitz assumptions**, the solution of the SDE $(E(a, \sigma))$ with $X_r = x$ satisfies the flow property: for $t_0 \geq r \geq 0$ and $t \geq 0$,*

$$F(r, x, t_0 + t, [B_. - B_r]_{t_0+t}) = F(t_0, X_{t_0}^{r,x}, t_0 + t, [B_. - B_r]_{t_0+t})$$
$$\text{i.e.} \qquad X_{t_0+t}^{r,x} = X_{t_0+t}^{t_0, X_{t_0}^{r,x}}.$$

This property is extended to bounded stopping times: let T be a bounded stopping time, then

$$F(r, x, T + t, [B_. - B_r]_{T+t}) = F(T, X_T^{r,x}, T + t, [B_. - B_r]_{T+t})$$
$$\text{i.e.} \qquad X_{T+t}^{r,x} = X_{T+t}^{T, X_T^{r,x}}.$$

7.4 Strong Markov property for homogeneous SDEs

In this section, we assume always the **Lipschitz assumptions**. We also assume that the coefficients of the SDE do not depend on time i.e. $a(t, y) = a(y)$ and $\sigma(t, y) = \sigma(y)$. In this case, the SDE is said homogeneous.
For any $x \in \mathbb{R}^d$, we denote by $\mathbb{P}_x$ the distribution on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ of the solutions of $(E_x(a, \sigma))$, since by Theorem 7.8, we have $\mathbb{P}_x = \mathbb{W}F_x^{-1}$.
The point ii) of Theorem 7.8 shows that $x \mapsto \mathbb{P}_x$ is continuous for the topology of weak convergence. We deduce by a monotone class argument that for any Borel function ϕ defined on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ with values in $\mathbb{R}$, the map

$$x \mapsto \mathbb{E}_x[\phi]$$

is measurable. Here $\mathbb{E}_x[.]$ denotes the expectation against the measure $\mathbb{P}_x$.

Theorem 7.11 *(Strong Markov property for homogeneous SDEs)*

Let X be a solution of $(E(a,\sigma))$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and let T be a stopping time which is finite almost surely. Then, for any Borel function $\Phi : \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d) \longrightarrow \mathbb{R}_+$, we have

$$\mathbb{E}\big[\Phi(X_{T+t} : t \geq 0)|\mathcal{F}_T\big] = \mathbb{E}_{X_T}\big[\Phi\big]$$

,

that is for any positive $\mathcal{F}_T$-measurable random variable U,

$$\mathbb{E}\big[U\Phi(X_{T+t} : t \geq 0)\big] = \mathbb{E}\big[U\mathbb{E}_{X_T}[\Phi]\big]$$

.

In other words, the conditional distribution of $(X_{T+t} : t \geq 0)$ knowing $\mathcal{F}_T$ (the past) is equal to the distribution of $(X_t : t \geq 0)$ starting from X_T (which depends only on the present at time T).

Lemma(Change of variable for the stochastic integral)
If h is a continuous adapted process and T is a stopping time which is finite almost surely, then for all $t \geq 0$,

$$\int_T^{T+t} h_s dB_s = \int_0^t h_{T+s} dB_s^{(T)},$$

where $B_t^{(T)} := B_{T+t} - B_T, t \geq 0$.

7.5 Probabilistic representation for solutions to partial differential equationss

$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ denotes again a filtered probability space satisfying the usual conditions and B is a given $(\mathcal{F}_t)$-Brownian motion. We assume that the coefficients do not depend on time and that the **Lipschitz assumptions** hold.
We know that there exists a bicontinuous process $(X_t^x : t \geq 0, x \in \mathbb{R}^d)$ such that for $t \geq 0$,

$$X_t = x + \int_0^t a(X_s)ds + \int_0^t \sigma(X_s)dB_s.$$

Moreover, we have for every $p \geq 0$ and all $T \geq 0$, $\mathbb{E}\big[\sup_{t\leq T} |X_t^x|^p\big] < \infty$. We introduce L as the second order differential operator given by

$$L := \sum_{i=1}^d a_i(x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^d (\sigma(x)\sigma^\top(x))_{ij}\frac{\partial^2}{\partial x_i \partial x_j}.$$

The main result is the aim of the next theorem.

Theorem 7.12 *Let $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a Borel function with polynomial growth i.e. there exists $k \geq 0$ such that*

$$|f(x)| \lesssim (1 + |x|^k), \text{ for all } x \in \mathbb{R}^d.$$

Let $u : \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{R}$ be a solution of the Cauchy problem:

$$\frac{\partial u}{\partial t}(t,x) = Lu(t,x),$$

$$u(0,x) = f(x).$$

If there exist a locally integrable function C and $k' \geq 0$ such that for all $t \geq 0$ and all $x \in \mathbb{R}^n$,

$$|\nabla u(t,x)| \leq C(t)(1 + |x|^{k'}),$$

then $u(t,x) = \mathbb{E}[f(X_t^x)]$.

Last lecture: Regularization by noise

The following is only for entertainment and not exam relevant.

1 - ODEs with more than one solution

Consider the ordinary differential equation for $x : [0,T] \to \mathbb{R}$

$$\partial_t x(t) = b(x(t)), \qquad x(0) = x_0.$$

If $b : \mathbb{R} \to \mathbb{R}$ is Lipschitz-continuous, then we know from analysis (or from the fact that this is a special case of an SDE) that there exists a unique solution. If b is continuous and bounded, then one can approximate b by a sequence (b_n) of Lipschitz-continuous functions and apply the Arzela-Ascoli theorem to show that the associated solutions (x_n) are relatively compact in $C([0,T],\mathbb{R})$, and that any limit point x solves the equation. But in general there is no uniqueness: Consider for example

$$b(x) = 2\,\mathrm{sign}(x)\sqrt{|x|}.$$

Then for $x(0) = 0$ there are infinitely many solutions: for example

$$x(t) = t^2, \qquad x(t) = -t^2, \qquad x(t) = 0$$

or more generally

$$x(t) = \pm \mathbb{I}_{t \geqslant t_0}(t - t_0)^2$$

for all $t_0 \in [0,T]$. Of course this b is unbounded, but this is not why uniqueness fails (indeed we can replace b by $(b \wedge (T^2+1)) \vee (-T^2-1)$ and note that the truncation is never active on $[0,T]$). Uniqueness fails because b is not Lipschitz-continuous in 0. On the other hand if say $x_0 > 0$, then there exists a unique solution $(x(t))_{t\in[0,T]}$ to

$$\partial_t x(t) = 2\,\mathrm{sign}(x(t))\sqrt{|x(t)|}, \qquad x(0) = x_0.$$

Indeed, we have $x(t) = \left(\sqrt{x_0} + t\right)^2$.

2 - Addition of noise restores uniqueness in law

Let us see what happens if we add a Brownian motion to the right hand side of our equation. We first show that then the law of any solution is unique. So we fix $b : \mathbb{R} \to \mathbb{R}$ and we consider a solution X to the equation

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}B_t, \qquad X_0 = x,$$

where we assume that b is bounded and measurable. Define a new probability measure $\mathbb{Q}$ by

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \exp\left(\int_0^T -b(X_s)\mathrm{d}B_s - \frac{1}{2}\int_0^T b^2(X_s)\mathrm{d}s\right).$$

Then under $\mathbb{Q}$ the process

$$B_t + \int_0^t b(X_s)\mathrm{d}s = X_t - x, \qquad t \in [0,T],$$

is a Brownian motion by Girsanov's theorem. So let Γ be in the Kolmogorov sigma algebra $\mathcal{B}(\mathbb{R})^{\otimes[0,T]}$. Then

$$\mathbb{P}(X \in \Gamma) = \mathbb{E}_{\mathbb{Q}}\left[\mathbb{I}_{X\in\Gamma}\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\right] = \mathbb{E}_{\mathbb{Q}}\left[\mathbb{I}_{X\in\Gamma}\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right)^{-1}\right] \tag{26}$$

$$= \mathbb{E}_{\mathbb{Q}}\left[\mathbb{I}_{X\in\Gamma}\exp\left(\int_0^T b(X_s)\mathrm{d}B_s + \frac{1}{2}\int_0^T b^2(X_s)\mathrm{d}s\right)\right] \tag{27}$$

$$= \mathbb{E}_{\mathbb{Q}}\left[\mathbb{I}_{X\in\Gamma}\exp\left(\int_0^T b(X_s)\mathrm{d}X_s - \frac{1}{2}\int_0^T b^2(X_s)\mathrm{d}s\right)\right], \tag{28}$$

where in the last step we used that $\mathrm{d}B_s = \mathrm{d}X_s - b(X_s)\mathrm{d}s$. But since under $\mathbb{Q}$ the process $X - x$ is a Brownian motion, the right hand side is given by

$$\mathbb{E}\left[\mathbb{I}_{x+B\in\Gamma}\exp\left(\int_0^T b(x+B_s)\mathrm{d}B_s - \frac{1}{2}\int_0^T b^2(x+B_s)\mathrm{d}s\right)\right],$$

and does not depend on the specific solution X we started from. So every solution to our SDE has the same law!

In particular, in the $b(x) = 2\,\mathrm{sign}(x)\sqrt{|x|}$, $X_0 = 0$ example it cannot happen that one solution stays almost surely around 0, another one around $+t^2$ and another one around $-t^2$. On a quite intuitive level the reason is that the Brownian forcing pushes the solution away from 0 before it is able to note the singularity of b in that point.

Note also that starting from a Brownian motion and performing a change of measure, we easily get the existence of a weak solution to our SDE (exercise!).

3 - Addition of noise gives pathwise uniqueness

So far we showed that contrary to the deterministic case, the law of the solution $(X_t)_{t\in[0,T]}$ to

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}B_t, \qquad X_0 = x,$$

is uniquely determined whenever $b : \mathbb{R} \to \mathbb{R}$ is bounded and measurable. Let us next show that if b is continuous and bounded, then we even have strong uniqueness. For that purpose consider the solution

$$s(x) = \int_0^x \exp\left(\int_0^y -2b(z)\mathrm{d}z\right)\mathrm{d}y$$

to

$$b(x)s'(x) + \frac{1}{2}s''(x) = 0.$$

Of course there are many solutions to that equation because we did not fix any initial conditions, but this specific one will do. Since

$$s'(x) = \exp\left(\int_0^x -2b(z)\mathrm{d}z\right) > 0,$$

we know that s is strictly increasing and in particular there exists an inverse function s^{-1}. Moreover, since b is continuous we have that $s \in C^2(\mathbb{R}, \mathbb{R})$. So by Itô's formula, the process $s(X)$ solves

$$\mathrm{d}s(X_t) = s'(X_t)\mathrm{d}X_t + \frac{1}{2}s''(X_t)\mathrm{d}\langle X, X\rangle_t = \left(s'(X_t)b(X_t) + \frac{1}{2}s''(X_t)\right)\mathrm{d}t + s'(X_t)\mathrm{d}B_t \quad (29)$$

$$= s'(s^{-1}(s(X_t)))\mathrm{d}B_t. \quad (30)$$

In other words, $Y = s(X)$ is a solution to the SDE

$$\mathrm{d}Y_t = (s' \circ s^{-1})(Y_t)\mathrm{d}B_t, \qquad Y_0 = s(x). \quad (31)$$

But the function $s' \circ s^{-1}$ is Lipschitz-continuous: by the inverse function theorem we have

$$(s' \circ s^{-1})' = \frac{s'' \circ s^{-1}}{s' \circ s^{-1}} = \left(\frac{s''}{s'}\right) \circ s^{-1},$$

and

$$\frac{s''(x)}{s'(x)} = -2b(x),$$

which is bounded uniformly in $x \in \mathbb{R}$ because b is bounded. So the solution Y to (31) is unique and therefore $X = s^{-1}(Y)$ is uniquely determined as well, and not just its law. This observation is due to Zvonkin (1974). Actually the same argument also gives the existence of X, but existence can anyways be easily obtained (as explained above).

The phenomenon that the addition of noise can restore the uniqueness of solutions to otherwise ill-posed deterministic equations is called *regularization by noise* and also today it is an active research field.

Appendix

A Monotone class theorem

Theorem A.1 *(Ethier and Kurtz, Appendix, Corollary 4.4)*
Let H be a linear space of bounded functions on Ω such that

- *H contains all constant functions;*
- *if $(h_n)_{n\in\mathbb{N}} \subset H$ and there exists $h : \Omega \to \mathbb{R}$ with $\sup_{\omega\in\Omega} |h_n(\omega) - h(\omega)| \to 0$, then $h \in H$;*
- *if $(h_n)_{n\in\mathbb{N}} \subset H$ is such that $-C \leqslant h_1 \leqslant h_2 \leqslant \ldots \leqslant C$ for some $C \in \mathbb{R}$ and such that there exists $h : \Omega \to \mathbb{R}$ with $h_n(\omega) \to h(\omega)$ for all $\omega \in \Omega$, then $h \in H$.*

Let $H_0 \subset H$ be closed under multiplication (that is $fg \in H_0$ whenever $f, g \in H_0$). Then H contains all $\sigma(H_0)$-measurable functions.