Lectures on Stochastic Analysis
Autumn 2014 version

Xue-Mei Li
The University of Warwick

Typset: January 22, 2017

Contents

1 Introduction **6**

1.1 Theory of Integration (Lecture 1) 6

1.1.1 Appendix 7

1.2 Stochastic Processes, Brownian Motions (Lecture 1) 8

1.2.1 Itô's Integration Theory w.r.t. Brownian Motion (Lecture 2) 10

1.2.2 Stochastic Integral Equations and Parabolic PDE (Lecture 2) 11

1.3 Appendix. Sample Paths of a Brownian Motion 13

1.4 References 14

2 Stochastic Processes **16**

2.1 Lecture 3. Kolmogorov's Extension Theorem 16

2.2 Lecture 4. Komogorov's Continuity Theorem 18

2.3 Wiener space and Wiener Measure (Lecture 5) 19

2.4 Construction by White noise (Lecture 5) 21

2.5 Appendix A. Functions on the Wiener Space 22

2.6 Appendix B. Borel Measures and Tensor σ-algebras 22

3 Conditional Expectations and Uniform Integrability **24**

3.1 Conditional Expectations (Lecture 6) 24

3.2 Properties of Conditional Expectations (Lecture 6-7) 25

3.2.1 Disintegration and Orthogonal Projection (Lecture 7) . . . 27

3.2.2 Appendix* 28

3.3 Uniform Integrability (Lecture 7) 28

3.4 Appendix 29

3.5 Appendix. Absolute continuity of measures 31

4 Martingales **33**

4.1 Definitions (Lecture 7) 33

4.2 Discrete time martingales(Lecture 8) 34

4.3 Discrete Integrals(Lecture 8) 35
4.4 The Upper Crossing Theorem and Martingale Convergence Theorem(Lecture 9) 36
4.5 Stopping Times (Lecture 10) 38
4.6 The Optional Stopping Theorems (Lecture 11) 41
4.7 Doob's Optional Stopping Theorem (Lecture 12) 42
4.8 Right End of a Martingale and OST II (Lecture 13) 43
4.9 Martingale Inequalities (Lecture 14-15) 45

5 Continuous Local Martingales and The quadratic Variation Process 48
5.1 Lecture 14-15. Local Martingales 48
5.2 The Quadratic Variation Process (Lecture 15-16) 51
5.3 Local Martingale Inequality and Lévy's Martingale Characterization Theorem. (Lecture 18) 53
5.3.1 Appendix 54
5.4 The Hilbert space of L^2 bounded martingale (Lecture 18) 55

6 Stochastic Integration 57
6.1 Introduction (Lectures 18-19) 57
6.1.1 Integration w.r.t. Stochastic Processes of Finite Variation . 59
6.2 Space of Integrands (Lecture 18-19) 61
6.3 Lecture 20. Characterization of Stochastic Integrals 66
6.4 Integration w.r.t. Semi-martingales (Lecture 21) 68
6.5 Stochastic Integration w.r.t. Semi-Martingales (Lecture 22) 72
6.5.1 Appendix 73
6.6 Itô's Formula (Lecture 22-23) 73

7 Stochastic Differential Equations 78
7.1 Stochastic processes defined up to a random time (Lecture 24) . . 78
7.2 Concepts 79
7.3 Stochastic Integral Equations (Lectures 22-26) 80
7.4 Examples 86
7.5 Notions of Solutions (Lectures 26-27) 87
7.6 Notions of uniqueness (Lecture 27) 90
7.6.1 The Yamada-Watanabe Theorem 91
7.7 Markov process and Transition function (Lecture 26) 92
7.7.1 Semigroup and Generators 94
7.7.2 Solutions of SDE as Markov process 95
7.8 Existence of Solutions and The Martingale Problem 97
7.9 Localisation 98

8 Girsanov Transform **101**
8.1 Girsanov Theorem For Martingales (Lecture 28) 101
8.2 Girsanov for Martingales 102

9 Appendix **105**
9.1 Lyapunov Function Test 105
9.1.1 Strong Completeness, flow 107

Prologue

Prerequisites: a working knowledge of probability theory, measure theory, theory of integration, functional analysis, and metric spaces is required.

What do we cover in this course and why?

We cover the theory of martingales, basics of Brownian motions, theory of stochastic integration, basic theory of stochastic differential equations. This will provide the foundation for advancing to topics offered on stochastic flows, geometry of stochastic differential equations and leading to stochastic partial differential equations and Malliavin calculus.

What are Brownian motions? They result from summing many small and independent influential factors (law of large numbers) over a time interval $[0, t]$, $t \geq 0$. So we are talking about Gaussian laws that change with time t.

What are martingales? A stochastic process is a martingale if, roughly speaking, the conditioned average value at a future time t given its value at s is the value at s. On average you expect to see what is already statistically known. Continuous martingales and local martingales can be represented as stochastic integrals with respect to a Brownian motion (Integral Representation Theorem or Clark-Ocone formula).

What are Markov processes? The conditional average of the future value of a Markov process given knowledge of its past up to now is the same as the conditional average of the future value of the Markov process given knowledge on its present status only. The Dubin-Schwartz Theorem says that a martingale is a time change of a Brownian motion, e.g. a Brownian motion run at a random clock. The random clock is the quadratic variation of the martingale. However the time change may not be Markovian, and hence the process may not be a Markov process.

Acknowledgement

This note benefited from readings by those who attended my lectures. I would like to specially thank Michael Coffey, Owen Daniel and Wojciech Ozanski.

Chapter 1

Introduction

1.1 Theory of Integration (Lecture 1)

People have been toying with various concepts of integration theory. We will explore the concept of stochastic integration which is not covered by any of the theories below.

Definition 1.1 A function $f : [a, b] \to \mathbf{R}$ is Riemann integrable if there exists a number I s.t. for any number $\epsilon > 0$, there exists $\delta > 0$ s.t. for any tagged partition $\Delta : a = t_0 < t_1 < \cdots < t_n = b$, $t_i^* \in [t_{i-1}, t_i]$ with mesh $\max_i(t_i - t_{i-1}) < \delta$,

$$\left| \sum_{i=1}^{n} f(t_i^*)(t_i - t_{i-1}) - I \right| < \epsilon.$$

The number I, which will be denoted by $\int_a^b f(t)dt$ is the Riemannian integral of f.

Definition 1.2 Let $f, g : [a, b] \to \mathbf{R}$ be bounded functions. We say f is Riemann-Stieltjes integrable w.r.t. g if for for any number $\epsilon > 0$, there exists $\delta > 0$ s.t. for any tagged partition $\Delta : a = t_0 < t_1 < \cdots < t_n = b$, $t_i^* \in [t_{i-1}, t_i]$ with mesh $\max_i(t_i - t_{i-1}) < \delta$,

$$\left| \sum_{i=1}^{n} f(t_i^*)(g(t_i) - g(t_{i-1})) - I \right| < \epsilon.$$

The number I is the Riemannian-Stieljes integral of f w.r.t. g and will be denoted by $\int_a^b f(t)dg(t)$.

If μ is a finite measure on $[a, b]$ and f is a bounded measurable function we can define $\int_{[a,b]} f_s \mu(ds)$. The so called regulated functions are in the closure of step functions in the uniform topology. Regulated functions are hence bounded Lebesgue measurable functions and are integrable with respect to the Lebesque measure.

Definition 1.3 A function $g : [a, b] \to \mathbf{R}$ has bounded variation if

$$g_{TV}([a,b]) \equiv \mathrm{Var(g, [a, b])} := \sup_{\mathrm{P}} \{ \sum_{\mathrm{i=1}}^{\mathrm{n}} |\mathrm{g(t_{i+1}) - g(t_i)}| \} < \infty$$

where $\Delta : a = t_0 < t_1 < \cdots < t_n = b$, $t_i^* \in [t_{i-1}, t_i]$. The collection of such functions is denoted by $BV([a, b])$.

Both $g_{TV}([a, b])$ and $\mathrm{Var(g, [a, b])}$ are commonly used notations for the total variation of g over $[a, b]$. If g is continuous and of finite total variation, the variation can be obtained by taking a sequence of partitions $\{\Delta_n\}$ whose mesh converges to zero and take the limit $\lim_{|\Delta_n| \to 0} \sum_{i=1}^n |g(t_{i+1}) - g(t_i)|$.

Theorem 1.1 *A real valued function of bounded variation on* $[a, b]$ *is the difference of two monotone functions on* $[a, b]$.

Theorem 1.2 *There is a one to one correspondence between a function* $g \in BV(\mathbf{R}_+)$ *which is also right continuous and a Radon measure* μ_g *on* $\mathbf{R}_+$*,*

$$g(t) - g(0) = \mu_g([0, t]).$$

If f *is integrable with respect to* μ_g*, we say* f *is Stieltjes integrable w.r.t.* g *and define* $\int f_s dg_s = \int f_s d\mu_s$.

If $g(x) = x$ we have Lebesgue integrals.

Young Integral: If f is α-regular and g is β regular with $\alpha + \beta > 1$, Young integral $\int f dg$ can be defined.

Itô integral: we do not assume much on the regularity of the integrand (f_s), left continuous adapted is sufficient, and the integrator (B_s) is almost surely not Hölder continuous of order $\alpha > \frac{1}{2}$.

1.1.1 Appendix

A measure is Radon if it is inner regular, i.e. for any Borel set B and $\epsilon > 0$ there exists a compact $K \subset B$ with $\mu(B \setminus K) < \epsilon$. Note that if g is of bounded variation, for any t_1, t_2 positive, $\mu_g((t_1, t_2]) := g(t_2) - g(t_1)$; $\mu_g(\{t\}) = \mu_g((t, t]) = g(t) - g(t-)$ is the jump of g at t.

Example 1.1 If g is increasing $g_{TV}([a,b]) = g(b) - g(a)$; If $f \in C^1([a,b])$ then $f \in BV([a,b])$. If μ is a finite positive measure, set $f(x) = \mu((-\infty, x])$. Then f is of finite total variation, increasing and right continuous and $\lim_{x \to -\infty} f(x) = 0$.

If $f \in BV([a,b])$ it has derivatives at almost surely all $x \in [a,b]$.

Theorem 1.3 *If f is Lebesque integrable on $[a,b]$, then $\int_a^x f(t)dt$ is a continuous function of finite variation. If $g \in BV([a,b])$ and $f \in C([a,b]; \mathbf{R})$ then f is Riemann-Stieltjes integrable with respect to g, and*

$$|\int_0^t f_s dg_s| \le g_{TV}([a,b]) \cdot |f|_\infty.$$

1.2 Stochastic Processes, Brownian Motions (Lecture 1)

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let E be a separable complete metric space with a σ-algebra $\mathcal{B}$ which is usually the Borel σ-algebra. A function $f : \Omega \to E$ is said to be measurable if the pre-image of any measurable set B, $f^{-1}(B) = \{\omega : f(\omega) \in B\}$, is a measurable set, i.e. belongs to $\mathcal{F}$. Measurable functions on a probability space are also called random variables. The concept of measurability is close to that of continuity: the first determined by σ-algebras and the latter by topologies.

Let I be an index set, indicating time, e.g. $[0,T]$, $[a,b]$, $[0,\infty)$ or $\mathcal{N}$.

Definition 1.4 A stochastic process on a separable metric space E is a map $X : I \times \Omega \to E$ s.t. for any $t \in I$, $X(t) : \Omega \to E$ is measurable.

Remark. In another word, a stochastic process consists of a family of measurable functions $X_t : (\Omega, \mathcal{F}) \to (E, \mathcal{B})$. Recall that the tensor σ-algebra is the smallest one such that for all $\alpha \in I$, the mapping $\pi_\alpha : (E, \otimes_{\alpha \in I} \mathcal{F}_\alpha) \to (E_\alpha, \mathcal{F}_\alpha)$ is measurable. For each ω, we may view the function $t \in I \mapsto X_t(\omega) \in E$ an an element of S^I. Then $X_\cdot : (\Omega, \mathcal{F}) \to (E^I, \otimes_I \mathcal{B})$ is measurable if and only if each $X_t : (\Omega, \mathcal{F}) \to (E, \mathcal{B})$ is measurable.

Example 1.2 (1) Take $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$, and P the Lebesque measure. Take $I = \{1, 2, \dots\}$. Define $X_n(\omega) = \frac{\omega}{n}$. These are continuous functions from $[0,1] \to \mathbf{R}$ and are Borel measurable.

(2) Take $I = [0,3]$. Let $X, Y : \Omega \to \mathbf{R}$ be two random variables on a measure space $(\Omega, \mathcal{F})$. Then $X_t(\omega) = X(\omega)\mathbf{1}_{[0,\frac{1}{2}]}(t) + Y(\omega)\mathbf{1}_{(\frac{1}{2},3]}(t)$ is a stochastic process.

Definition 1.5 Let I be an interval.

(1) A stochastic process $(X_t, t \in I)$ with state space E is said to be sample continuous (or path continuous or a continuous process) if $t \mapsto X_t(\omega)$ is continuous for almost surely all ω.

(2) A stochastic processes is càdlag if $t \mapsto X_t(\omega)$ has left limit and is right continuous for a.s. all ω. Càdlág processes have jumps at the point of discontinuity.

(3) A stochastic process $(X_t, t \in I)$ is said to have independent increments if for any finite number of disjoint intervals $\{[u_i, v_i], i = 1, \dots, n\}$, $\{X_{u_i} - X_{v_i}\}_{i=1}^n$ are independent random variables.

(4) A stochastic process $(X_t, t \geq 0)$ is Gaussian if for any $n \in \mathcal{N}$ and any numbers $0 \leq t_1 < \cdots < t_n$, the distribution of the random variable $(X_{t_1}, \dots X_{t_n})$, with values in $\mathbf{R}^n$, is Gaussian.

Definition 1.6 A stochastic process $(B_t : t \geq 0)$ on $\mathbf{R}^1$ is the standard Brownian motion if $B_0 = 0$ and the following holds:

(1) it is sample continuous,

(2) it has independent increments,

(3) for any $0 \leq s < t$, the distribution of $B_t - B_s$ is $N(0, t - s)$.

Let W_0^d or $C_0([0, \infty), \mathbf{R}^d)$ denote the space of continuous paths over $\mathbf{R}^d$ with initial value 0:

$$W_0^d = C_0([0, \infty), \mathbf{R}^d) := \{\sigma : \mathbf{R}_+ \to \mathbf{R}^d : \sigma \text{ is continuous and} \sigma(0) = 0\}.$$

We may treat (B_t) as a measurable function on the Banach space W_0^d with its Borel σ-algebra.

$$\begin{aligned} B : \Omega &\mapsto W_0^d \\ \omega &\mapsto (B_t(\omega), t \geq 0) \end{aligned}$$

It induces a measure on $(W_0^d, \mathcal{B}(W_0^d))$ which will be called the Wiener measure. The probability space $(W_0^d, \mathcal{B}(W_0^d), \mu)$ is called the Wiener space. The evaluation map

$$ev_t : (W_0^d, \mathcal{B}(W_0^d), \mu) \to \mathbf{R}$$

given by $\mathrm{ev}_t(\sigma) = \sigma(t)$ is a Brownian motion on the Wiener space. Let us visualize a basket of continuous curves, dropping down according to μ, what we see will be the sample paths of the Brownian motion.

How does a typical Brownian path look like? We have the following facts:

Proposition 1.4 *(1) For a.s. all ω and any pair of positive numbers $a < b$,* $\mathrm{Var}(\mathrm{B_t}(\omega), [\mathrm{a}, \mathrm{b}]) = \infty$.

(2) For a.s. all ω, $B_t(\omega)$ cannot have Hölder continuous path of order $\alpha > \frac{1}{2}$.

The integration theory we mentioned earlier fails to define $\int_0^t B_s(\omega) dB_s(\omega)$, path by path.

1.2.1 Itô's Integration Theory w.r.t. Brownian Motion (Lecture 2)

Let $I \subset \mathbf{R}$. If $(X_t, t \in I)$ is a stochastic process.

Definition 1.7 (a) A family $\{\mathcal{F}_t\}_{t \in I}$ of non-decreasing sub-σ- algebras of $\mathcal{F}$ is a filtration if $\mathcal{F}_s \subset \mathcal{F}_t$ whenever $s, t \in I, s < t$.

(b) $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is a filtered probability space.

(c) $(X_t : t \in I)$ is $\mathcal{F}_t$-adapted if X_t is $\mathcal{F}_t$ measurable for each $t \in I$.

(d) The natural filtration, $\mathcal{F}_t^X$, of $(X_t, t \geq 0)$, is the smallest σ-algebra w.r.t. which each X_s, $s \leq t$, is measurable.

Adapted means that the process does not look into the future.

Definition 1.8 An $\mathcal{F}_t$ adapted stochastic process (X_t) is a $(\mathcal{F}_t)$ Brownian motion if it is a Brownian motion and for each $t \geq 0$, $(B_{t+s} - B_s)$ is independent of $\mathcal{F}_s$.

Let K_t be a stochastic processes that is piecewise constant

$$K_t(\omega) = K_{-1}(\omega)\mathbf{1}_{\{0\}}(t) + \sum_{i=0}^{\infty} K_i(\omega)\mathbf{1}_{(t_i, t_{i+1}]}(\omega),$$

where $0 = t_0 < t_1 < t_2 < \dots$ with $\lim_{n\to\infty} t_n = \infty$. If $t \in (t_n, t_{n+1}]$, we define an elementary integral:

$$\int_0^t K_s dB_s = \sum_{i=1}^{n} K_i(\omega)(B_{t_{i+1}}(\omega) - B_{t_i}(\omega)) + K_n(\omega)(B_t(\omega) - B_{t_n}(\omega)).$$

If f is a left continuous and adapted stochastic process, we wish to define $\int_0^t f_s dB_s$. Let Δ_n be a partition of $[0, t]$ with mesh converging to zero. On each partition we have a piecewise constant function and an elementary integral. We define :

$$\int_0^t K_s dB_s = \lim_{|\Delta_n| \to 0} \sum_{t_i \in \Delta_n}^{n} K_{t_i}(\omega)(B_{t_{i+1}}(\omega) - B_{t_i}(\omega)) + K_n(\omega)(B_t(\omega) - B_{t_n}(\omega)).$$

This convergence will be in probability and we do not hope, in general, that we have almost sure convergence. Such integrals will be local martingales.

1.2.2 Stochastic Integral Equations and Parabolic PDE (Lecture 2)

Let $\sigma, \sigma_0 : \mathbf{R} \to \mathbf{R}$ be Lipschitz continuous functions, and (B_t) a Brownian motion, we seek a stochastic process (x_t) that satisfies the stochastic integral equation

$$x_t = x_0 + \int_0^t \sigma(x_s)dB_s + \int_0^t \sigma_0(x_s)ds.$$

For each initial value x_0, we denote the solution by $(F_t(x_0), t \geq 0)$, whose probability distribution is denoted by $P(t, x_0, dy)$. Assume that the solution is unique and exists for all time (non-explosion). Let $f : \mathbf{R} \to \mathbf{R}$ be bounded Borel measurable, We define

$$P_t f(x) := \mathbf{E} f(F_t(x_0)) = \int_{\mathbf{R}} f(y) P(t, x, dy).$$

Then the Chapman-Kolmogorov equation holds,

$$P(t + s, x, A) = \int_{\mathbf{R}} P(t, y, A)\, P(s, x, dy).$$

and $(F_t(x_0))$ is a Markov process. Then

$$\begin{aligned} P_{t+s} f(x) &= \int_{\mathbf{R}} f(z) P(t + s, x, dz) = \int_{\mathbf{R}} f(z) \int_{\mathbf{R}} P(s, x, dy) P(t, y, dz) \\ &= \int_{\mathbf{R}} P_t f(z) P(s, x, dy) = P_s P_t f(x). \end{aligned}$$

Let us define

$$\mathcal{A} f(x) := \lim_{t \to 0} \frac{P_t f(x) - f(x)}{t},$$

whenever the limit exists. We say that $\mathcal{A}$ is the (infinitesimal) generator of the Markov processes whose domain consists of functions f for which the limit exists.

Then under suitable conditions, $P_t f$ solves the Kolmogorov equation

$$\frac{d}{dt} P_t f = P_t(\mathcal{A} f)$$

and the partial differential equation:

$$\frac{d}{dt} P_t f = \mathcal{A}(P_t f).$$

The generator $\mathcal{A}$ has a formal expression:

$$\mathcal{A} f = \frac{1}{2} (\sigma(x))^2 f''(x) + \sigma_0(x) f'(x).$$

Example 1.3 If (B_t) is a d dimensional Brownian motion, it solves $dx_t = dB_t$ with $x_0 = 0$.

$$P_t f(x) = \mathbf{E} f(x+B_t) = \int_{\mathbf{R}^d} f(x+y) \frac{1}{(2\pi t)^{\frac{d}{2}}} e^{-\frac{|y|^2}{2t}} dy = \int_{\mathbf{R}^d} f(y) \frac{1}{(2\pi t)^{\frac{d}{2}}} e^{-\frac{|y-x|^2}{2t}} dy.$$

If we differentiate $P_t f$ for suitable f, we see that $P_t f$ solves the heat equation,

$$\frac{\partial}{\partial t} = \frac{1}{2} \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}.$$

For $x, y \in \mathbf{R}^d$ let

$$p(t,x,y) = \frac{1}{(2\pi t)^{\frac{d}{2}}} e^{-\frac{|x-y|^2}{2t}}.$$

and let $K_t(x) = p(t, 0, x)$. we define a probability measure $P(t, x, \cdot)$ by

$$P(t,x,A) = \int_A p(t,x,y) dy = \frac{1}{\sqrt{2\pi t}^{\frac{d}{2}}} \int_A e^{\frac{-|y-x|^2}{2t}} dy.$$

Each measure $P(t, x, \cdot)$ is Gaussian measure, with variance t and mean x. This family of measures $\{P(t, x, \cdot)\}$ are called the heat kernel measures.

Exercise 1.1 Prove that $p(t, x, y)$ satisfies the Chapman-Kolmogorov equation

$$\int_{\mathbf{R}^d} p(s,x,y) p(t,y,z) dy = p(s+t,x,z).$$

We quote a standard theorem from PDE, which together with Itô's formula gives the Kolmogorov equation:

Theorem 1.5 *If* $f \in L^p(\mathbf{R}^d, \mathbf{R})$ *where* $1 \le p \le \infty$. *Then*

$$K_t * f(x) := \int_{\mathbf{R}^d} f(y) K_t(x-y) dy = \int_{\mathbf{R}^d} f(y) p(t,x,y) dy$$

satisfies the heat equation

$$\frac{du_t}{dt} = \frac{1}{2} \Delta u_t, \qquad u_0(x) = f(x).$$

on $\mathbf{R}^d \times (0, \infty)$. *If*

(1) $1 \le p < \infty$, *then* $K_t * f \to f$ *in* L^p *as* $t \to 0$.

(2) $f \in L^\infty \cap C(\mathbf{R}^d, \mathbf{R})$, *then* $K_t * f$ *is continuous on* $\mathbf{R}^d \times [0, \infty)$ $(K_0 * f = f)$.

1.3 Appendix. Sample Paths of a Brownian Motion

Proposition 1.6 *Let $\Delta^n : a = t_0^n < t_1^n < \cdots < t_{M_n+1}^n = b$ be a sequence of partitions of $[a,b]$ with $|\Delta_n| \to 0$. Define*

$$T_n = \sum_{i=0}^{M_n} (B_{t_{i+1}^n} - B_{t_i^n})^2.$$

Then

$$\lim_{n\to\infty} \mathbf{E}(T_n - (b-a))^2 = 0.$$

In particular T_n converges in probability to $b-a$. There is a sub-sequence of partitions Δ^{n_k}, such that

$$T_{n_k} \to b-a, \qquad a.s.$$

Proof Firstly,

$$\mathbf{E}T_n = \sum_{i=0}^{M_n} \mathbf{E}(B_{t_{i+1}^n} - B_{t_i^n})^2 = \sum_{i=0}^{M_n} (t_{i+1}^n - t_i^n) = b-a.$$

By the independent increment property of the BM,

$$\begin{aligned}
\mathbf{E}(T_n - (b-a))^2 = \mathbf{var}(T_n) &= \sum_{i=0}^{M_n} \mathbf{var}\left((B_{t_{i+1}^n} - B_{t_i^n})^2\right) \\
&= \sum_{i=0}^{M_n} \mathbf{var}\left((t_{i+1}^n - t_i^n) B_1^2\right), \qquad \text{since } B_t - B_s \overset{d}{=} \sqrt{t-s} B_1 \\
&= \sum_{i=0}^{M_n} \left(t_{i+1}^n - t_i^n\right)^2 (\mathbf{var} B_1^2) \\
&\le \max_i \left(t_{i+1}^n - t_i^n\right) \sum_{i=0}^{M_n} \left(t_{i+1}^n - t_i^n\right) (\mathbf{var} B_1^2) \\
&= \max_i \left(t_{i+1}^n - t_i^n\right) (b-a)(\mathbf{var} B_1^2) \to 0.
\end{aligned}$$

The first statement of Proposition 1.6 holds. Now L^2 convergence implies convergence in probability, and so there is a sub-sequence that is convergent almost surely. □

If the partition is a dyadic partition, i.e. divide each interval by 2 each time, the whole sequence converge almost surely, [14]. Recall Definition 1.3 for the total variation of a function.

Proposition 1.7 *For almost surely all ω, the Brownian paths $t \mapsto B_t(\omega)$ have infinite total variation on any interval $[a, b]$. And $B_t(\omega)$ cannot have Hölder continuous path of order $\alpha > \frac{1}{2}$.*

Proof Fix an ω. Since B_t has almost surely continuous paths we only consider all such ω with $t \mapsto B_t(\omega)$ continuous.

(1) Suppose that $B(\omega)_{TV}([a, b]) < \infty$. Let us consider a sequence of partitions Δ^n such that $T_n = \sum_{i=0}^{M_n}(B_{t_{i+1}^n} - B_{t_i^n})^2$ converges almost surely, see Proposition 1.6. Then

$$\begin{aligned}\sum_{i=0}^{M_n}(B_{t_{i+1}^n} - B_{t_i^n})^2 &\leq \max_i \left|B_{t_{i+1}^n}(\omega) - B_{t_i^n}(\omega)\right| \cdot \sum_{i=0}^{M_n}\left|B_{t_{i+1}^n} - B_{t_i^n}\right| \\ &\leq \max_i \left|B_{t_{i+1}^n}(\omega) - B_{t_i^n}(\omega)\right| \cdot B(\omega)_{TV}([a, b]) \to 0\end{aligned}$$

The convergence follows from the fact that $B_t(\omega)$ is uniformly continuous on $[a, b]$. This contradicts that $\sum_{i=0}^{n}(B_{t_{i+1}^n}(\omega) - B_{t_i^n}(\omega))^2$ converges to $b - a$.

(2) Suppose that $\left|B_{t_{i+1}^n}(\omega) - B_{t_i^n}(\omega)\right| \leq C(\omega)|t - s|^\alpha$, where $C(\omega)$ is a constant for each ω, for some $\alpha > \frac{1}{2}$.

$$\begin{aligned}&\sum_{i=0}^{M_n}|B_{t_{i+1}^n}(\omega) - B_{t_i^n}(\omega)|^2 \leq C^2(\omega)\sum_{i=0}^{M_n}|t_{i+1}^n - t_i^n|^{2\alpha} \\ &\leq C^2(\omega)|\Delta^n|^{2\alpha-1}\sum_{i=0}^{M_n}(t_{i+1}^n - t_i^n) \\ &\leq C^2(\omega)(b-a)|\Delta^n|^{2\alpha-1} \to 0,\end{aligned}$$

as $2\alpha - 1 > 0$. This contradicts with Proposition 1.6. □

1.4 References

For a comprehensive study of martingales we refer to "Continuous Martingales and Brownian Motion" by D. Revuz and M. Yor [24]. An enjoyable read for introduction to martingales is the book "Probability with martingales" by D. Williams [30]. For further reads on Brownian motions check on M. Yor's recent books, e.g. [31] also [18] by R. Mansuy-M. Yor, and also [19] by P. Morters and Y. Peres.

For an overall reference for stochastic differential equations, we refer to "Stochastic differential equations and diffusion processes, second edition" by N. Ikeda and S. Watanabe [13]. The small book [16] by H. Kunita is nice to read. There are two lovely books by A. Friedman "Stochastic differential equations and applications" [9, 10], and "Stochastic differential equations " by I. Gihman and A.V. Skorohod [11]. Another book that is good for working out examples is "Stochastic stability of differential equations" by R. Z. Khasminskii [12]. Two books that are good for the beginners are "Stochastic Differential Equations" by B. Oksendale [20] and "Brownian Motion and Stochastic Calculus" by I. Karatzas and S.E. Shreve [15]. The book by Oksendale has 6 editions. I like edition three and edition four: they are neat and compact. For further studies there are "Diffusions, Markov processes and Diffusions" by C. Rogers and D. Williams [26, 25]. Another lovely reference book is "Foundations of Modern Probability" by Kallenberg [14]. It would work great as a reference book. For stochastic integrals for stochastic processes with jumps read Protter [22]. For SDEs driven by space time martingales see "Stochastic Flows and Stochastic Differential Equations" by H. Kunita [17]. For SDEs on manifolds see "Stochastic differential equations on manifolds" by K. D. Elworthy [4]. For work from the point of view of random dynamics see "Random Dynamical systems" by L. Arnold [1] and "Random perturbations of dynamical systems" by M. I. Freidlin and A.D. Wentzell. For further work on the geometry of SDEs have a look at the books "On the geometry of diffusion operators and stochastic flows" [7] and "The geometry of filtesing" [5] by K. D. Elworthy, Y. LeJan and X.-M. Li. For a theory on Markov processes and especially the treatment of the Martingale problem see "Multidimensional diffusion processes" by D. Stroock and S. R.S. Varadhan [28]. There are a number of nice and slim books by the two authors, see D.W. Stroock [27] and S. R.S. Varadhan [29].

If you wish to review the theory of integration, try Royden's book "Real Analysis". It is easy to read and useful as a reference. For further study on measures see "Real Analysis" by Folland [8]. Have a read of "Probability measures on metric spaces" by Parthasarathy [21] for a deep theory on measures. The books "Measure Theory, vol 1&2" by Bogachev [3] is quite useful. For some aspects measure on the Wiener space see "Convergence of Probability measures" by Billingsley [2].

Chapter 2

Stochastic Processes

In lectures 3-4, we discuss the existence of a Brownian motion on $\mathbf{R}$.

2.1 Lecture 3. Kolmogorov's Extension Theorem

Let $(X, \mathcal{B}_1)$ and (Y, B_2) be measurable spaces and μ a measure on $(X, \mathcal{B}_1)$. Let $\Phi : X \to Y$ be a measurable function. It induces a pushed forward measure on $(Y, \mathcal{B}_2)$:

$$(\Phi_* \mu)(A) = \mu(\{x : \Phi(x) \in A\}).$$

If $f : Y \to \mathbf{R}$ be an $\Phi_*(\mu)$-integrable function then

$$\int_X (f \circ \Phi)(x)\ \mu(dx) = \int_Y f(y)\ (\Phi_* \mu)(dy).$$

A measurable function $f : \Omega \to E$ induces a measure on E which is called the probability distribution of f and will be denoted by $\hat{P}_X$.

Definition 2.1 Let $(X_t, t \in [0, \infty))$ be a stochastic process on a metric space S. For $n \in \mathcal{N}$, and $0 \le t_1 < t_2 < \cdots < t_n$ we denote by $\mu_{t_1, \ldots, t_n}$ the probability measure on S^n pushed forward by

$$(X_{t_1}, X_{t_2}, \ldots, X_{t_n}).$$

The family of probability measures $\{\mu_{t_1, \ldots, t_n}\}$ are the finite dimensional distributions of the stochastic process (X_t).

Example 2.1 Let (B_t) be a one dimensional Brownian motion, and $\{A_i\}$ Borel sets in $\mathbf{R}$, then

$$\begin{aligned}&P(B_{t_1} \in A_1, \dots, B_{t_k} \in A_k)\\&= \int_{A_1} \dots \int_{A_k} p_{t_1}(0, y_1) p_{t_2 - t_1}(y_1, y_2) \dots p_{t_k - t_{k-1}}(y_{k-1}, y_k) dy_k \dots dy_1.\end{aligned}$$

Proof I prove this for $k = 2$, the rest is left as an exercise. Let $f, g : \mathbf{R} \to \mathbf{R}$ be bounded measurable functions. Then for $s \le t$,

$$\begin{aligned}\mathbf{E} f(B_s) g(B_t) &= \mathbf{E}\left(\{\mathbf{E} f(B_s) g(B_t - B_s + B_s) | \mathcal{F}_s\}\right)\\&= \mathbf{E}\left(f(B_s) \mathbf{E}\{g(B_t - B_s + B_s) | \mathcal{F}_s\}\right)\\&= \mathbf{E}\left(f(B_s) \int_{\mathbf{R}^d} g(z + B_s) p_{t-s}(0, z) dz\right)\\&= \mathbf{E} \int_{\mathbf{R}} \int_{\mathbf{R}} f(x) g(y) p(s, 0, x) p(t - s, x, y) dy dx.\end{aligned}$$

Hence (B_s, B_t) is distributed as $p(s, 0, x)p(t - s, x, y)dydx$.

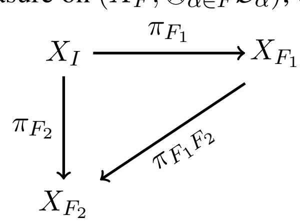
□

Let I be an arbitrary index and let $(X_\alpha, \mathcal{B}_\alpha)_{\alpha \in I}$ be a family of measurable spaces. The Cartesian product $X_I = \Pi_{\alpha \in} X_\alpha$ is the set of all maps x defined on I with $x(\alpha) \in X_\alpha$. We define the coordinate map $\pi_\alpha : X_I \to X_\alpha$ by $\pi_\alpha(x) = x(\alpha)$. The tensor σ-algebra, also called the product σ-algebra, on X_I is the smallest σ-algebra such that each π_α is measurable:

$$\otimes_{\alpha \in I} \mathcal{B}_\alpha = \sigma\{\pi_\alpha^{-1}(A_\alpha) : A_\alpha \in \mathcal{B}_\alpha, \alpha \in I\}.$$

For any $I_2 \subset I_1 \subset I$ let $\pi_{I_1, I_2}(x)$ be the restriction of x in X_{I_1} to X_{I_2}.

A family of measures $\{\mu_F, F \subset I, \#|F| < \infty\}$ is consistent if (1) μ_F is a measure on $(X_F, \otimes_{\alpha \in F} \mathcal{B}_\alpha)$; (2) For any $F_2 \subset F_1 \subset I$, $(\pi_{F_1 F_2})_* \mu_{F_1} = \mu_{F_2}$.



We note that a separable complete metric space with Borel σ-algebra is a standard measure space. See Parthasarathy [21] for detail and for a proof of the following theorem.

Theorem 2.1 (Kolmogorov's Extension Theorem) *Let $(X_\alpha, \mathcal{B}_\alpha, \alpha \in I)$ be 'standard' measure spaces. Given a consistent family of probability measures $\{\mu_F, F \subset I, \#|F| < \infty\}$, there exists a unique probability measure μ on X_I s.t. $(\pi_F)_*\mu = \mu_F$.*

Example 2.2 Let us define a family of finite dimensional probability measures $\{\mu_{t_1,\dots,t_n}, 0 < t_1 < \dots < t_n, n \in \mathcal{N}\}$ as below. Let $A_i \in \mathcal{B}(\mathbf{R}^d)$,

$$\mu_{t_1,\dots,t_n}(\Pi_{j=1}^n A_j)$$
$$= \int_{A_1} \dots \int_{A_n} p(t_1, 0, y_1)p(t_2 - t_1, y_1, y_2) \dots p(t_n - t_{n-1}, y_{n-1}, y_n) dy_n \dots dy_1.$$

Let $E_\alpha = \mathbf{R}, \alpha \in [0,1]$, then $E_I = \mathbf{R}^{[0,1]}$ and the coordinate maps are $\pi_t : x \in E_I \mapsto x(t)$. This is a consistent family of probability measures (exercise). By Kolmogorov's Extension Theorem, there exists a measure $\tilde{\mu}$ on $(\mathbf{R}^{[0,1]}, \otimes_{[0,1]}\mathcal{B}(\mathbf{R}))$ such that its pushed forward measure by the map $\pi_{t_1,\dots,t_n}$ is $\mu_{t_1,\dots,t_n}$. Then $(\pi_t, t \geq 0)$ is a stochastic processes on the probability space $(\mathbf{R}^{[0,1]}, \otimes_{[0,1]}\mathcal{B}(\mathbf{R}), \tilde{\mu})$ with the property that it has independent increments, $\pi_0(x) = 0$ almost surely, $\pi_t - \pi_s \sim N(0, t-s)$ (exercise).

2.2 Lecture 4. Komogorov's Continuity Theorem

Definition 2.2

1. Two stochastic processes X_t and Y_t on the same probability space are *modifications* of each other if for each t, $P(X_t = Y_t) = 1$. The exceptional set $\{\omega : X_t(\omega) \neq Y_t(\omega)\}$ may depend on t.
2. Two stochastic processes X_t and Y_t on the same probability space are *indistinguishable* of each other if $P(X_t = Y_t, \ \forall \text{t}) = 1$.

Let E be a Banach space with norm $\| - \|$, e.g. $E = \mathbf{R}^d$.

Definition 2.3 Let $\alpha \in (0,1)$ and I be an interval of $\mathbf{R}$.

(1) A function $f : I \to E$ is Hölder continuous of exponent α if for all $t, s \in I$,

$$|f(t) - f(s)| \leq C|t-s|^\alpha.$$

(2) A function $f : I \to E$ is locally Hölder continuous of exponent α if on any compact subinterval $[a,b] \subset I$,

$$\sup_{t \neq s, t, s \in [a,b]} \frac{|f(t) - f(s)|}{|t-s|^\alpha} < \infty.$$

Theorem 2.2 (Kolmogorov's Continuity Theorem) *Let $(x_t, t \in I)$ be a stochastic process with values in a separable Banach space $(E, |-|)$. Suppose that there exist positive constants p, δ and C such that for all $s, t \in I$,*

$$\mathbf{E}|x_t - x_s|^p \leq C|t-s|^{1+\delta}.$$

Then there is a continuous modification $(\tilde{x}_t, t \in I)$ of $(x_t, t \in I)$, s.t. for any $\alpha \in (0, \frac{\delta}{p})$ and $[a, b] \subset I$,

$$\mathbf{E} \sup_{s \neq t, s, t \in [a,b]} \left(\frac{|\tilde{x}_s - \tilde{x}_t|}{|t-s|^\alpha} \right)^p < \infty.$$

Example 2.3 Let (x_t) be a stochastic process with $x_t - x_s \sim N(0, t-s)$. Then for any $p \geq 1$,

$$\begin{aligned}
\mathbf{E}|x_t - x_s|^p &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{\infty} |y|^p e^{-\frac{|y|^2}{2(t-s)}} \, dy \\
&= |t-s|^{\frac{p}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z|^p e^{-\frac{|z|^2}{2}} dz \\
&= \mathbf{E}(|x_1|^p)|t-s|^{\frac{p}{2}} < \infty.
\end{aligned}$$

2.3 Wiener space and Wiener Measure (Lecture 5)

Let us consider the separable Banach space

$$W_0^d = C_0([0,1]; \mathbf{R}^d) = \{\omega : [0,1] \to \mathbf{R}^d \text{ continuous }, \omega(0) = 0\}.$$

with the uniform norm $\|\omega\| = \sup_{0 \leq t \leq 1} |\omega(t)|$ and distance

$$d(\omega_1, \omega_2) = \sup_{0 \leq t \leq 1} |\omega_1(t) - \omega_2(t)| = \sup_{t_i \in Q} |\omega_1(t_i) - \omega_2(t_i)|.$$

The Borel σ-algebra on W_0^d is generated by open balls. Let $\{f_i, i \in \mathcal{N}\}$ be a dense set of W_0^d. Since

$$\{\omega \in W_0^d : d(\omega, \omega_0) < a\} = \cap_{t_i \in Q \cap [0,1]} \{\omega \in W_0^d : |\omega(t_i) - \omega_0(t_i)| < a\},$$

$$\mathcal{B}(W_0^d) = \sigma\{\{\omega \in W_0^d : |\omega(t_i) - f_k(t_i)| < a\} : t_i \in Q, k \in \mathcal{N}\}.$$

To ease notation, take $d = 1$ and write $W_0 := W_0^1$.

Theorem 2.3 *Let $\pi_t : W_0 \to \mathbf{R}$ be the evaluation maps: $\pi_t(\omega) = \omega_t$. There is a probability measure μ on $(W_0, \mathcal{B}(W_0))$ such that for any $0 < t_1 < \cdots < t_n$, and any $A_i \in \mathcal{B}(\mathbf{R})$,*

$$
\begin{aligned}
&(\pi_{t_1}, \ldots, \pi_{t_n})_*(\mu)(\Pi_{i=1}^n A_k) \\
&= \int_{A_1} \cdots \int_{A_k} p(t_1, 0, y_1) p(t_2 - t_1, y_1, y_2) \ldots p(t_k - t_{k-1}, y_{k-1}, y_k) dy_k \ldots dy_1.
\end{aligned}
$$

In particular, $(\pi_t, t \le T)$ is a standard Brownian motion on $(W_0, \mathcal{B}(W_0), \mu)$.

Proof Let us take $T = 1$ for simplicity and let $E = \{f_k\}$ a countable dense set of W_0. Open balls in W_0 are determined by 'cylindrical sets' of the following form

$$\{\omega \in W_0 : |\omega(t_i) - f(t_i)| \le r, f \in E, r > 0, 1 \le i \le n, n \in \mathcal{N}\}.$$

These sets are in $\otimes_{[0,1]}\mathcal{B}(\mathbf{R})$. A continuous path is determined by its values on $Q \cap [0,1]$; however we cannot determine whether an arbitrary function from $[0,1]$ to $\mathbf{R}$ is continuous by a countable number of evaluations. Hence $W_0 \notin \otimes_{[0,1]}\mathcal{B}(\mathbf{R})$.

We construct a map

$$\Phi : (\mathbf{R}^{[0,1]}, \otimes_{[0,1]}\mathcal{B}(\mathbf{R})) \to (W_0, \mathcal{B}(W_0))$$

in the following way. If $x : [0,1] \to \mathbf{R}$ is continuous when restricted to Q, we set $\Phi(x)(t_i) = x(t_i)$ and continuously extend the value of $\phi(x))$ to irrational numbers:

$$\Phi(x)(t) = \lim_{t_i \to t, t_i \in Q} x(t).$$

Otherwise we set $\Phi(x)(t) = 0$ for all $t \in [0,1]$. Let $\tilde{\mu}$ be the measure on $\otimes_{[0,1]}\mathcal{B}(\mathbf{R})$ given in Example 2.2. By Kolmogorov's continuity theorem $\tilde{\mu}(x : \Phi(x) \neq x) = 0$. We add all subsets of measurable sets of zero measure to obtain a completion of the σ-algebra $\otimes_{[0,1]}\mathcal{B}(\mathbf{R})$. It is clear that Φ is a measurable map. For any $q \in Q$ take $B_q \in \mathcal{B}(\mathbf{R})$. Then $\cap_{q \in Q \cap [0,1]}\{x : x_q \in B_q\}$ belongs to $\otimes_{[0,1]}\mathcal{B}(\mathbf{R})$. Let $\mu = \Phi_*(\tilde{\mu})$. This is is the required measure. For any $n \in \mathcal{N}$ and $\{t_1, \ldots, t_n\} \in [0,1]$,

$$\mu\left(\cap_{i=1}^n \{x \in \mathbf{R}^{[0,1]} : \pi_{t_i}(x) \in B_i\}\right) = \tilde{\mu}\left(\cap_{i=1}^n \{x \in \mathbf{R}^{[0,1]} : \pi_{t_i}(\Phi(x)) \in B_i\}\right).$$

By Example 2.2, the required property of μ follows. □

2.4 Construction by White noise (Lecture 5)

Let (e_i) be an o.n.b. of $L^2([0,1];\mathbf{R})$. Let

$$x_t = \sum_{i=1}^{\infty} \xi_i \int_0^t e_i(s)ds$$

where $\{\xi_i\}$ are independent random variables with distribution $N(0,1)$. Then for each t, the sum converges in L^2, i.e.

$$\lim_{n\to\infty} \mathbf{E}\left(\sum_{i=n}^{n+m} \xi_i \int_0^t e_i(s)ds\right) = 0.$$

To see this we note that

$$\int_0^t e_i(s)ds = \langle \mathbf{1}_{[0,t]}(s), e_i(s)\rangle_{L^2([0,1];\mathbf{R})}.$$

Since $e_i \in L^2$, by Parseval's theorem,

$$\sum_{i=1}^{\infty}\left(\int_0^t e_i(s)ds\right) < \infty.$$

Let

$$x_t^{(n)} = \sum_{i=1}^{n} \xi_i \int_0^t e_i(s)ds.$$

For each t, there exists a subsequence $\{x_t^{n_k}, k \in \mathcal{N}\}$ which converges almost surely to the limit x_t. It is easy to compute the distribution of $x_t^{(n)}$, it is a mean zero Gaussian random variable and has variance

$$\sum_{i=0}^{n}\left(\int_0^t e_i(s)ds\right)^2.$$

For any $\lambda \in \mathbf{R}$,

$$\begin{aligned}\mathbf{E}e^{i\lambda x_t^{(n)}} &= e^{-\frac{1}{2}\lambda^2 \sum_{i=0}^{n}\left(\int_0^t e_i(s)ds\right)^2}\\ &\to e^{-\frac{1}{2}\lambda^2 \left|\int_0^1 \mathbf{1}_{s\le t}ds\right|_{L^2}} = e^{-\frac{1}{2}\lambda^2 t}.\end{aligned}$$

Since $\mathbf{E}e^{i\lambda x_t^{(n)}} \to \mathbf{E}e^{i\lambda x_t}$ we see that $\mathbf{E}e^{i\lambda x_t} = e^{-\frac{1}{2}\lambda^2 t}$.

A similar computation shows that (x_t) has independent increments:

$$\mathbf{E}e^{i\lambda\sum(x_{t_j}-x_{t_{j-1}})} = \Pi_j e^{-\frac{\lambda^2}{2}(t_j-t_{j-1})} = \Pi_j \mathbf{E}e^{i\lambda(x_{t_j}-x_{t_{j-1}})}.$$

We choose a special basis of L^2. Let $\{e_n\}$ be the Haar functions so that $S_n = \int_0^t e_n(s)ds$ is the Schauder basis. Then the convergence can be shown to be uniform in t on compact subinterval from which it follows that (x_t) has continuous sample paths. This proves that (x_t) is a Brownian motion.

2.5 Appendix A. Functions on the Wiener Space

Let $0 = t_0 < t_1 < \cdots < t_k$ and $g : (\mathbf{R}^d)^k \to \mathbf{R}$ a Borel measurable function. Then functions of the type $f(\omega) = g(\omega_{t_1}, \ldots, \omega_{t_k})$ are called cylindrical functions.

Example 2.4 1. Cylindrical : (a) $f(\omega) = \omega(2)$; (b) $f(\omega) = \omega(1) + (\omega(1)^2)$;

2. Not cylindrical : (c) $f(\omega) = \max_{0\le s\le 1}\omega(s)$; (d) $f(\omega) = \int_0^1 \omega_s ds$.

Let us integrate an cylindrical function:

$$\int_{W_0} g(\omega_{t_1}, \ldots, \omega_{t_k})d\mu(\omega) = \int_{W_0} g(\pi_{t_1,\ldots,t_k}(\omega))d\mu(\omega)$$

$$= \int_{(\mathbf{R}^d)^k} g(y)d(\pi_{t_1,\ldots,t_k})_*\mu(y)$$

$$= \int_{(\mathbf{R}^d)^k} g(y_1, \ldots, y_k)\Pi_{i=0}^k p(t_i - t_{i-1}, y_{k-1}, y_k)dy.$$

where $y_0 = 0$, $t_0 = 0$ and $dy = \Pi_{i=1}^n dy_i$. In particular if $s < t$,

$$\mathbf{E}\pi_t = \int_{W_0} \omega_t \, d\mu(\omega) = \int_{\mathbf{R}^d} y\, p(t, 0, y)dy = 0,$$

$$\int_{W_0} g(\omega_s, \omega_t)\, d\mu(\omega) = \int_{\mathbf{R}^2} g(x, y)p(s, 0, x)p(t-s, x, y)\, dy\, dx.$$

2.6 Appendix B. Borel Measures and Tensor σ-algebras

Let $(E_\alpha, \mathcal{F}_\alpha, \alpha \in I)$ be measurable spaces. The tensor or product σ-algebra of the σ-algebras $\{F_\alpha, \alpha \in I\}$ is

$$\mathcal{F}_I = \sigma\{\pi_\alpha^{-1}(A_\alpha) : A_\alpha \in \mathcal{F}_\alpha, \alpha \in I\}$$

where $\pi_\alpha : E_I \to E_\alpha$ denotes the projection given by the formula $\pi_\alpha(x) = x(\alpha)$.

Proposition 2.4 *For each $\alpha \in I$ let G_α be a generating set of $\mathcal{F}_\alpha$. Then*

$$\mathcal{F}_I = \sigma\{\pi_\alpha^{-1}(A_\alpha) : A_\alpha \in \mathcal{G}_\alpha, \alpha \in I\}.$$

If I is a countable set then,

$$\mathcal{F}_I = \sigma\{\Pi_{\alpha \in I} A_\alpha, A_\alpha \in \mathcal{F}_\alpha\}.$$

Proof (1) It is clear that $C_I := \sigma\{\pi_\alpha^{-1}(A_\alpha) : A_\alpha \in \mathcal{G}_\alpha, \alpha \in I\} \subset \mathcal{F}_I$. But C_I is a σ-algebra containing each $\mathcal{F}_\alpha$.

(2) It is clear that $\mathcal{F}_I \subset \sigma\{\Pi_{\alpha \in I} A_\alpha, A_\alpha \in \mathcal{F}_\alpha\}$. We observe that $\Pi_{\alpha \in I} A_\alpha = \cap_{\alpha \in I} \pi_\alpha^{-1}(A_\alpha)$. Since I is countable, the latter belongs to $\mathcal{F}_I$. $\square$

Let X be a metric space and $\mathcal{B}(X)$ its Borel σ-algebra. A measure on the Borel σ-algebra is a Borel measure. If X is a separable metric space, the metric topology satisfies the second axiom of countability, i.e. there exists a countable base. This countable base generates $\mathcal{B}(X)$. Let $(X_\alpha, \alpha \in I)$ be separable metric spaces. The product topology on X_I is the coarsest topology such that the projections are continuous, it is generated by sets of the form $\{\Pi_\alpha^{-1}(A_\alpha)\}$ where $\alpha \in I$ and A_α are open sets of X_α. The coordinate mappings are measurable with respect to the Borel σ algebra on the product space and $\otimes_\alpha \mathcal{B}(X_\alpha) \subset \mathcal{B}(\Pi_{\alpha \in I} X_\alpha)$.

Proposition 2.5 (Thm 1.10 in [21]) *Let $(X_1, X_2, \dots)$ be separable metric spaces and $X = \Pi_{i=1}^\infty X_i$. Then $\mathcal{B}(X) = \otimes_{n=1}^\infty \mathcal{B}(X_n)$.*

Chapter 3

Conditional Expectations and Uniform Integrability

Definition 3.1 Let $p \geq 1$.

1. A family of Borel measurable functions $\{f_\alpha\}$ on a measure space is L^p bounded if $\sup_\alpha \int |f_\alpha|^p < \infty$.

2. A stochastic process (X_t) is L^p integrable if $\mathbf{E}(|X_t|^p) < \infty$ for all t; it is L^p bounded if $\sup_t \mathbf{E}(|X_t|^p) < \infty$.

3.1 Conditional Expectations (Lecture 6)

Definition 3.2 Let $X \in L^1(\Omega, \mathcal{F}, P)$ be a r.v.. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. A conditional expectation of X given $\mathcal{G}$ is any $\mathcal{G}$-measurable integrable random variable Y such that

$$\int_A X dP = \int_A Y dP, \qquad \forall A \in \mathcal{G} \tag{3.1}$$

Theorem 3.1 *Let $X \in L^1(\Omega, \mathcal{F}, P)$.*

(1) If $Y_1, Y_2 \in L^1(\Omega, \mathcal{G}, P)$ are conditional expectations of X then $Y_1 = Y_2$ a.s.

(2) If $a, b \in \mathbf{R}$, $X_1, X_2 \in L^1(\Omega, \mathcal{F}, P)$ then $\mathbf{E}(aX_1 + bX_2|\mathcal{G}) = a\mathbf{E}(X_1|\mathcal{G}) + b\mathbf{E}(X_2|\mathcal{G})$.

(3) The conditional expectation of X given $\mathcal{G}$ exists.

(4) If $X \geq 0$, $\mathbf{E}(X|\mathcal{G}) \geq 0$.

We denote by $\mathbf{E}(X|\mathcal{G})$ or $\mathbf{E}\{X|\mathcal{G}\}$ any version of the conditional expectation of X given $\mathcal{G}$.

Proof (1) We first prove uniqueness. Let Y_1, Y_2 be variables such that for any $A \in \mathcal{G}$,

$$\int_A (Y_1 - Y_2) dP = 0.$$

This implies that $Y_1 = Y_2$ a.s.

(2) The linearity follows from uniqueness.

(3) and (4). Assume that $X \geq 0$. Define $Q(A) = \int_A X(\omega) dP(\omega)$ for $A \in \mathcal{G}$. Then Q is a measure. The measure P restricts to a measure on $\mathcal{G}$. If $P(A) = 0$ then $Q(A) = 0$. By the Radon-Nikodym theorem, there exists a non-negative random variable $\frac{dQ}{dP}$, that belongs to $L^1(\Omega, \mathcal{G}, P)$, such that

$$Q(A) = \int_A X(\omega) dP(\omega) = \int_A \frac{dQ}{dP} dP.$$

Thus $\frac{dQ}{dP}$ satisfies (3.1) and is the conditional expectation of X given $\mathcal{G}$.

This proves (4).

Let $X \in L^1$. Then $X = X^+ - X^-$ where X^+, X^- are positive functions in L^1. By part (2) they have conditional expectations. We define

$$\mathbf{E}\{X|\mathcal{G}\} = \mathbf{E}\{X^+|\mathcal{G}\} - \mathbf{E}\{X^-|\mathcal{G}\}.$$

(The conditional expectation can also be obtained directly by Radon-Nikodym theorem for signed measures). This proves (3).

□

Proposition 3.2 *For all bounded $\mathcal{G}$-measurable functions g,*

$$\int_\Omega g(\omega) X(\omega) dP(\omega) = \int_\Omega g(\omega) \mathbf{E}\{X|\mathcal{G}\}(\omega) \, dP(\omega). \tag{3.2}$$

3.2 Properties of Conditional Expectations (Lecture 6-7)

Proposition 3.3 *Let $X, Y \in L^1(\Omega, \mathcal{F}, P)$ and $\mathcal{G}$ a sub-σ-algebra of $\mathcal{F}$.*

1. *Positivity Preserving. If $X \leq Y$, then $\mathbf{E}(X|\mathcal{G}) \leq \mathbf{E}(Y|\mathcal{G})$.*

2. *Linearity. For all $a, b \in \mathbf{R}$,*

$$\mathbf{E}(aX + bY|\mathcal{G}) = a\mathbf{E}(X|\mathcal{G}) + b\mathbf{E}(Y|\mathcal{G}).$$

3. $|\mathbf{E}(X|\mathcal{G})| \le \mathbf{E}(|X|\ |\mathcal{G})$.
4. *If* X *is* $\mathcal{G}$*-measurable,* $\mathbf{E}(X|G) = X$.
5. *If* $\sigma(X)$ *is independent of* $\mathcal{G}$*,* $\mathbf{E}(X|\mathcal{G}) = \mathbf{E}X$ *a.s.*
6. *Taking out what is known: If* X *is* $\mathcal{G}$ *measurable,* $XY \in L^1$ *then*

$$\mathbf{E}(XY|\mathcal{G}) = X\mathbf{E}(Y|\mathcal{G}).$$

7. $\mathbf{E}(\mathbf{E}(X|\mathcal{G})\ ) = \mathbf{E}X$.
8. *Tower property: If* $\mathcal{G}_1$ *is a sub* σ*-algebra of* $\mathcal{G}_2$ *then*

$$\mathbf{E}(X|\mathcal{G}_1) = \mathbf{E}\left(\mathbf{E}(X|\mathcal{G}_1)|\mathcal{G}_2\right) = \mathbf{E}\left(\mathbf{E}(X|\mathcal{G}_2)|\mathcal{G}_1\right).$$

9. *Conditional Jensen's Inequality. Let* $\phi : \mathbf{R}^d \to \mathbf{R}$ *be a convex function. Then*

$$\phi\left(\mathbf{E}(X|\mathcal{G})\right) \le \mathbf{E}(\phi(X)|\mathcal{G}).$$

   For $p \ge 1$, $\|\mathbf{E}(X|\mathcal{G})\|_{L_p} \le \|X\|_{L_p}$.
10. *Conditional dominated convergence Theorem. If* $|X_n| \le g \in L^1$ *then*

$$\mathbf{E}(X_n|\mathcal{G}) \to \mathbf{E}(X|\mathcal{G}).$$

11. L^1 *convergence. If* $X_n \to X$ *in* L^1 *then* $\mathbf{E}(X_n|\mathcal{G}) \to \mathbf{E}(X|\mathcal{G})$ *in* L^1.
12. *Monotone Convergence Theorem. If* $X_n \ge 0$ *and* X_n *increases with* n *then* $\mathbf{E}(X_n|\mathcal{G})$ *increases to* $\mathbf{E}(\lim_{n\to\infty} X_n|\mathcal{G})$.
13. *Fatou's Lemma. If* $X_n \ge 0$,

$$\mathbf{E}(\liminf_{n\to\infty} X_n|\mathcal{G}) \le \liminf_{n\to\infty} \mathbf{E}(X_n|\mathcal{G}).$$

14. *Suppose that* $\sigma(X) \vee \mathcal{G}$ *is independent of* $\mathcal{A}$*, then* $\mathbf{E}(X|\mathcal{A} \vee \mathcal{G}) = \mathbf{E}(X|\mathcal{G})$.

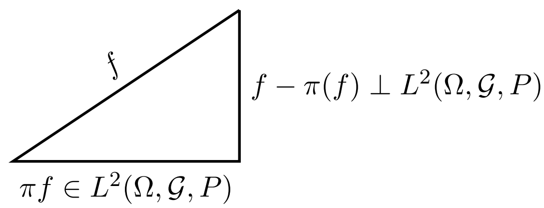
Proposition 3.4 *Let* $h : E \times E \to \mathbf{R}$ *be an integrable function on a metric space* E*. Let* X, Y *be random variables with state space* E *such that* $h(X,Y) \in L^1$*. Let* $H(y) = \mathbf{E}\left(h(X,y)\right)$*. Then*

$$\mathbf{E}(h(X,Y)|\sigma(Y)) = H(Y).$$

3.2.1 Disintegration and Orthogonal Projection (Lecture 7)

Let $\mathcal{G}$ be a sub-σ-algebra of a σ-algebra $\mathcal{F}$. Since $L^2(\Omega, \mathcal{F}, P)$ is a Hilbert space and $L^2(\Omega, \mathcal{G}, P)$ is a closed subspace of L^2, let π denote the orthogonal projection defined by the projection theorem (§II.2 Functional Analysis [23]),

$$\pi : L^2(\Omega, \mathcal{F}, P) \to L^2(\Omega, \mathcal{G}, P).$$



We will see below that the conditional expectation of an L^2 function is precisely its L^2 orthogonal projection to $L^2(\Omega, \mathcal{G}, P)$. We give below second proof for the existence of conditional expectations.

Proof (1) Let $X \in L^2(\Omega, \mathcal{F}, P)$. Then for any $h \in L^2(\Omega, \mathcal{G}, P)$,

$$\langle X - \pi X, h \rangle_{L^2(\Omega, \mathcal{F}, P)} = 0.$$

This is,

$$\int_\Omega X h dP = \int_\Omega \pi(X) h dP$$

Let $A \in \mathcal{G}$ and take $h = \mathbf{1}_A$ to see that

$$\pi X = \mathbf{E}\{X | \mathcal{G}\}.$$

(2) Let $X \in L^1$ with $X \geq 0$. Let $0 \leq X_1 \leq X_2 \leq \dots$ be a sequence of bounded positive functions (increasing with n) converging to X pointwise. Then $X_n \in L^2$, $\{\pi X_n\}$ exists, and are positive. Furthermore for any $A \in \mathcal{G}$,

$$\int_A X_n dP = \int_A \pi X_n dP$$

Since,

$$0 \leq \mathbf{1}_A X_1 \leq \mathbf{1}_A X_2 \leq \dots,$$

$\lim_{n \to \infty} \pi X_n$ exists. By the monotone convergence theorem,

$$\int_A X dP = \lim_{n \to \infty} \int_A X_n dP = \lim_{n \to \infty} \int_A \pi X_n dP = \int_A \lim_{n \to \infty} \pi X_n dP.$$

(3) Finally for $X \in L^1$ not necessarily positive, let $X = X^+ - X^-$ and define $\mathbf{E}\{X|\mathcal{G}\} = \mathbf{E}\{X^+|\mathcal{G}\} - \mathbf{E}\{X^-|\mathcal{G}\}$.

□

Remark 3.1 Let $X \in L^2(\Omega, \mathcal{F}, P)$. Then πX is the unique element of $L^2(\Omega, \mathcal{G}, P)$ such that

$$\mathbf{E}|X - \pi X|^2 = \min_{Y \in L^2(\Omega, \mathcal{G}, P)} \mathbf{E}|X - Y|^2.$$

3.2.2 Appendix*

At this point we note a simple problem from Filtering Theory. Let Y_t be the observation process of a signal process. What is the best estimation for X_t given $\{Y_s, s \le t\}$? We have seen that in the L^2 case, the conditional expectation is an L^2 minimizer. We therefore define the L^2 estimator to be:

$$\hat{X}_t := \mathbf{E}\{X_t | \sigma\{Y_s : 0 \le s \le t\}\}.$$

The concern in filtering is to find the conditional distribution, and the conditional density when it exists, of $X(t)$ given $Y(t)$.

In linear filtering, we assume that

$$X_t(\omega) = X_0(\omega) + W_t(\omega) + \int_0^t F(s) X_s(\omega) ds + \int_0^t f(s) ds \qquad (3.3)$$

$$Y_t(\omega) = \int_0^t H(s) X_s ds + \int_0^t h(s) ds + B_t(\omega). \qquad (3.4)$$

Here $\{(W_t), (B_t)\}$ are independent Brownian motions and both independent of X_0. We assume that $F, f, H, h : \mathbf{R}_+ \to \mathbf{R}$ are bounded measurable functions. This leads to Karman Filter, linear filtering and Zakai equation.

3.3 Uniform Integrability (Lecture 7)

Let $(\Omega, \mathcal{F}, \mu)$ be a (σ-finite) measure space, and I an index set.

Definition 3.3 A family of real-valued measurable functions $(f_\alpha, \alpha \in I)$ is uniformly integrable (u.i.) if

$$\lim_{C \to \infty} \sup_{\alpha \in I} \int_{\{|f_\alpha| \ge C\}} |f_\alpha| d\mu = 0.$$

Lemma 3.5 (Uniform Integrability of Conditional Expectations) *Let $X : \Omega \to \mathbf{R}$ be in L^1. Then the family of functions*

$$\{\mathbf{E}\{X|\mathcal{G}\} : \ \mathcal{G} \text{ is a sub } \sigma\text{-algebra of } \mathcal{F}\}$$

is uniformly integrable.

Lemma 3.6 *Let $X : \Omega \to \mathbf{R}$ be an integrable random function, then the family of functions*

$$\{\mathbf{E}(X|\mathcal{G}) : \ \mathcal{G} \text{ is a sub } \sigma\text{-algebra of } \mathcal{F}\}$$

is uniformly integrable.

Proof exercise. □

Theorem 3.7 *(Vitali Theorem) Let $f_n \in L^p(\mu)$, $p \in [1, \infty]$. Then the following is equivalent.*

1. $f_n \stackrel{L^p}{\to} f$, *i.e.* $\lim_{n\to\infty} \|f_n - f\|_p = 0$.
2. $\{|f_n|^p\}$ *is uniformly integrable and* $f_n \to f$ *in measure.*
3. $\int |f_n|^p d\mu \to \int |f|^p d\mu$ *and* $f_n \to f$ *in measure.*

3.4 Appendix

Let $(S, \mathcal{A}, \mu)$ be a measure space. Let $f, f_\alpha : S \to \mathbf{R}$ be Borel measurable functions.

Proposition 3.8 *If $f \in L^1(\mu)$ where μ is a σ-finite measure, for every $\epsilon > 0$ there is $\delta > 0$ such that for all A with $\mu(A) < \delta$,*

$$\int_A |f| d\mu < \epsilon.$$

Proof We define a measure $\nu(A) = \int_A f d\mu$. It is a signed measure with both the positive and negative part absolutely continuous w.r.t. μ. By considering ν^+, ν^- separately, we may and will assume that $f \geq 0$ and ν is a positive measure. If the conclusion does not hold, there exists a positive number ϵ such that for each n there is a set A_n with $\mu(A_n) < \frac{1}{2^n}$ and

$$\nu(A_n) = \int_{A_n} |f| d\mu \geq \epsilon.$$

Let $A = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k$. Then,

$$\mu(A) = \mu(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) = \lim_{n\to\infty} \mu\left(\cup_{k=n}^{\infty} A_k\right) = 0.$$

In particular $\int_A f d\mu = 0$. But,

$$\nu(A) = \lim_{n\to\infty} \nu\left(\cup_{k=n}^{\infty} A_k\right) \geq \nu(A_n) \geq \epsilon.$$

This gives a contradiction. □

Definition 3.4 A family of integrable real valued random functions $\{f_\alpha\}$ is uniformly absolutely continuous if for every $\epsilon > 0$ there is a number $\delta > 0$ such that if a measurable set A has $\mu(A) < \delta$ then for all $\alpha \in I$

$$\int_A |f_\alpha| d\mu < \epsilon.$$

Proposition 3.9 *Let μ be a finite measure. Let $(f_\alpha, \alpha \in I)$ be a family of integrable real valued functions. The following statements are equivalent:*

(1) $(f_\alpha, \alpha \in I)$ is uniformly integrable (u.i.)

(2) $(f_\alpha, \alpha \in I)$ is L^1 bounded and uniformly absolutely continuous.

(3) (de la Vallee-Poussin criterion) There exists an increasing convex function $\Phi : \mathbf{R}_+ \to \mathbf{R}_+$ such that $\lim_{x\to\infty} \frac{\Phi(x)}{x} = \infty$ and $\sup_\alpha \mathbf{E}\left(\Phi(|f_\alpha|)\right) < \infty$.

Proposition 3.10 *Let $(S, \mathcal{A}, \mu)$ be a measure space. Suppose that $f_n : S \to \mathbf{R}$ belongs to L^1.*

1. *If $f_n \to f$ in L^1 then $\{f_n\}$ is L^1 bounded.*

2. *If $f_n \to f$ in L^1, then $\{f_n\}$ is uniformly absolutely continuous. See exercise 11, section 3.2 in [8].*

3. *Suppose that μ is a finite measure. If $f_n \to f$ in measure and $\{f_n\}$ is uniformly absolutely continuous then $f_n \to f$ in L^1.*

Proof By Riesz-Fisher theorem, the L^1 space is a complete Banach space. (1) is obvious.

(2)Suppose that $f_n \to f$ in L^1. For any $\epsilon > 0$ there is $N(\epsilon)$ such that

$$\sup_{n\geq N} \int |f_n - f| d\mu < \epsilon/2.$$

Let $\alpha > 0$ be such that if $\mu(A) < \alpha$ then

$$\int_A |f| d\mu < \epsilon/2, \qquad \sup_{k \le N-1} \int_A |f_k| d\mu < \epsilon.$$

For $n \ge N$,

$$\int_A |f_n| d\mu \le \int |f_n - f| d\mu + \int_A |f| d\mu < \epsilon.$$

(3) We may assume that $\mu = P$ is a probability measure.

Suppose that $\{f_n\}$ is uniformly absolutely continuous and $f_n \to f$ in measure, i.e. for any $\epsilon > 0$,

$$\lim_{n \to \infty} P(|f_n - f| > \frac{\epsilon}{3}) = 0.$$

Let $\epsilon > 0$. Choose $\delta(\epsilon) > 0$, such that if E is a measurable set with $\mu(E) < \delta$,

$$\sup_n \int_E |f_n| dP < \epsilon/3, \quad \int_E |f| dP < \epsilon/3.$$

There exists $N(\epsilon, \delta)$ such that for $P(|f_n - f| > \epsilon/3) < \delta$ whenever $n \ge N(\delta, \epsilon)$. For such n,

$$\int |f_n - f| dP \le \int_{|f_n - f| \le \frac{\epsilon}{3}} |f_n - f| dP + \int_{|f_n - f| > \frac{\epsilon}{3}} |f_n| dP + \int_{|f_n - f| > \frac{\epsilon}{3}} |f| dP < \epsilon. \tag{3.5}$$

It follows that $f_n \to f$ in L^1.

□

3.5 Appendix. Absolute continuity of measures

Definition 3.5 1. Let $(\Omega, \mathcal{F})$ be a measurable space. Given two measures P and Q. The measure Q is said to be absolutely continuous with respect to P if $Q(A) = 0$ whenever $P(A) = 0$, $A \in \mathcal{F}$. This will be denoted by $Q << P$.

2. They are said to be equivalent, denoted by $Q \sim P$, if they are absolutely continuous with respect to the other.

Theorem 3.11 (Radon-Nikodym Theorem) *If $Q << P$, there is a nonnegative measurable function $\Omega \to \mathbf{R}$, , which we denote by $\frac{dQ}{dP}$, such that for each measurable set A we have*

$$Q(A) = \int_A \frac{dQ}{dP}(\omega) dP(\omega).$$

The function $\frac{dQ}{dP} : \Omega \to \mathbf{R}$ is called the Radon-Nikodym derivative of Q with respect to P. We also say that $\frac{dQ}{dP}$ is the density of Q with respect to P. This function is unique.

Note that if Q is a finite measure then $\frac{dQ}{dP} \in L^1(\Omega, \mathcal{F}, P)$. If P is a probability measure, and $\int_\Omega \frac{dQ}{dP}(\omega)\, dP(\omega) = 1$, then Q is a probability measure.

If furthermore $\frac{dQ}{dP} > 0$, then

$$\int_A dP = \int_A \frac{1}{\frac{dQ}{dP}} \frac{dQ}{dP} dP = \int_A \frac{1}{\frac{dQ}{dP}} dQ.$$

Since $Q(A) = 0$, it follows that $P(A) = \int_A \frac{1}{\frac{dQ}{dP}} dQ = 0$ and $P << Q$. The two measures are equivalent and $\frac{dP}{dQ} \cdot \frac{dQ}{dP} = 1$.

Example 3.1 Let $\Omega = [0, 1)$ and P the Lebesgue measure. Let $A_i^n = [\frac{i}{2^n}, \frac{i+1}{2^n})$, $i = 0, 1, \dots, 2^n - 1$. and $\mathcal{F}_n = \sigma\{A_0^n, A_1^n, \dots, A_{2^n-1}^n\}$. Let μ be a measure on $\mathcal{F}_n$. Check that

$$\frac{d\mu}{dP}(x) = \sum_i \frac{\mu(A_i^n)}{P(A_i^n)} \mathbf{1}_{A_i^n}(x), x \in [0, 1).$$

Two measures Q_1 and Q_2 are singular if $Q_1(A) = 0$ whenever $Q_2(A) \neq 0$ and $Q_2(A) = 0$ whenever $Q_1(A) \neq 0$.

Example 3.2 Let $\Omega = [0, 1]$ and P the Lebesgue measure . Define Q_1 by $\frac{dQ_1}{dP} = 2\mathbf{1}_{[0,\frac{1}{2}]}$. Then $Q_1 << P$ and P is not absolutely continuous with respect to Q_1. Define Q_2 by $\frac{dQ_2}{dP} = 2\mathbf{1}_{[\frac{1}{2},1]}$. The two measures Q_1 and Q_2 are singular.

Chapter 4

Martingales

A filtration $(F_t, t \geq 0)$ is right continuous if $F_{t+} := \cap_{h>0}\mathcal{F}_{t+h}$ equals $\mathcal{F}_t$. Let $\mathcal{F}_\infty = \vee_{t\geq 0}\mathcal{F}_t = \sigma(\cup_{t\geq 0}\mathcal{F}_t)$, the smallest σ algebra containing every σ-algebra $\mathcal{F}_t$, $t \geq 0$. The completion of a σ-algebra $\mathcal{F}_t$ is normally obtained by adding all null sets in $\mathcal{F}_\infty$ whose measure is zero and is called the augmented σ-algebra.

The standard assumption on the filtration is that it is right continuous and each σ-algebra is complete.

The filtration $\{\mathcal{G}_t : \mathcal{G}_t = \mathcal{F}_{t+}\}$ is right continuous. The natural filtration of a continuous process is not necessarily right continuous. Let $\mathcal{F}_s^B := \sigma\{B_r : 0 \leq r \leq s\}$ be the natural filtration of (B_t) complete with respect to P. Then $\mathcal{F}_s^B$ is right continuous. This is due to Blumenthal's $0-1$ law.

Definition 4.1 An $(\mathcal{F}_t)$ adapted stochastic process is a $\mathcal{F}_t$-Brownian motion if it is a Brownian motion and if for every pair of numbers $0 \leq s < t$, $(X_{t+s} - X_s)$ is independent of $\mathcal{F}_s$.

4.1 Definitions (Lecture 7)

Definition 4.2 Let $\mathcal{F}_t$ be a filtration on $(\Omega, \mathcal{F}, P)$. An adapted stochastic process $(X_t, t \in I)$

(1) is a -martingale if $\mathbf{E}|X_t| < \infty$ and

$$\mathbf{E}\{X_t|\mathcal{F}_s\} = X_s, \qquad \forall s \leq t.$$

(2) is a (integrable) sub-martingale, if $\mathbf{E}|X_t| < \infty$ and $\mathbf{E}\{X_t|\mathcal{F}_s\} \geq X_s$ for all $s \leq t$. (In [24], $X_t^+ \in L^1$ is assumed instead of $X_t \in L^1$)

(3) is a (integrable) super martingale if $\mathbf{E}|X_t| < \infty$ and $\mathbf{E}\{X_t|\mathcal{F}_s\} \le X_s$ for all $s \le t$. (In [24], $X_t^- \in L^1$ is assumed instead of $X_t \in L^1$)

If (X_t) is a super-martingale then $(-X_t)$ is a sub-martingale. If (X_t) is both a sub-martingale and a super-martingale, it is a martingale.

Example 4.1 Let $f \in L^1$ and $f_t = \mathbf{E}\{f|\mathcal{F}_t\}$ then f_t is a martingale.

Example 4.2 Take $\Omega = [0,1]$ and define $\mathcal{F}_1$ to be the Borel sets of $[0,1]$ and P the Lebesgue measure. Define $\mathcal{F}_t$ to be the σ-algebra generated by the collection of functions which are Borel measurable when restricted to $[0,t]$ and constant on $[t,1]$. Let $f : [0,1] \to \mathbf{R}$ be an integrable function and define $M_t = \mathbf{E}\{f|\mathcal{F}_t\}$. Then

$$M_t(x) = \begin{cases} f(x), & \text{if } x \le t \\ \frac{1}{1-t}\int_t^1 f(r)dr & \text{if } x > t. \end{cases}$$

Check that for $s < t$,

$$\begin{aligned}
\mathbf{E}\{M_t|\mathcal{F}_s\}(x) &= \begin{cases} f(x), & \text{if } x \le s \\ \frac{1}{1-s}[\int_s^t f(r)dr + \int_t^1 M_t(r)dr] & \text{if } x > s. \end{cases} \\
&= \begin{cases} f(x), & \text{if } x \le s \\ \frac{1}{1-s}[\int_s^t f(r)dr + \int_t^1 (\frac{1}{1-t}\int_t^1 f(u)du)dr] & \text{if } x > s. \end{cases} \\
&= \begin{cases} f(x), & \text{if } x \le s \\ \frac{1}{1-s}[\int_s^1 f(r)dr] & \text{if } x > s. \end{cases} \\
&= M_s(x).
\end{aligned}$$

4.2 Discrete time martingales(Lecture 8)

Proposition 4.1 *If $(X_n, n \in I)$ where I is a countable set is an $\mathcal{F}_n$-martingale if and only if for all $n \in I$,*

$$\mathbf{E}(X_{n+1}|\mathcal{F}_n) = X_n.$$

This can be prove by induction.

Example 4.3 Let $\{X_n, n \in \mathcal{N}\}$, be a sequence of independent integrable random variables. Let $\mathcal{F}_k = \sigma\{X_1, X_2, \dots, X_k\}$.

1. Suppose that $\mathbf{E}(X_n) = 0$. Then $S_n = \sum_{j=1}^n X_j$ is a martingale:

   $$\mathbf{E}(S_n|\mathcal{F}_{n-1}) = \mathbf{E}(X_n|\mathcal{F}_{n-1}) + S_{n-1} = \mathbf{E}(X_n) + S_{n-1} = S_{n-1}.$$

   If $X_n : \Omega \to \{1, -1\}$ are Bernoulli variables, S_n is said to be a simple random walk.

2. Let $\tilde{X}_n = X_n + 1$. Then

$$\tilde{S}_n = \sum_{k=1}^{n} \tilde{X}_n = \sum_{k=1}^{n} X_k + n = S_n + n$$

is a sub-martingale.

3. Suppose that $X_i \geq 0$ and $\mathbf{E}(X_n) = 1$. Then $M_n = \Pi_{i=1}^n X_i$ is a discrete time martingale.

4.3 Discrete Integrals(Lecture 8)

Let X_n be the value of an asset at time n, we may take $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$, then $X_n \in \mathcal{F}_n$. Let H_n be the number of stakes one puts down at time $n-1$, based on the values of $\{X_1, \dots, X_{n-1}\}$, i.e. $H_n \in \mathcal{F}_{n-1}$. Stochastic process $\{H_n\}$ with $H_n \in \mathcal{F}_{n-1}$ is said to be previsible. The total winning at time n will be denoted by $H \cdot X$.

$$\begin{aligned}
&(H \cdot X)_0 = 0 \\
&(H \cdot X)_1 = H_1(X_1 - X_0) \\
&(H \cdot X)_n = H_1(X_1 - X_0) + \dots + H_n(X_n - X_{n-1}), \qquad n \geq 1
\end{aligned}$$

These are discrete 'stochastic integrals' and will be denoted by $\int_0^n H_s dX_s$.

Lemma 4.2 *Let (H_n) be previsible with $|H_n(\omega)| \leq K$ for some $K > 0$.*

(1) If (X_n) is an $(\mathcal{F}_n)$ martingale then $H \cdot X$ is a martingale.

(2) Suppose that $H_n \geq 0$. If (X_n) is a super-martingale then so is $H \cdot X$.

Proof (1) Since H_n is bounded, $(H \cdot X)_n \in L^1$ for each n. Furthermore,

$$\mathbf{E}\{(H \cdot X)_{n+1} | \mathcal{F}_n\} = (H \cdot X)_n + H_{n+1}\mathbf{E}\{X_{n+1} - X_n | \mathcal{F}_n\}.$$

If (X_n) is a martingale, the last term vanishes and

$$\mathbf{E}\{(H \cdot X)_{n+1} | \mathcal{F}_n\} = (H \cdot X)_n.$$

(2) If $H_n \geq 0$ and X_n is a super-martingale, $H_{n+1}\mathbf{E}\{X_{n+1} - X_n | \mathcal{F}_{n-1}\} \leq 0$, hence

$$\mathbf{E}\{(H \cdot X)_{n+1} | \mathcal{F}_n\} \leq (H \cdot X)_n.$$

□

4.4 The Upper Crossing Theorem and Martingale Convergence Theorem(Lecture 9)

Let $a < b$. By 'an upper crossing' by (X_n) we mean a journey starting from below a and ends above b. Let us say $X_n < a$ and $m = \inf_{m>n}\{X_m > b\}$. Then connecting the points $X_n, X_{n+1}, \ldots, X_m$ gives us an 'upper crossing' in graph.

Lemma 4.3 *Let $(X_n, n \in \mathcal{N})$ be a super-martingale and let $U_N([a,b])(\omega)$ be the number of up-crossings of $[a,b]$ made by a stochastic process $\{X_n\}$ by time N. Then*

$$(b-a)\mathbf{E}U_N([a,b]) \le \mathbf{E}(X_N - a)^-.$$

Proof (Sketch) Let H be a betting strategy that you play 1 unit when $X_n < a$, plays until X gets above b and stop playing. Then

$$(H \cdot X)_N \ge (b-a)\left(U_N([a,b])\right) - [X_N(\omega) - a]^-.$$

Taking expectation, using the fact that $(H \cdot X)$ is a super-martingale (Lemma 4.2), to see that

$$0 = \mathbf{E}(H \cdot X)_0 \ge \mathbf{E}(H \cdot X)_N(\omega) \ge (b-a)\mathbf{E}\left(U_N([a,b])\right) - \mathbf{E}[X_N(\omega) - a]^-.$$

□

If (a_n) is a sequence that crosses from a to b infinitely often for some $a < b$, then a_n cannot have a limit. If a_n does not have a limit, there will be a number $a < b$ such that a_n crosses it infinitely often. This is the philosophy behind the following martingale convergence theorem.

Theorem 4.4 *Let $(X_n, n \in \mathcal{N})$ be a discrete time super-martingale.*

(1) Suppose that $\sup_n \mathbf{E}(X_n^-) < \infty$. Then $X_\infty := \lim_{n\to\infty} X_n$ exists.

(2) Assume that $\sup_n \mathbf{E}|X_n| < \infty$ then X_∞ is in L^1.

Proof If $\lim_{n\to\infty} X_n(\omega)$ does not exists, there are two rational numbers $a(\omega) < b(\omega)$ such that

$$\liminf_{N\to\infty} X_N(\omega) < a(\omega) < b(\omega) < \limsup_{N\to\infty} X_N(\omega).$$

Let A be the set of ω such that $\lim_{N\to\infty} X_N$ does not exist. It is clear that

$$A \subset \cup_{a,b\in Q, a<b}\left\{\omega : \liminf_{N\to\infty} X_N(\omega) < a < b < \limsup_{N\to\infty} X_N(\omega)\right\}.$$

Let us prove that for each pairs of rational numbers (a, b), $P(\Lambda_{a,b}) = 0$ where

$$\Lambda_{a,b} = \left\{\omega : \liminf_{N\to\infty} X_N(\omega) < a < b < \limsup_{N\to\infty} X_N(\omega)\right\}.$$

If $\omega \in \Lambda_{a,b}$, there must be an infinite number of visits, $U_N(\omega)$, to below a and to above b: $\lim_{N\to\infty} U_N([a,b]) = \infty$. Hence

$$\Lambda_{a,b} \subset \left\{\omega : \lim_{N\to\infty} U_N([a,b]) = \infty\right\}.$$

By Doob's upper crossing Lemma (Lemma 4.3),

$$(b-a) \lim_{N\to\infty} \mathbf{E}U_N([a,b]) \le \sup_N \mathbf{E}(X_N - a)^- \le \sup_N \mathbf{E}((X_N)^- + |a|) < \infty.$$

By the monotone convergence theorem,

$$\mathbf{E} \lim_{N\to\infty} U_N([a,b]) = \lim_{N\to\infty} \mathbf{E}\, U_N([a,b]) < \infty.$$

In particular $\lim_{N\to\infty} U_N([a,b]) < \infty$ almost surely and $P(\Lambda_{a,b}) = 0$.

If (X_n) is L^1 bounded, we apply Fatou's lemma

$$\mathbf{E}|\lim_{N\to\infty} X_N| \le \mathbf{E} \lim_{N\to\infty} |X_N| \le \liminf_{N\to\infty} \mathbf{E}|X_N| \le \sup_N \mathbf{E}|X_N| < \infty.$$

□

Remark 4.1 If (X_n) is a sub-martingale, we must control its positive part. If $\sup_n \mathbf{E}(X_n)^+ < \infty$ then $\lim_{n\to\infty} X_n$ exist a.s.. Note also that $|X_n| = (X_n^+) + (X_n^-)$. So (X_n) is L^1 bounded if and only if

$$\sup_n \mathbf{E}(X_n)^- < \infty, \qquad \sup_n \mathbf{E}(X_n)^+ < \infty.$$

If (X_t) is a continuous time martingale, it converges along every increasing sequence $\{t_k\}$ by applying the above convergence theorem to the discrete time super-martingale$\{X_{t_k}\}$. Let $f : \mathbf{R}_+ \to \mathbf{R}$ be a function then $\lim_{t\to T} f(t)$ exists if and only if for any sequence $t_n \to T$, $\lim_{n\to\infty} f(t_n)$ exists and have the same limit. However we do not have control over the exceptional sets on which this does not hold, unless some continuity conditions is imposed, in which case the values of the process will be determined by theirs values on Q and the countable points of jumps.

Let $T \in \mathbf{R}_+ \cup \{\infty\}$ and $I = (0, T)$.

Theorem 4.5 *Let $(X_t, t \in I)$ be a right continuous stochastic process. Then $\lim_{t\to T} X_t$ exists almost surely if one of the following conditions hold:*

(1) (X_t) is a super martingale with $\sup_{t<T} \mathbf{E}(X_t^-) < \infty$,

(2) (X_t) is a sub-martingale with $\sup_t \mathbf{E}(X_t^+) < \infty$.

It is easy to see this. For a.s. ω, functions $X_t(\omega)$ has a limit along its increasing sequence of times. Let us simply arrange the rational number and the set of discontinuities of $X_t(\omega)$ in an increasing order.

Corollary 4.6 *If the filtration $(\mathcal{F}_t)$ satisfies the usual assumptions, and $Y \in L^1$, we may choose Y_t among versions of $\mathbf{E}(Y|\mathcal{F}_t)$ such that (Y_t) is a càdlàg martingale.*

4.5 Stopping Times (Lecture 10)

A stopping time is, roughly speaking, the time that an event has arrived. This time is ∞ if the event does not arrive. Let $I \subset \mathbf{R}_+$ and $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ a filtered probability space.

Definition 4.3

1. A function $T : \Omega \to I \cup \{\infty\}$ is a $(\mathcal{F}_t, t \in I)$ stopping time if $\{\omega : T(\omega) \le t\} \in \mathcal{F}_t$ for all $t \in I$.
2. Given a stochastic process $(X_t, t \in I)$. The stopped process X^T is defined by $X_t^T(\omega) = X_{T(\omega)\wedge t}(\omega)$.

Example 4.4

(a) A constant time is a stopping time.

(b) $T(\omega) \equiv \infty$ is also a stopping time.

Proposition 4.7 *A function $T : \Omega \to \mathcal{N}$ is an $\{\mathcal{F}_n, n \in \mathcal{N}\}$ stopping time if and only if $\{T(\omega) = n\} \in \mathcal{F}_n$ for all n.*

Proof If T is a stopping time, $\{T = n\} = \{T \le n\} \cap \{T \le n-1\}^c \in \mathcal{F}_n$. Conversely, $\{T \le n\} = \cup_{i=1}\{T = i\} \in \mathcal{F}_n$ if $\{T(\omega) = n\} \in \mathcal{F}_n$ for all n. □

Let (X_t) be a stochastic process on S. For $B \in \mathcal{B}(S)$ let

$$T_B(\omega) = \inf\{t > 0 : X_t(\omega) \in B\}$$

T_B be the hitting time of B by (X_t). By convention, $\inf(\emptyset) = +\infty$.

Example 4.5 Suppose that (X_n) is $(\mathcal{F}_n)$ adapted. Let B be a measurable set. Then T_B is an $\mathcal{F}_n$ stopping time:

$$\{T_B \le n\} = \cup_{k\le n}\{\omega : X_k(\omega) \in B\} \in \mathcal{F}_n.$$

If (X_t) is an right continuous $(\mathcal{F}_t)$-adapted stochastic process, the hitting time of an open set is an $\mathcal{F}_t^+$-stopping time. Recall one of the usual assumptions: $\mathcal{F}_t = \mathcal{F}_t^+$. The first hitting time of closed set by a continuous $(\mathcal{F}_t)$-adapted stochastic process is an $\mathcal{F}_t$- stopping time.

Proposition 4.8 *Let S, T, T_n be stopping times.*

(1) Then $S \vee T = \max(S,T)$, $S \wedge T = \min(S,T)$ are stopping times.

(2) $\limsup_{n\to\infty} T_n$ and $\liminf_{n\to\infty} T_n$ are stopping times.

Proof Part (1) follows from the following observations:

$$\{\omega : \max(S,T) \le t\} = \{S \le t\} \cap \{T \le t\}, \quad \{\omega : \min(S,T) \le t\} = \{S \le T\} \cup \{T \le t\}.$$

Since

$$\limsup_{n\to\infty} T_n = \inf_{n\ge1} \sup_{k\ge n} T_n, \qquad \liminf_{n\to\infty} T_n = \sup_{n\ge1} \inf_{k\ge n} T_n$$

we only proof that if T_n is an increasing sequence, $\sup_n T_n$ is a stopping time; and if S_n is a decreasing sequence of stopping times with limit S, $\inf_n S_n$ is a stopping time. These follows from

$$\{\sup_n T_n \le t\} = \cap_n \{T_n \le t\}, \quad \{\inf S \le t\} = \cup_n \{S_n \le t\}.$$

□

Definition 4.4 Let T be a stopping time. Define

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap \{T \le t\} \in \mathcal{F}_t, \forall t \ge 0\}.$$

If $T = t$ is a constant time, $\mathcal{F}_T$ agrees with $\mathcal{F}_t$. For T takes values in $\mathcal{N}$, $\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap \{T = n\} \in \mathcal{F}_n, \forall n \in \mathcal{N}\}$.

Theorem 4.9 *(1) If T is a stopping time and (X_t) is a progressively measurable stochastic process, then X_T if $\mathcal{F}_T$-measurable.*

(2) If T is finite stopping time, then $\mathcal{F}_T = \sigma\{X_T : X$ is càdłàg$\}$.

For a proof see Revuz-Yor [24] and Protter [22].

Proposition 4.10 *Let S, T be stopping times.*

(1) If $S \leq T$ then $\mathcal{F}_S \subset \mathcal{F}_T$.

(2) Let $S \leq T$ and $A \in \mathcal{F}_S$. Then $S\mathbf{1}_A + T\mathbf{1}_{A^c}$ is a stopping time.

(3) S is $\mathcal{F}_S$ measurable.

(4) $\mathcal{F}_S \cap \{S \leq T\} \subset \mathcal{F}_{S\wedge T}$.

Proof

(1) If $A \in \mathcal{F}_S$,

$$A \cap \{T \leq t\} = (A \cap \{S \leq t\}) \cap \{T \leq t\} \in \mathcal{F}_t$$

and hence $A \in \mathcal{F}_T$.

(2) Since $\mathcal{F}_S \subset \mathcal{F}_T$,

$$\{S\mathbf{1}_A + T\mathbf{1}_{A^c} \leq t\} = (\{S \leq t\} \cap A) \cup (\{T \leq t\} \cap A^c) \in \mathcal{F}_T.$$

(3) Let $r, t \in \mathbf{R}$, $\{S \leq r\} \cap \{S \leq t\} = \{S \leq \min(r,t)\} \in \mathcal{F}_t$. Hence $\{S \leq r\} \in \mathcal{F}_r$.

(4) Take $A \in \mathcal{F}_S$ and $t \geq 0$. Then

$$A \cap \{S \leq T\} \cap \{S \wedge T \leq t\} = (A \cap \{T \leq t\}) \cap \{S \leq t\} \cap \{S \wedge t \leq T \wedge t\} \in \mathcal{F}_t.$$

which follows as $S \wedge t$ and $T \wedge t$ are $\mathcal{F}_t$-measurable. Hence $A \cap \{S \leq T\} \in \mathcal{F}_{S\wedge T}$.

□

For a nice account of stopping times see Kallenberg [14].

Proposition 4.11 *Let T be a stopping time.*

(1) If (X_n) is a martingale so is the stopped process (X_n^T).

(2) If (X_n) is a super-martingale so is (X_n^T).

Proof For any $n \in \mathcal{N}$,

$$X_n^T = \sum_{i=1}^{n-1} X_i \mathbf{1}_{\{T=i\}} + X_n \mathbf{1}_{\{T\ge n\}}.$$

Hence for every n, $\mathbf{E}|X_n^T| \le \sum_{i=1}^n \mathbf{E}|X_i| < \infty$. We observe that $\mathbf{1}_{\{T\ge n\}} = 1 - \mathbf{1}_{\{T\le n-1\}}$ is $\mathcal{F}_{n-1}$-measurable. Hence

$$\begin{aligned}
\mathbf{E}(X_n^T|\mathcal{F}_{n-1}) &= \sum_{i=1}^{n-1} X_i \mathbf{1}_{\{T=i\}} + \mathbf{E}(X_n \mathbf{1}_{\{T\ge n\}}|\mathcal{F}_{n-1}) \\
&= \sum_{i=1}^{n-1} X_i \mathbf{1}_{\{T=i\}} + \mathbf{1}_{\{T\ge n\}}\mathbf{E}(X_n|\mathcal{F}_{n-1}) \\
&= \sum_{i=1}^{n-2} X_i \mathbf{1}_{\{T=i\}} + X_{n-1}\mathbf{1}_{\{T=n-1\}} + \mathbf{1}_{\{T\ge n\}} X_{n-1} \\
&= X_{n-1}^T.
\end{aligned}$$

In case (2), $\mathbf{1}_{T\ge n}\mathbf{E}(X_n|\mathcal{F}_{n-1}) \le X_{n-1}\mathbf{1}_{T\ge n}$ and $\mathbf{E}(X_n^T|\mathcal{F}_{n-1}) \le X_{n-1}^T$.

□

4.6 The Optional Stopping Theorems (Lecture 11)

By a bounded stopping time T we mean that there exists a number C such that $S(\omega) \le T(\omega) \le C$ a.s. Let I be an ordered countable set of positive real numbers.

Proposition 4.12 (Doob's Elementary Optional Stopping Theorem) *Let $(X_r, r \in I)$ be a super-martingale. Let $S \le T$ be bounded stopping times. Then*

$$\mathbf{E}(X_T) \le \mathbf{E}(X_S).$$

If furthermore $(X_r, r \in I)$ is a martingale, equality holds.

Proof Let $H_n = \mathbf{1}_{\{T\ge n\}} - \mathbf{1}_{\{S\ge n\}}$. Since $S \le T \le N$,

$$(H\cdot X)_N := \sum_{S<i\le T} (X_i - X_{i-1}) = (X_T - X_{T-1}) + \cdots + (X_{S+1} - X_S) = X_T - X_S.$$

Since H is non-negative, (X_n) a super-martingale, then $H\cdot X$ is a super-martingale. Thus $\mathbf{E}(H\cdot X)_N \le \mathbf{E}(H\cdot X)_0 = 0$ and $\mathbf{E}(X_T) \le \mathbf{E}(X_S)$. □

4.7 Doob's Optional Stopping Theorem (Lecture 12)

Let I be any index set.

Proposition 4.13 *Let $(X_t : t \in I)$ be an integrable right or left continuous (or progressively measurable) stochastic process.*

(1) Suppose that for all bounded stopping times $S \le T$, $\mathbf{E}X_T = \mathbf{E}X_S$. Then

$$\mathbf{E}\{X_T|\mathcal{F}_S\} = X_S.$$

(2) Suppose that for all bounded stopping times $S \le T$, $\mathbf{E}X_T \le \mathbf{E}X_S$. Then

$$\mathbf{E}\{X_T|\mathcal{F}_S\} \le X_S.$$

Proof Let $S \le T$ be two stopping times bounded by C. Let $A \in \mathcal{F}_S$. Define $\tau = S\mathbf{1}_A + T\mathbf{1}_{A^c} \le C$. It is a stopping time by Proposition 4.10. Then

$$\begin{aligned}\mathbf{E}X_T &= \mathbf{E}[X_T\mathbf{1}_A] + \mathbf{E}[X_T\mathbf{1}_{A^c}]\\ \mathbf{E}X_\tau &= \mathbf{E}[X_S\mathbf{1}_A] + \mathbf{E}[X_T\mathbf{1}_{A^c}].\end{aligned}$$

(1) For the first statement, $\mathbf{E}X_\tau = \mathbf{E}X_T$ by assumption. Thus $\mathbf{E}[X_T\mathbf{1}_A] = \mathbf{E}[X_S\mathbf{1}_A]$ for all $A \in \mathcal{F}_S$. It follows that $\mathbf{E}\{X_T|\mathcal{F}_S\} = X_S$.

(2) For the second statement, $\mathbf{E}X_\tau \ge \mathbf{E}X_T$ by the assumption giving that

$$\mathbf{E}(X_T\mathbf{1}_A) \le \mathbf{E}(X_S\mathbf{1}_A).$$

Since $\mathbf{E}(X_T\mathbf{1}_A) = \mathbf{E}\left(\mathbf{E}\{X_T|\mathcal{F}_S\}\mathbf{1}_A\right)$, we have

$$\mathbf{E}\left(\left(\mathbf{E}\{X_T|\mathcal{F}_S\} - X_s\right)\mathbf{1}_A\right]) \le 0$$

for any A. Hence $\mathbf{E}\{X_T|\mathcal{F}_S\} \le X_S$.

□

Note that if $X_\infty = \lim_{t\to\infty}$ exists, the above works with T replaced by ∞.

Theorem 4.14 (Doob's Optional Stopping Theorem) *Let S and T be two bounded stopping times such that $S \le T$.*

(1) Let $(X_t, t \ge 0)$ be a right continuous martingale. Then

$$\mathbf{E}\{X_T|\mathcal{F}_S\} = X_S, a.s.$$

(2) Let $(X_t, t \geq 0)$ be a right continuous super-martingale. Then

$$\mathbf{E}\{X_T | \mathcal{F}_S\} \leq X_S$$

almost surely.

Proof We prove part (1). Let $K \in \mathbf{R}$ be such that $S(\omega) \leq T(\omega) \leq K$ a.s.. Let

$$S_n = \frac{1}{2^n}[2^n S + 1].$$

In other words,

$$S_n(\omega) = \frac{j+1}{2^n}, \text{ if } S(\omega) \in \left[\frac{j}{2^n}, \frac{j+1}{2^n}\right), \qquad j = 0, 1, 2 \ldots.$$

Then S_n decreases with n and $|S_n - S| \leq \frac{1}{2^n} \to 0$. If $t \in [\frac{m}{2^n}, \frac{m+1}{2^n})$,

$$\{S_n(\omega) \leq t\} = \{S(\omega) \leq \frac{m}{2^n}\} \in \mathcal{F}_{\frac{m}{2^n}} \subset \mathcal{F}_t.$$

So S_n are stopping times. Recall that (X_t) is integrable. By Doob's elementary optional stopping theorem

$$X_{S_n} = \mathbf{E}\{X_K | \mathcal{F}_{S_n}\}.$$

Since $\mathbf{E}\{X_K | \mathcal{F}_{S_n}, n \in \mathcal{N}\}$ is uniformly integrable, see Lemma 3.5, by Proposition 3.10,

$$\mathbf{E}X_S = \mathbf{E} \lim_{n\to\infty} X_{S_n} = \lim_{n\to\infty} \mathbf{E}X_{S_n} = \lim_{n\to\infty} \mathbf{E}\left(\mathbf{E}\{X_K | \mathcal{F}_{S_n}\}\right) = \mathbf{E}X_K.$$

We have used right continuity of the process. That $\mathbf{E}\{X_T | \mathcal{F}_S\} = X_S$ follows from Proposition 4.13. □

4.8 Right End of a Martingale and OST II (Lecture 13)

Theorem 4.15 (Closability of Martingale) *If $(X_t, t \in [0, T))$ is a right continuous martingale, the following statements are equivalent:*

(1) X_t converges to a r.v. X_T in L^1 (i.e. $\lim_{t\to T} \mathbf{E}|X_t - X_T| = 0$).

(2) There exists a L^1 random variable X_T s.t. $X_t = \mathbf{E}\{X_T | \mathcal{F}_t\}$ for any $0 \leq t \leq T$.

(3) $(X_t, t < T)$ *is uniformly integrable.*

Proof In all cases $\sup_t \mathbf{E}|X_t| < \infty$ and by the convergence theorem (Proposition 4.5), $X_T := \lim_{t\to\infty} X_t$ exists almost surely.

- That (1) is equivalent to (3) is standard, c.f. part (5) of Proposition 3.10
- Assume (2). By Lemma 3.5, the conditional random variables $(X_t, t \geq 0)$ are uniformly integrable, hence (3) holds.
- Assume (3): $(X_t, t < T)$ is uniformly integrable. By Proposition 3.9, $\sup_t \mathbf{E}|X_t|$ is bounded, By Theorem 4.5, $\lim_{t\to T} X_t$ belongs to L^1. By the martingale property, for any $T > u > t$, $X_t = \mathbf{E}\{X_u|\mathcal{F}_t\}$. By the uniform integrability,
$$X_t = \lim_{u\to T} \mathbf{E}(X_u|\mathcal{F}_t) = \mathbf{E}(X_T|\mathcal{F}_t),$$
giving (2).

□

We define $X_T = \lim_{t\to\infty} X_t$ when the limit exists If $\lim_{t\to\infty} X_t$ exists and T a stopping time, we define $X_T = X_\infty$ on $\{T = \infty\}$.

Theorem 4.16 (The Optional Stopping Theorem II) *Let* $(X_t, t \geq 0)$ *be a uniformly integrable sub-martingale. Let* $S \leq T$ *be stopping times (not necessarily bounded). Then*

$$\mathbf{E}\{X_T|\mathcal{F}_S\} \geq X_S, \qquad \mathbf{E}(X_\infty|\mathcal{F}_T) \geq X_T.$$

If $\{X_t, t \geq 0\}$ *is furthermore a uniformly integrable martingale, then*

$$\mathbf{E}\{X_\infty|\mathcal{F}_S\} = X_S, \qquad \mathbf{E}\{X_\infty|\mathcal{F}_S\} = X_S.$$

Proof We only prove the case of a martingale (the sub-martingale case is left as an exercise).

We only need to prove $\mathbf{E}(X_S) = \mathbf{E}(X_\infty)$ for any stopping time S, see the proof of Proposition 4.13 with T replaced by ∞.

By Theorem 4.15, for all $t > 0$,

$$X_t = \mathbf{E}(X_\infty|\mathcal{F}_t).$$

Let $A_k^n = \{S \in [\frac{k}{2^n}, \frac{k+1}{2^n})\}$ and $S_n = \sum \frac{k}{2_n} \mathbf{1}_{A_k^n}$. Let $A \in \mathcal{F}_{S_n}$, Then

$$\mathbf{E}(X_{S_n} \mathbf{1}_{A_k^n} \mathbf{1}_A) = \mathbf{E}(X_{\frac{k}{2^n}} \mathbf{1}_{A_k^n} \mathbf{1}_A) = \mathbf{E}(X_\infty \mathbf{1}_{A_k^n} \mathbf{1}_A).$$

Summing over all k,
$$\mathbf{E}(X_{S_n}\mathbf{1}_A) = \mathbf{E}(X_\infty \mathbf{1}_A). \tag{4.1}$$
In particular for any $B \in \mathcal{F}_S \subset \mathcal{F}_{S_n}$,
$$\mathbf{E}(X_{S_n}\mathbf{1}_A\mathbf{1}_B) = \mathbf{E}(X_\infty \mathbf{1}_A\mathbf{1}_B).$$
Hence $X_{S_n}\mathbf{1}_B = \mathbf{E}(X_\infty \mathbf{1}_B|\mathcal{F}_{S_n})$. Thus $\{X_{S_n}\mathbf{1}_B, n \in \mathcal{N}\}$ is u.i. Taking $n \to \infty$ we see that,
$$\mathbf{E}(X_S\mathbf{1}_B) = \lim_{n\to\infty} \mathbf{E}(X_{S_n}\mathbf{1}_B) = \mathbf{E}(X_\infty \mathbf{1}_B).$$
This proves that $\mathbf{E}(X_\infty|\mathcal{F}_S) = X_S$. □

Corollary 4.17 *(1) If $(M_t, 0 \le t \le a)$ is a right continuous martingale then $\{M_T : T$ is a stopping time, $T \le a\}$ is uniformly integrable.*

(2) If $(M_t, 0 \le t \le \infty)$ is a uniformly integrable right continuous martingale then $\{M_T : T$ any stopping time$\}$ is uniformly integrable.

Proof (1) $M_t = \mathbf{E}(M_a|\mathcal{F}_t)$ and (2) $M_T = \mathbf{E}\{M_\infty|\mathcal{F}_T\}$. □

4.9 Martingale Inequalities (Lecture 14-15)

The inequalities in this section are important. The proofs are straight forward and I only cover the proof of the maximal inequality in the lecture.

Lemma 4.18 (Maximal Inequality) *If (X_k) is a sub-martingale (or a martingale),*
$$\lambda P\left(\max_{0\le k\le N} X_k \ge \lambda\right) \le \mathbf{E}\left(X_N \mathbf{1}_{\{\max_{0\le k\le N} X_k \ge \lambda\}}\right).$$

Proof This follows by letting $T = \inf\{k : X_k \ge \lambda\}$, and set $T = N$ if $T \ge N$. Then T is a bounded stopping time. By the optional stopping theorem,
$$\begin{aligned}\mathbf{E}(X_N) \ge \mathbf{E}(X_T) &= \mathbf{E}\left(X_T \mathbf{1}_{\{\max_{0\le k\le N} X_k \ge \lambda\}}\right) + \mathbf{E}\left(X_T \mathbf{1}_{\{\max_{0\le k\le N} X_k < \lambda\}}\right)\\ &\ge \lambda P\left(\max_{0\le k\le N} X_k \ge \lambda\right) + \mathbf{E}\left(X_N \mathbf{1}_{\{\max_{0\le k\le N} X_k < \lambda\}}\right).\end{aligned}$$

It follows that,
$$\lambda P\left(\max_{0\le k\le N} X_k \ge \lambda\right) \le \mathbf{E}\left(X_N \mathbf{1}_{\{\max_{0\le k\le N} X_k \ge \lambda\}}\right).$$
□

Lemma 4.19 *Suppose that (X_k) is a martingale or a positive sub-martingale. Let $X^* = \sup_{0\le k\le n} |X_k|$. Then*

$$\mathbf{E} \sup_{0\le k\le n} |X_k|^p \le \left(\frac{p}{p-1}\right)^p \mathbf{E}|X_k|^p.$$

Proof If (X_k) is a martingale, $|X_k|$ is a sub-martingale. It is sufficient to work with positive sub-martingales. For any constant C

$$\mathbf{E}(X^* \wedge C)^p = \mathbf{E}\int_0^{X^*\wedge C} pt^{p-1}dt = \int_0^C pt^{p-1}\mathbf{E}\mathbf{1}_{\{t\le X^*\}}dt.$$

Apply the maximal inequality

$$\begin{aligned}&\int_0^C pt^{p-1}\mathbf{E}\mathbf{1}_{\{t\le X^*\}}dt \le \int_0^C pt^{p-2}\mathbf{E}\left(|X_n|\mathbf{1}_{\{X^*\ge t\}}\right)dt\\ &= \mathbf{E}\left(|X_n|\int_0^C pt^{p-2}\mathbf{1}_{\{X^*\ge t\}}dt\right) = \frac{p}{p-1}\mathbf{E}\left(|X_n|(X^*\wedge C)^{p-1}\right).\end{aligned}$$

By Hölder inequality,

$$\mathbf{E}\left(|X_n|(X^*\wedge C)^{p-1}\right) \le \left(\mathbf{E}|X_n|^p\right)^{\frac{1}{p}}\left[\mathbf{E}(X^*\wedge C)^p\right]^{\frac{p-1}{p}}.$$

To summarise we have

$$\mathbf{E}(X^*\wedge C)^p \le \left(\frac{p}{p-1}\right)^p \mathbf{E}|X_n|^p]$$

The required identity follows by taking C to infinity. □

We return to the continuous time stochastic processes. Let $I = [a,b]$, $I = [a,b)$ or $I = [a,\infty)$. Let $(X_t, t\in I)$ be a right continuous martingale or a positive sub-martingale. For $p > 1$, $|X_t|^p$ is a sub-martingale. In particular $\mathbf{E}|X_t|^p$ increases with t. Set

$$X^* = \sup_{t\in I}|X_t|.$$

If (B_t) is a Brownian motion, then

$$P(\sup_{s\le t} B_s \ge a) = 2P(B_t\ge a) = P(|B_t|\ge a) \le \frac{1}{a^p}\mathbf{E}|B_t|^p.$$

We do not study this equality in this course, and will instead study the L^p inequality below.

Proposition 4.20 *Let $(X_t, t \in I)$ be a right continuous martingale or a positive sub-martingale. Let $t \in I$, an interval.*

(1) Maximal Inequality. For $p \geq 1$,

$$P\Big(\sup_{t\in I} |X_t| \geq \lambda\Big) \leq \frac{1}{\lambda^p} \sup_{t\in I} \mathbf{E}|X_t|^p, \qquad \lambda > 0$$

(2) Doob's L^p inequality. Let $p > 1$. Then

$$\mathbf{E}\left(\sup_{t\in I} |X_t|^p\right) \leq \left(\frac{p}{p-1}\right)^p \sup_{t\in I} \mathbf{E}|X_t|^p.$$

Proof If (X_k) is a martingale or a positive sub-martingale, for $p \geq 1$, $(|X_k|^p)$ is a sub-martingale, $\sup_{0\leq k\leq n} \mathbf{E}|X_k|^p = \mathbf{E}|X_n|^p$. The maximal inequality for a finite index set extends to stochastic processes with a countable index set. As the values of (X_t) is determined by $(X_t, t \in Q)$, the required inequality follows.

□

Chapter 5

Continuous Local Martingales and The quadratic Variation Process

5.1 Lecture 14-15. Local Martingales

We fixed a filtered probability space $(\omega, \mathcal{F}, \mathcal{F}_t, P)$.

Definition 5.1 Let (X_t) be an $\mathcal{F}_t$-adapted stochastic process. If there exists a sequence of stopping times $\{T_n\}$ with $T_n \le T_m$ for $n \le m$ and $\lim_{n\to\infty} T_n = \infty$ a.s. and the property that for each n, $(X_t^{T_n}\mathbf{1}_{\{T_n>0\}}, t \ge 0)$ is a uniformly integrable martingale, we say that (X_t) is a local martingale and that T_n reduces X.

For $t = 0$, $X_t^{T_n}\mathbf{1}_{\{T_n>0\}} = X_0\mathbf{1}_{\{T_n>0\}}$.

A sample continuous martingale is a local martingale (Take $T_n = n$). A bounded continuous local martingale is a martingale. The following definition comes from [6].

Definition 5.2 (1) A local martingale that is not a martingale is a strictly local martingale.

(2) An n dimensional stochastic process $(X_t^1, \dots, X_t^n)$ is a $\mathcal{F}_t$ local-martingale if each component is a $\mathcal{F}_t$ local-martingale.

(3) A (special/decomposable) semi-martingale is an adapted stochastic process such that $X_t = X_0 + M_t + A_t$ where (X_0) is a random variable, (M_t) is a local martingale and (A_t) a process of finite total variation, $M_0 = A_0 = 0$. The decomposition is called the Doob-Meyer decomposition

(4) A continuous semi-martingale X_t is of the form $X_t = X_0 + M_t + A_t$ where M_t and A_t are continuous.

The terminology semi-martingale was traditionally introduced to define stochastic integrals. If a stochastic process can be decomposed as in (3), it is indeed a semi-martingale in the traditional sense. A process of the form $X_t = X_0 + M_t + A_t$ are traditionally called special or decomposable semi-martingales. We will drop the qualifier 'special' or 'decomposable'.

Remark 5.1 (1) A local martingale (M_t) is a martingale if for each t, and the reducing sequence of stopping times T_n, $\{X_t^{T_n}, n \geq 0\}$ is uniformly integrable. It is a martingale if for all $t > 0$,

$$\{M_T : T \text{ bounded stopping times}, T \leq t\}$$

is uniformly integrable.

(2) A local martingale (M_t) is a martingale If $|M_t| \leq Z$ where $Z \in L^1$. In particular a bounded local martingale is a martingale.

(3) If X_t is a martingale then $\mathbf{E}X_t = \mathbf{E}X_0$. If X_t is a local martingale this no longer holds. Furthermore given any function $m(t)$ of bounded variation there is a local martingale such that $m(t)$ is its expectation process. A local martingale which is not a martingale is called a strictly local martingale, otherwise it is a true martingale, see Elworthy-Li-Yor [6] for discussions related to this.

(4) A stochastic integral, as to be defined in the next chapter, with respect to a continuous local martingale is a martingale.

Theorem 5.1 *Let $(M_t, t \leq T)$ be a continuous local martingale. Let $A = \{\omega : M(\omega)_{TV} < \infty\}$. Then for almost surely all $\omega \in A$, $M_t(\omega) = M_a(\omega)$ for all $t \in [a, b]$.*

Appendix

Proof of Theorem 5.1. This proof is the same as that for Brownian motions. We may assume that $M_0 = 0$. Let $t \leq T$. First let

$$M_{TV}(t, \omega) = \sup_{\Delta} \sum_{j=0}^{N-1} |M_{t_{j+1}}(\omega) - M_{t_j}(\omega)|.$$

where Δ ranges through all partitions $0 = t_0 < t_1 < \cdots < t_N = t$ of $[0, t]$. It is increasing and continuous in t. Let

$$T_n = \inf\{t : M_{TV}(t) \geq n\}.$$

Fix n, write

$$X_t = M_t^{T_n}.$$

Then X_t is bounded by n. Since $M_0 = 0$, $|X_t| = |M_{T_n \wedge t}| \leq |M_{T_n \wedge t} - M_0| \leq M_{TV}(t) \leq n$. Also (X_t) is a martingale: For a martingale, $\mathbf{E}X_t^2 = \mathbf{E}\sum_{i=0}^{N-1} |X_{t_{i+1}} - X_{t_i}|^2$. Indeed,

$$\begin{aligned}
\mathbf{E}\sum_{i=0}^{N-1} |X_{t_{i+1}} - X_{t_i}|^2 &= \sum_{i=0}^{N-1} \mathbf{E}\left(\mathbf{E}\left\{|X_{t_{i+1}} - X_{t_i}\Big|^2 |\mathcal{F}_{t_i}\right\}\right) \\
&= \sum_{i=0}^{N-1} \mathbf{E}\left(\mathbf{E}\left\{[X_{t_{i+1}}^2 - 2X_{t_{i+1}}X_{t_i} + X_{t_i}^2]\Big|\mathcal{F}_{t_i}\right\}\right) \\
&= \sum_{i=0}^{N-1} \left(\mathbf{E}X_{t_{i+1}}^2 - \mathbf{E}X_{t_i}^2\right) = \mathbf{E}X_t^2
\end{aligned}$$

$$\begin{aligned}
\mathbf{E}X_t^2 &= \mathbf{E}\sum_{i=0}^{N-1} |X_{t_{i+1}} - X_{t_i}|^2 \leq \mathbf{E}\left(\max |X_{t_{i+1}} - X_{t_i}| \sum_{i=0}^{N-1} |X_{t_{i+1}} - X_{t_i}|\right) \\
&\leq \max_{\omega} X_{TV}(\omega)\, \mathbf{E}\left(\max_i |X_{t_{i+1}} - X_{t_i}|\right) \\
&\leq n\mathbf{E}\left(\max_i |X_{t_{i+1}} - X_{t_i}|\right).
\end{aligned}$$

Since X_t is uniformly continuous on $[0, t]$ and bounded, by the dominated convergence theorem, $\mathbf{E}\left(\max_i |X_{t_{i+1}} - X_{t_i}|\right) \to 0$. Hence $\mathbf{E}(X_t^2) = 0$ and $\mathbf{E}(M_t^{T_n}) = 0$. This implies that $M_t = 0$ on $\{t < T_n\}$ for any n. On $\{\omega : M_{TV}(\omega) < \infty\}$, $T_n \to \infty$. We take $n \to \infty$ to see that if

$$P((M_{TV}(\omega)([a, b]) < \infty) > 0,$$

on this set, $M_t(\omega) = M_a(\omega)$ for all $t \in [a, b]$.

$\square$

5.2 The Quadratic Variation Process (Lecture 15-16)

Definition 5.3 Let (X_t) and (Y_t) be two continuous processes. If for any sequence of partitions with $|\Delta_n| \to 0$,

$$\lim_{n\to\infty} \sum_{j=0}^{\infty} (X_{t\wedge t_{j+1}^n} - X_{t\wedge t_j^n})(Y_{t\wedge t_{j+1}^n} - Y_{t\wedge t_j^n})$$

exists in probability, we define the limit to be $\langle X, Y\rangle_t$.

In particular,

$$\langle X, X\rangle_t \stackrel{P}{=} \lim_{n\to\infty} \sum_{j=0}^{\infty} (X_{t\wedge t_{j+1}^n} - X_{t\wedge t_j^n})^2.$$

Theorem 5.2 *For any continuous local martingales (M_t) and (N_t), there exists a unique continuous process $\langle M, N\rangle_t$ of finite variation vanishing at 0 such that $M_tN_t - \langle M, N\rangle_t$ is a continuous local martingale. This process is called the bracket process or the quadratic variation of (M_t) and (N_t).*

Theorem 5.3 *The stochastic process $(\langle M, N\rangle_t)$ has the following properties:*

(1) It is symmetric and bilinear, and

$$\langle M, N\rangle = \frac{1}{4}[\langle M+N, M+N\rangle - \langle M-N, M-N\rangle]. \tag{5.1}$$

(2) $\langle M - M_0, N - N_0\rangle_t = \langle M, N\rangle_t$.

(3) $\langle M\rangle_t \equiv \langle M, M\rangle_t$ is increasing.

(4) If (M_t) is bounded, $M_t^2 - \langle M\rangle_t$ is a martingale.

(5) For any t and any sequence of partitions $\{\Delta_n : 0 \le t_0^n < \cdots < t_{M_n}^n = t\}$ of $[0, t]$ with $|\Delta_n| \to 0$,

$$\lim_{n\to\infty} \sup_{s\le t} \left| \sum_{j=0}^{M_n} (M_{s\wedge t_{j+1}^n} - M_{s\wedge t_j^n})(N_{s\wedge t_{j+1}^n} - N_{s\wedge t_j^n}) - \langle M, N\rangle_s \right| = 0.$$

The convergence is in probability.

Proposition 5.4 • *If A_t is continuous process of finite variation and X_t is a continuous semi-martingale then*

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

- *If $X_t = M_t + A_t$ and $Y_t = N_t + C_t$ be two continuous semi-martingales with local martingale parts M and N,*

$$\langle X, Y\rangle_t = \langle M, N\rangle_t.$$

Proposition 5.5 *Let T be a stopping time, M and N are local continuous martingales, then*

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

Proof If (M_t) is bounded, then (M_t^T) is a martingale and $(M_t^T)^2 - \langle M^T, M^T\rangle_t$ is a martingale,. Since $(M_t^T)^2 = (M^2)_{T\wedge t}$, we see that

$$\langle M^T, M^T\rangle_t = \langle M, M\rangle_{T\wedge t}.$$

For non-bounded local martingales, take a localising sequence T_n such that $(M_t^{T_n})$ is bounded. Then

$$\langle M^{T_n\wedge T}, M^{T_n\wedge T}\rangle_t = \langle M, M\rangle_{T\wedge T_n\wedge t}.$$

Thus $(M_t^{T\wedge T_n})^2 - \langle M, M\rangle_{T\wedge T_n\wedge t}$ is amartingale. Hence $(M_t^T)^2 - \langle M, M\rangle_{T\wedge t}$ is a local martingale which means that $\langle M^T, M^T\rangle = \langle M, M\rangle^T$.

Now

$$M_t^T N_t^T - M_0 N_0 - \frac{1}{4}[\langle M+N, M+N\rangle^T + \langle M-N, M-N\rangle^T$$

is a local martingale. Hence $\langle M^T, N^T\rangle = \langle M, N\rangle^T$. Similarly $\langle M, N^T\rangle^T = \langle M^T, N^T\rangle = \langle M, N\rangle^T$. It follows that $\langle M, N^T\rangle = \langle M, N\rangle$. □

Proposition 5.6 *If M is a continuous local martingale then $\langle M, M\rangle_t = 0$ if and only if $M_s = M_0$ for $s \in [0, t]$.*

Proof Assume that $\langle M, M\rangle_t = 0$. We first suppose that M_t is bounded and that $M_0 = 0$. Let $s \le t$. Then $\mathbf{E}(M_s - M_0)^2 = \mathbf{E}(M_s)^2 - \mathbf{E}(M_0)^2 = \mathbf{E}\langle M, M\rangle_s = 0$. Then $M_s = M_0$ for all $s \le t$. Otherwise let T_n be a reducing sequence of stopping times then $M_t^{T_n} - M_0 = 0$ almost surely for each n. Take $n \to \infty$ to complete the proof. □

See Revuz-Yor, Proposition 1.13.

5.3 Local Martingale Inequality and Lévy's Martingale Characterization Theorem. (Lecture 18)

Let $I = [0, T]$ if $T \in \mathbf{R}_+$ or $I = [0, \infty)$. Let $X^* = \sup_{t \in I} |X_t|$. In this section we state, without proof, some important theorems.

Theorem 5.7 *[Burkholder-Davis-Gundy Inequality] For every $p > 0$, there exist universal constants c_p and C_p such that if (M_t) is continuous local martingales with $M_0 = 0$,*

$$c_p \mathbf{E}\Big(\langle M, M\rangle_T\Big)^{\frac{p}{2}} \le \mathbf{E}(\sup_{t \in I} |M_t|)^p \le C_p \mathbf{E}\Big(\langle M, M\rangle_T\Big)^{\frac{p}{2}}.$$

Remark 5.2 Let (M_t) is a continuous local martingale with $M_0 = 0$. If $\sup_{t<\infty} M_t \in L^1$ then (M_t) is a martingale.

Let τ be a stopping time, note that

$$\sup_{t<\infty} \sup_{\tau} |M_{t\wedge\tau}|^p \le \sup_{t<\infty} |M_t|^p, \quad \sup_{\tau} |M_\tau|^p \le \sup_{t<\infty} |M_t|^p,$$

Theorem 5.8 *[Lévy's martingale characterization Theorem] An $\mathcal{F}_t$ adapted continuous real valued stochastic process B_t vanishing at 0 is a standard $\mathcal{F}_t$-Brownian motion if and only if (B_t) is an $\mathcal{F}_t$-martingale with quadratic variation t.*

Theorem 5.9 *Let T be a finite stopping time. Then $(B_{T+s} - B_T, s \ge 0)$ is a Brownian motion.*

Definition 5.4 An n dimensional stochastic process $(X_t^1, \dots, X_t^n)$ is a $\mathcal{F}_t$ local-martingale if each component is a $\mathcal{F}_t$ local-martingale.

Multi dimensional version:

Theorem 5.10 *[Lévy's Martingale Characterization Theorem] An $(\mathcal{F}_t)$ adapted sample continuous stochastic process (B_t) in $\mathbf{R}^d$ vanishing at 0 is a $(\mathcal{F}_t)$-Brownian motion if and only if each (B_t) is a $(\mathcal{F}_t)$ local martingale and $\langle B^i, B^j\rangle_t = \delta_{i,j} t$.*

Theorem 5.11 (Dambis, Dubins-Schwartz) *Let $\mathcal{F}_t$ be a right continuous filtration. Let (M_t) be a continuous local martingale vanishing at 0 such that $\langle M, M\rangle_\infty = \infty$. Define*

$$T_t = \inf\{s : \langle M, M\rangle_s > t\}.$$

Then M_{T_t} is an $\mathcal{F}_{T_t}$ Brownian motion and $M_t = B_{\langle M,M\rangle_t}$ a.s..

The condition on the bracket assures that the time change T_t is almost surely finite for all t. Apply Lévy's Characterization Theorem, Theorem **??**, for Brownian motions.

5.3.1 Appendix

The following proof use Itô formulas. **Proof** of Theorem 5.8. If (B_t) is an $\mathcal{F}_t$ BM we already know that it is a martingale with $\mathbf{E}(B_t)^2 = t$ a martingale. Suppose that (B_t) is a martingale with quadratic variation t. Let a be a real number, We apply Itoˆ's formula $(x, y) \mapsto e^{iax+\frac{a^2}{2}y}$ and the stochastic process (B_t, t):

$$e^{iaB_t-\frac{a^2t}{2}} = 1 + ia\int_0^t e^{iB_s-\frac{s}{2}}dB_s + \frac{a^2}{2}\int_0^t e^{iB_s-\frac{s}{2}}ds + \frac{1}{2}(ia)^2\int_0^t e^{iB_s-\frac{s}{2}}ds.$$

Then

$$e^{iaB_t-\frac{a^2t}{2}} = 1 + ia\int_0^t e^{iB_s-\frac{s}{2}}dB_s$$

and $e^{iaB_t-\frac{a^2t}{2}}$ is a martingale. This means that

$$\mathbf{E}\{e^{ia(B_t-B_s)}|\mathcal{F}_s\} = e^{-\frac{a^2(t-s)}{2}}.$$

Let $\phi(a) := \mathbf{E}\{e^{ia(B_t-B_s)}|\mathcal{F}_s\}$. Since $\phi(a)$ is non-random,

$$\mathbf{E}e^{ia(B_t-B_s)} = \mathbf{E}\{e^{ia(B_t-B_s)}|\mathcal{F}_s\}.$$

Thus $B_t - B_s$ is independent of $\mathcal{F}_s$ and is a Gaussian random variable with distribution $N(0, t-s)$. □

Proof of Theorem 5.10 Proof of the 'if part': for any $\lambda \in \mathbf{R}^d$, $Y_t = \langle \lambda, X_t \rangle_{\mathbf{R}^d} = \sum \lambda_j X_t^j$ is a local martingale with bracket $|\lambda|^2 t$. The exponential martingale $\exp^{i\langle\lambda,X_t\rangle+\frac{1}{2}|\lambda|^2t}$ is a martingale as it is bounded on any compact time interval, hence

$$\mathbf{E}\{\exp^{i\langle\lambda,X_t-X_s\rangle}|\mathcal{F}_s\} = e^{-\frac{1}{2}|\lambda|^2(t-s)}.$$

This is sufficient to show that $X_t - X_s$ is independent of $\mathcal{F}_s$ and

$$\mathbf{E}\exp^{i\langle\lambda,X_t-X_s\rangle} = e^{-\frac{1}{2}|\lambda|^2(t-s)},$$

which implies that $X_t - X_s \sim N(0, t-s)$.

Proof of 'only if part'. First for $s < t$,

$$\mathbf{E}\{X_t^i|\mathcal{F}_s\} = \mathbf{E}\{X_t^i - X_s^i + X_s^i|\mathcal{F}_s\} = \mathbf{E}(X_t^i - X_s^i) + X_s^i = X_s^i$$

and each X_t is a martingale. For $s < t$,

$$\mathbf{E}\{(X_t^i)^2 - (X_s^i)^2 | \mathcal{F}_s\} = \mathbf{E}\{(X_t^i - X_s^i)^2 | \mathcal{F}_s\} = \mathbf{E}(X_t^i - X_s^i)^2 = t - s.$$

Then $\langle X^i, X^i \rangle_t = t$ and the bracket of independent Brownian motions is zero. $\square$

5.4 The Hilbert space of L^2 bounded martingale (Lecture 18)

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered probability space, $(\mathcal{F}_t)$ satisfies the usual right continuity and completion assumptions. If $f \in L^1$, $Y_t = \mathbf{E}(f|\mathcal{F}_t)$ can be chosen in such a way that (Y_t) is right continuous and cádlág.

Theorem 5.12 *Let $\mathbb{H}^2$ be the space of L^2 bounded right continuous martingales. Let*

$$\|M\|_{\mathbb{H}^2} = \sqrt{\mathbf{E}(M_\infty)^2} = \lim_{t \to \infty} \sqrt{\mathbf{E}(M_t)^2}.$$

Then $\mathbb{H}^2$ is a Hilbert space and the space H^2 of continuous L^2 bounded martingales is closed in $\mathbb{H}^2$.

Proof Let $(M_t) \in \mathbb{H}^2$. Let $M_\infty = \lim_{t \to \infty} M_t$ exists and $M_t = \mathbf{E}(M_\infty | \mathcal{F}_t)$. By Doob's L^2 inequality,

$$\mathbf{E} \sup_t (M_t)^2 \le 4 \sup_t \mathbf{E}(M_t)^2 < \infty.$$

Hence $(M_t)^2$ is uniformly integrable and $(M_\infty)^2 \in L^2$. There is a one to one correspondence between $\mathbb{H}^2$ and $L^2(\Omega, \mathcal{F}, P)$.

Since $\langle M, M \rangle_t$ increases with t and is positive, By Fatou's lemma,

$$\mathbf{E}\langle M, M \rangle_\infty = \mathbf{E} \lim_{t \to \infty} \langle M, M \rangle_t \le \lim_{t \to \infty} \mathbf{E}\langle M, M \rangle_t = \sup_t \mathbf{E}(M_t)^2 < \infty.$$

This means that $(M_t)^2 - \langle M, M \rangle_t$ is an uniformly integrable martingale with limit as $t \to \infty$. In particular,

$$\mathbf{E}(M_\infty)^2 = \mathbf{E}\langle M, M \rangle_\infty.$$

To see that H^2 is closed, let $M_t^{(n)} \subset H^2 \to (M_t)$ in $\mathbb{H}^2$. By Doob's L^2 inequality (Proposition 4.20 (2)) ,

$$\mathbf{E} \sup_t (M_t^n - M_t)^2 \le 4 \sup_t \mathbf{E}(M_t^n - M_t)^2 \to 0.$$

In particular there is a subsequence $\{M^{n_k}\}$ such that

$$\sup_t (M_t^{n_k} - M_t)^2 \to 0$$

almost surely. This implies that (M_t) is continuous almost surely and H^2 is closed in $\mathbb{H}^2$. □

Theorem 5.13 *If (M_t) is a continuous local martingale with $M_0 = 0$, the following statements are equivalent.*

1. *(M_t) is L^2 bounded.*
2. *$\mathbf{E}\langle M, M\rangle_\infty < \infty$.*

If either holds, $(M_t)^2 - \langle M, M\rangle_t$ is an uniformly integrable martingale and

$$\mathbf{E}(M_t)^2 = \mathbf{E}\langle M, M\rangle_t, \qquad \mathbf{E}(M_\infty)^2 = \mathbf{E}\langle M, M\rangle_\infty.$$

It follows if $f \in H^2$,

$$\|f\|_{H^2} = \mathbf{E}\langle M, M\rangle_\infty. \tag{5.2}$$

Chapter 6

Stochastic Integration

We aim to define the stochastic integral $\int_0^t f_s dX_s$ where $X_s = X_0 + M_s + A_s$ is a right continuous semi-martingale (Definition 5.2) and (f_t) is a left continuous measurable process. In this chapter we we assume that (M_s) is continuous, in which case the bracket process is also continuous.

We take the index set to be $[0, \infty)$ or $[0, T]$ for a real number T. For simplicity, in most of the cases, we deal with only $[0, \infty)$. The $[0, T]$ case can be either dealt with analogously or follows trivially by a localization $\mathbf{1}_{[0,T]}$. This is important as Brownian motion is not L^2 bounded on $\mathbf{R}_+$, it is L^2 bounded on any bounded time interval.

6.1 Introduction (Lectures 18-19)

Let $\mathcal{E}$ denotes the collection of elementary processes:

$$\mathcal{E} = \left\{ (K_t) : K_t(\omega) = K_{-1}(\omega)\mathbf{1}_{\{0\}}(t) + \sum_{i=0}^{\infty} K_i(\omega)\mathbf{1}_{(t_i, t_{i+1}]}(t), \right\}$$

where $0 = t_0 < \cdots < t_n < \ldots$ is any sequence of positive numbers increasing to infinity, $K_{-1} \in \mathcal{F}_0, K_i \in \mathcal{F}_{t_i}$, $|K_i|$ are bounded. Let $K \in \mathcal{E}$ and let (X_s) be a stochastic process. We previously defined elementary integration

$$(I_M)_t(K) := (K \cdot M)_t = H_0 K_0 + \sum_{i=1}^{\infty} K_i (M_{t_{i+1} \wedge t} - M_{t_i \wedge t}).$$

We seek a class of stochastic processes (f_s) with the property that there exists a sequence of stochastic processes $H^n \in \mathcal{E}$ with H^n converges to f (in some sense), and $\int_0^t H_n(s)dX_s$ converges (in some sense) to a limit, the limit will be a candidate for the stochastic integral $\int_0^t f_s dX_s$. This is a continuity property. The standard convergence used for stochastic processes is the uniform convergence on compacta in probability (u.c.p.). In other words, a sequence of stochastic processes f_n is said to converge in the u.c.p. topology to a stochastic process f if for all $t \geq 0$,

$$\lim_{n\to\infty} \sup_{0\leq s\leq t} |f_n(s) - f(s)| = 0$$

in probability.

Let α and β be numbers satisfying that $\alpha + \beta > 1$. It is not possible to find a universal bi-linear operation from the space of 'C^α regular functions' and the space of 'C^β regular functions' to the space of 'C^β regular functions' such that the linear operation is continuous. Hence a stochastic integral cannot be, in general, interpreted as a pathwise integral.

Question

(1) For which class of stochastic processes (X_t), the map $H \in \epsilon \to \int_0^t H_s dX_s$ is continuous from $\mathcal{E}$ to the set $\mathbb{D}$ of cádlág adapted processes?

(2) What is the closure of $\mathcal{E}$ in D with respect to the u.c.p. topology? And what are suitable step processes which can be used in the continuity argument?

The search for an answer to Question leads to the coinage of the terminology 'semi-martingales'. If $X_t = X_0 + M_t + A_t$, then it is a 'semi-martingale'. From now on we will be concerned only with continuous special semi-martingales and we omit the word special. The second question will lead to the terminology of simple predictable stochastic processes,

$$S = \left\{ (K_t) : K_t(\omega) = K_{-1}(\omega)\mathbf{1}_{\{0\}}(t) + \sum_{i=1}^n K_i(\omega)\mathbf{1}_{(T_i,T_{i+1}]}(t), \right\}$$

where $0 = T_0 < T_1 < \dots$ is an increasing sequence of stopping times converging to infinity, with K_i finite valued and $K_i \in \mathcal{F}_{T_i}$, $K_{-1} \in \mathcal{F}_0$. This set is dense in $\mathbb{L}$, the set of left continuous right limit adapted stochastic processes in the u.c.p. topology.

We will be dealing with the simpler case where the integrators are sample continuous. To begin with, we will use a Hilbert space construction and a suitable L^2 norm for the convergence.

6.1.1 Integration w.r.t. Stochastic Processes of Finite Variation

If $(A_s, s \geq 0)$ is a right continuous function of finite variation with $A_0 = 0$. There is an associated Borel measure μ_A on $[0,\infty)$ determined by $\mu_A((c,d]) = A(d) - A(c)$. Note that $\mu(\{d\}) = A(d) - A(d-)$. If A is continuous the measure does not charge a singleton.

Write

$$A_s = \frac{A_{TV}(s) + A_s}{2} - \frac{A_{TV}(s) - A_s}{2}.$$

where A_{TV} is the total variation process. Recall a signed measure μ decomposed as difference of two positive measures: $\mu = \mu^+ - \mu^-$. For the Radon measure μ_A, $|\mu_A|$ is the measure determined by A_{TV}. Let $f : \mathbf{R}_+ \to \mathbf{R}$ be integrable with integral denoted by $\int_{[0,\infty)} f_s dA_s$. Let $\int_0^t f_s dA_s = \int_{[0,\infty)} \mathbf{1}_{(0,t]}(s) f_s d\mu_A(s)$. If $f : \mathbf{R}_+ \to \mathbf{R}$ is left continuous,

$$\int_0^t f_s dA_s = \lim_{|\Delta_n| \to 0} \sum_j f(t_j^n) \left(A(t_{j+1}^n) - A(t_j^n) \right).$$

We may allow f_s and A_s random. The above procedure holds for each ω for which $A_s(\omega)$ is of finite variation. There is however the added complication of measurability. We assume that f is progressively measurable (by progressive measurability we include the assumption that f is universally measurable, i.e. $f : \mathbf{R}_+ \otimes \Omega \to \mathbf{R}$ is measurable with respect to $\mathcal{B}(\mathbf{R}_+) \otimes \mathcal{F}_\infty$ and A is a right continuous finite variation process (recall, in particular, A_s is adapted). Then $\int_0^t f_s(\omega) dA_s(\omega)$ is a process of finite variation and is right continuous. The integral is furthermore continuous if (A_s) is sample continuous.

Recall that $\langle M, M \rangle$ correspond to a positive measure and $\langle M, N \rangle$ a signed measure, written as $\mu^+ - \mu^-$ where μ^+, μ^- are positive measures. By $|\langle M, N \rangle|$ we mean the measure corresponds to $\mu^+ + \mu^-$.

Lemma 6.1 *Let $s \leq t$, we define $\langle M, N \rangle_t^s := \langle M, N \rangle_t - \langle M, N \rangle_s$.*

$$\langle M, N \rangle_t^s \leq \sqrt{\langle M, M \rangle_t - \langle M, M \rangle_s} \sqrt{\langle N, N \rangle_t - \langle N, N \rangle_s}.$$

Proof For any a, $\langle M - aN \rangle_t \geq 0$. This means $\langle M, M \rangle_t + a^2 \langle N, N \rangle_t \geq 2a \langle M, N \rangle_t$. Take $a = \sqrt{\frac{\langle M, M \rangle_t}{\langle N, N \rangle_t}}$ to see that $\langle M, N \rangle_t \leq \sqrt{\langle M, M \rangle_t \langle N, N \rangle_t}$. A similar proof shows that for $s < t$:

$$\langle M, N \rangle_t - \langle M, N \rangle_s \leq \sqrt{\langle M, M \rangle_t - \langle M, M \rangle_s} \sqrt{\langle N, N \rangle_t - \langle N, N \rangle_s}.$$

□

Let H_s, K_s be measurable functions by which we mean they are Borel measurable functions from $(\mathbf{R}_+ \times \Omega, \mathcal{F}_\infty \otimes \mathcal{B}(\mathbf{R}_+)$ to $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$. Approximating them by elementary functions leads to the following theorem:

Theorem 6.2 *Let (M_t) and (N_t) be two continuous local martingales. Let (H_t) and (K_t) be measurable processes. Then for t $\le \infty$,*

$$\int_0^t |H_s||K_s| \; d|\langle M, N\rangle|_s \le \sqrt{\int_0^t |H_s|^2 d\langle M, M\rangle_s}\sqrt{\int_0^t |K_s|^2 d\langle N, N\rangle_s}, \qquad a.s.$$

The inequality states in particular that the left hand side is finite if the right hand side is. If furthermore $H \in L^1(d\langle M, M\rangle_s)$ and $K \in L^1(d\langle N, N\rangle_s)$,

$$\left|\int_0^t H_s K_s \; d\langle M, N\rangle_s\right| \le \sqrt{\int_0^t |H_s|^2 d\langle M, M\rangle_s}\sqrt{\int_0^t |K_s|^2 d\langle N, N\rangle_s}, \qquad a.s.$$

Proof This is a schematic proof. Let (H_s) and (K_s) be from $\mathcal{E}$, elementary processes. Let $0 = t_1 < \cdots < t_{N+1}$ be a partition such that on each sub-interval, both $H_s(\omega)$ and $K_s(\omega)$ are constant in s. We write, for $H_0, K_0 \in \mathcal{F}_0, H_i, K_i \in \mathcal{F}_{t_i}$,

$$H_t(\omega) = H_0(\omega)\mathbf{1}_{\{0\}}(t) + \sum_{i=1}^N H_i(\omega)\mathbf{1}_{(t_i, t_{i+1}]}(t),$$

and

$$K_t(\omega) = K_0(\omega)\mathbf{1}_{\{0\}}(t) + \sum_{i=1}^N K_i(\omega)\mathbf{1}_{(t_i, t_{i+1}]}(t),$$

Then

$$\begin{aligned}
&\left|\int_0^t H_s(\omega)K_s(\omega)d\langle M, N\rangle_s(\omega)\right| \\
&= \left|\sum_i H_i(\omega)K_i(\omega)\langle M, N\rangle_{t_{i+1}}^{t_i}(\omega)\right| \le \sum_i |H_i(\omega)||K_i(\omega)||\langle M, N\rangle_{t_{i+1}}^{t_i}(\omega)| \\
&\le \sqrt{\sum_i |H_i(\omega)|^2\langle M, M\rangle_{t_{i+1}}^{t_i}(\omega)}\sqrt{\sum_i |K_i(\omega)|^2\langle N, N\rangle_{t_{i+1}}^{t_i}(\omega)} \\
&= \left(\int_0^\infty (H_s(\omega))^2 d\langle M, M\rangle_s(\omega)\right)^{\frac{1}{2}} \left(\int_0^\infty (K_s(\omega))^2 d\langle N, N\rangle_s(\omega)\right)^{\frac{1}{2}}.
\end{aligned}$$

Take appropriate limit to see the second inequality holds.

Let $\tilde{H}_s = H_s \text{sign}(H_s K_s) \frac{d\langle M,N\rangle_s}{|d\langle M,N\rangle_s|}$, we see that

$$\int_0^t |H_s||K_s| \; d|\langle M,N\rangle|_s = \int_0^t \tilde{H}_s K_s d\langle M,N\rangle_s$$

and apply the second inequality we see the first inequality holds for all bounded measurable functions (H_s) and (K_s). If they are not bounded, we take a sequence of cut-off functions for H_s and K_s to see that first inequality always holds. They may however be infinite. □

Apply Hölder Inequality to the above inequality to obtain the following.

Corollary 6.3 *[Kunita-Watanabe Inequality] For* $t \le \infty$, *and* $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$,

$$\begin{aligned} &\mathbf{E} \int_0^t |H_s||K_s||d\langle M,N\rangle|_s \\ &\le \left(\mathbf{E} \left(\int_0^t |H_s|^2 d\langle M,M\rangle_s \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} \left(\mathbf{E} \left(\int_0^t |K_s|^2 d\langle N,N\rangle_s \right)^{\frac{q}{2}} \right)^{\frac{1}{q}} \end{aligned}$$

6.2 Space of Integrands (Lecture 18-19)

Definition 6.1 Let E be a metric space. A stochastic process $X : I \times \Omega \to E$ is progressively measurable if

(1) $X : I \times \Omega \to E$ is measurable

(2) for each $t > 0$, $X : [0,t] \times \Omega \to E$ is a measurable map with respect to the product σ-algebra $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$.

Proposition 6.4 *Left or right continuous adapted processes are progressively measurable.*

Proof exercise. □

Let H^2 be the space of L^2 bounded continuous martingales.

Definition 6.2 For $M \in H^2$, define $L^2(M)$ to be the space of progressively measurable stochastic process (f_t) such that

$$\|f\|^2_{L^2(M)} := \mathbf{E} \int_0^\infty (f_s)^2 d\langle M,M\rangle_s < \infty.$$

$$L^2(M) = \{(f_t) : \; f \text{ is progressively measurable}, \mathbf{E} \int_0^\infty (f_s)^2 d\langle M,M\rangle_s < \infty.\}.$$

For $H, K \in L^2(M)$, we define

$$\langle H, K\rangle_{L^2(M)} = \mathbf{E}\int_0^\infty H_s K_s d\langle M\rangle_s.$$

Let P_M be the measure on $\mathbf{R}_+ \times \Omega$ determined by $\Gamma \in \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{F}_\infty$,

$$P_M(\Gamma) = \mathbf{E}\int_0^\infty \mathbf{1}_\Gamma d\langle M, M\rangle_s.$$

Then $L^2(M) = L^2(\mathbf{R}_+ \times \Omega, \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{F}_\infty, P_M)$. We see that $L^2(M)$ is a Hilbert space. By standard estimate, c.f. Kunita-Watanabe inequality (Corollary 6.3),

$$|\langle H, K\rangle_{L^2(M)}| \le \|H\|_{L^2(M)} \|K\|_{L^2(N)}.$$

Let N be an integer and $0 = t_1 < \cdots < t_{N+1}$. Let

$$K_t(\omega) = K_0(\omega)\mathbf{1}_{\{0\}}(t) + \sum_{i=1}^N K_i(\omega)\mathbf{1}_{(t_i, t_{i+1}]}(t),$$

where $K_0 \in \mathcal{F}_0, K_i \in \mathcal{F}_{t_i}$. Let $M \in H^2$ then

$$\left|\mathbf{E}\int_0^\infty K_s^2 d\langle M, M\rangle_s\right| \le |K|_\infty \mathbf{E}\Big(\langle M, M\rangle_{t_{N+1}}\Big)$$

and $\mathcal{E} \subset \cap_{M \in H^2} L^2(M)$. In other words, every elementary process belongs to $L^2(M)$ for any $M \in H^2$.

Proposition 6.5 *The set of elementary processes are dense in* $L_2(M)$.

Proof We prove the case when $f \in L^2(M)$ is left continuous. First assume f is bounded and let

$$g_n(s, \omega) = f_0(\omega)\mathbf{1}_{\{0\}}(s) + \sum_{j \ge 1} f_{\frac{j}{2^n}}(\omega)\mathbf{1}_{(\frac{j}{2^n} \le s < \frac{j+1}{2^n}]}.$$

Note that $f_{\frac{j}{2^n}}$ is $\mathcal{F}_{j/2^n}$ measurable. Since f is left continuous and bounded, $|g_n|_\infty \le |f|_\infty$, by the dominated convergence theorem, $g_n \to f$ in $L^2(M)$. If f is not bounded, let $f_n(s) = f_s \mathbf{1}_{|f_s| \le n}$. Then

$$\|f_n - f\|_{L^2(M)}^2 \le \mathbf{E}\int_0^\infty f_s^2(\omega)\mathbf{1}_{\{(s,\omega): |f(s,\omega)| \ge n\}} d\langle M, M\rangle_s(\omega) \stackrel{(n \to \infty)}{\to} 0.$$

□

Note to the Proof. If f is only assumed to be progressively measurable, the proof is as below. Let $f \in L^2(M)$ be a function orthogonal to $\mathcal{E}$. Hence for any $s < t$, $K \in \mathcal{F}_s$,

$$0 = \langle f, K\mathbf{1}_{(s,t]}\rangle_{L^2(M)} = \mathbf{E}\int_0^\infty f_r K\mathbf{1}_{(s,t]} d\langle M, M\rangle_r = \mathbf{E}K\int_s^t f_r d\langle M, M\rangle_r.$$

In particular for any $A \in \mathcal{F}_s$,

$$\mathbf{E}\left(\int_0^t f_r d\langle M, M\rangle_r \mathbf{1}_A\right) = \mathbf{E}\left(\int_0^s f_r d\langle M, M\rangle_r \mathbf{1}_A\right).$$

and $(\int_0^t f_r d\langle M, M\rangle_r, r \geq 0)$, which is integrable by the Kunita-Watanabe inequality (Corollary 6.3), is a continuous martingale and is also a finite variation process. Hence $f = 0$.

We also use the notation: $K \cdot M = \int_0^\cdot K_s dM_s$.

Proposition 6.6 *Let $M \in H^2$ and $K \in \mathcal{E}$. Then $(\int_0^t K_s dM_s, t \geq 0)$ is an L^2 bounded continuous martingale.*

Proof Let $K_t = K_0\mathbf{1}_{\{0\}} + \sum K_i \mathbf{1}_{(t_i, t_{i+1}]}(t)$, then

$$\int_0^t K_r dM_r = \sum_i K_i(M_{t_{i+1}\wedge t} - M_{t_i \wedge t}).$$

Let us consider the ith interval $[t_i, t_{i+1}]$. If $t_i \geq t$,

$$M_{t_{i+1}\wedge t} - M_{t_i\wedge t} = 0.$$

If $t \in (t_j, t_{j+1})$, then

$$M_{t_{j+1}\wedge t} - M_{t_j \wedge t} = M_t - M_{t_j}.$$

We may assume that the summation is from 1 to N and $t_{N+1} = t$ and so

$$\int_0^t K_r dM_r = \sum_{i=1}^N K_i(M_{t_{i+1}} - M_{t_i}).$$

Let $s < t$, we compute

$$\mathbf{E}\left\{\sum_{i=1}^N K_i(M_{t_{i+1}} - M_{t_i}) | \mathcal{F}_s\right\}.$$

Let us analyze the ith interval $I_i = (t_i, t_{i+1}]$. If $t_{i+1} \le s$, then

$$\mathbf{E}\left\{K_i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_s\right\} = K_i(M_{t_{i+1}} - M_{t_i}) = K_i(M_{t_{i+1}\wedge s} - M_{t_i\wedge s}).$$

If $s \le t_i$, then

$$\mathbf{E}\left\{K_i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_s\right\} = \mathbf{E}\left\{K_i\mathbf{E}\left\{(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_s\right\}|\mathcal{F}_{t_i}\right\} = 0 = K_i(M_{t_{i+1}\wedge s} - M_{t_i\wedge s}).$$

If $t_i < s < t_{i+1}$,

$$\mathbf{E}\{K_i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_s\} = K_i\mathbf{E}\{M_{t_{i+1}}|\mathcal{F}_s\} - K_iM_{t_i} = K_i(M_{t_{i+1}\wedge s} - M_{t_i\wedge s}).$$

Summing up the three cases to obtain

$$\mathbf{E}\{(K \cdot M)_t|\mathcal{F}_s\} = \sum_i K_i(M_{t_{i+1}\wedge s} - M_{t_i\wedge s}).$$

and $K \cdot M$ is a martingale. □

Definition 6.3 Denote by H_0^2 the subspace of H^2 whose elements M_t satisfies $M_0 = 0$.

Then H_0^2 is a closed subspace of H^2 and $\sqrt{\mathbf{E}\langle M, M\rangle_\infty} = \sqrt{\mathbf{E}(M_\infty)^2}$ is a norm for H_0^2. See Theorems 5.12 and 5.13.

Proposition 6.7 *Let $M \in H^2$ and $K \in \mathcal{E}$. Then the elementary integral $(\int_0^t K_s dM_s, t \ge 0) \in H^2$. The map*

$$K \in \mathcal{E} \mapsto K \cdot M \in H_0^2$$

is linear and

$$\left\|\int_0^\cdot K_s dM_s\right\|_{H^2} = \|K\|_{L^2(M)}, \qquad \textit{(Itô Isometry).}$$

Proof Linearity is clear. We prove the Itô isometry,

$$\begin{aligned}(\|(K \cdot M)\|_{H^2})^2 &= \mathbf{E}\left(\sum_{i=0}^\infty K_i(M_{t_{i+1}\wedge t} - M_{t_i\wedge t})\right)^2 = \sum_{i=0}^\infty \mathbf{E}\left(K_i^2\left((M_{t_{i+1}\wedge t})^2 - (M_{t_i\wedge t})^2\right)\right)\\ &= \sum_{i=0}^\infty \mathbf{E}\left(K_i^2\mathbf{E}\left\{(M_{t_{i+1}\wedge t})^2 - (M_{t_i\wedge t})^2\Big|\mathcal{F}_{t_i}\right\}\right)\\ &= \sum_{i=0}^\infty \mathbf{E}\left(K_i^2\left(\langle M, M\rangle_{t_{i+1}\wedge t} - \langle M, M\rangle_{t_i\wedge t}\right)\right).\end{aligned}$$

We use the fact that $M_t^2 - \langle M, M\rangle_t$ is a martingale, c.f. Theorem 5.13. The isometry follows from this and the following computation

$$(\|K\|_{L^2(M)})^2 = \mathbf{E}\int_0^\infty (K_s)^2 d\langle M, M\rangle_s = \mathbf{E}\sum_{i=1}^\infty (K_i)^2 \langle M, M\rangle_{t_{i+1}\wedge t}^{t_i\wedge t}.$$

□

Definition 6.4 (and Theorem) If $f \in L^2(M)$ and $\{f_n\} \subset \mathcal{E}$ is a sequence converging to f in $L^2(M)$. Then $\int_0^t f_n dM_s$ exists. We define the limit to be $\int_0^t f_s dM_s$. This limit is independent of the choices of the converging sequence.

Proof Since $f_n \to f$, $\{\int_0^\cdot f_s dM_s\}$ is a Cauchy sequence in H_0^2 which follows from

$$\begin{aligned}
&\| \int_0^\cdot f_n(s)dM_s - \int_0^\cdot f_m(s)dM_s\|_{H^2} = \left\| \int_0^\cdot (f_n(s) - f_m(s))dM_s \right\|_{H^2} \\
&= \mathbf{E}\int_0^\infty (f_n(s) - f_m(s))^2 d\langle M, M\rangle_s.
\end{aligned}$$

The uniqueness follows by the standard argument. Suppose that $g_n \to f$ in $L^2(M)$. By Proposition 6.7,

$$\begin{aligned}
&\left\| \int_0^\cdot f_n(s)dM_s - \int_0^\cdot g_n(s)dM_s \right\|_{H^2} \\
&= \| \int_0^\cdot (f_n(s) - g_n(s))dM_s\|_{H^2} \overset{Proposition 6.7}{=} \|f_n - g_n\|_{L^2(M)} \\
&= \mathbf{E}\int_0^\infty (f_n(s) - g_n(s))^2 d\langle M, M\rangle_s \\
&\le \mathbf{E}\int_0^\infty (g_n(s) - f(s))^2 d\langle M, M\rangle_s + \mathbf{E}\int_0^\infty (f_n(s) - f(s))^2 d\langle M, M\rangle_s.
\end{aligned}$$

□

Exercise 6.1 Let $M \in H^2$ and $K \in \mathcal{E}$. Prove that the elementary integral $\int_0^t K_s dM_s$ satisfies that for any $N \in H^2$,

$$\langle I, N\rangle_t = \int_0^t K_s \, d\langle M, N\rangle_s, \qquad \forall t \ge 0 \tag{6.1}$$

6.3 Lecture 20. Characterization of Stochastic Integrals

All separable infinite dimensional Hilbert spaces are isomorphic (take a set of basis in each space to construct the isometry). Hence $L^2(M) \overset{iso}{=} H_0^2$. Below we construct an explicit isometric map:

$$K \mapsto I(K, M).$$

We will call $I(K, M)$, the Itô integral. This will agree with elementary integrals if $K \in \mathcal{E}$ and denoted by

$$\int_0^t K_s dM_s.$$

Note that the value $\langle M, N\rangle_t$ is unchanged with the transformation $M \mapsto M + c$ where c is a constant.

Theorem 6.8 *Given $M \in H^2$ and $K \in L^2(M)$, there is a unique process $I \equiv I(K, M) \in H^2$, vanishing at 0, such that*

$$\langle I, N\rangle_t = \int_0^t K_s \, d\langle M, N\rangle_s, \qquad \forall N \in H^2, \ \forall t \geq 0. \tag{6.2}$$

Furthermore the map $K \in L^2(M) \mapsto I(K, M) \in H_0^2$ is a linear isometry.

The identity (6.2) holds is equivalent to the following holds,

$$\langle I, N\rangle_\infty = \int_0^\infty K_s \, d\langle M, N\rangle_s, \quad \forall N \in H^2. \tag{6.3}$$

Proof Since the bracket processes $\langle M, N\rangle = \langle M, N - N_0\rangle$, it is sufficient to prove that (6.2) holds for all $N \in H_0^2$.

(a) The uniqueness. Suppose that there are two martingales $I_1, I_2 \in H_0^2$ such that

$$\langle I_1, N\rangle_t = \int_0^t K_s \, d\langle M, N\rangle_s, \qquad \langle I_2, N\rangle_t = \int_0^t K_s \, d\langle M, N\rangle_s.$$

Then $\langle I_1 - I_2, N\rangle_t = 0$ for any $N \in H_0^2$. In particular $\mathbf{E}\langle I_1 - I_2, I_1 - I_2\rangle_\infty = 0$, which implies that $I_1 = I_2$.

(b) The existence.

(b1) Let $N \in H_0^2$. We define a real valued linear map:

$$U: \quad H_0^2 \longrightarrow \mathbf{R}$$
$$U(N) = \mathbf{E}\int_0^\infty K_s \, d\langle M, N\rangle_s.$$

(b2) We apply Kunita-Watanabe inequality:

$$\begin{aligned}|U(N)| &\le \mathbf{E}\left|\int_0^\infty K_s\ d\langle M, N\rangle_s\right| \\ &\le \left(\mathbf{E}\int_0^\infty K_s^2\ d\langle M, M\rangle_s\right)^{\frac{1}{2}}\ \sqrt{\mathbf{E}\langle N, N\rangle_\infty} \\ &\le |K|_{L^2(M)}|N|_{H_0^2} < \infty.\end{aligned} \tag{6.4}$$

This proves that U is a bounded linear operator. By the Riesz Representation Theorem for bounded linear operators, there is a unique element, I, of H_0^2, such that

$$U(N) = \langle I, N\rangle_{H_0^2} = \mathbf{E}\langle I(K, M), N\rangle_\infty.$$

This can be rewritten as:

$$\mathbf{E}\left(\int_0^\infty K_s\ d\langle M, N\rangle_s\right) = \mathbf{E}\left(\langle I(K, M), N\rangle_\infty\right). \tag{6.5}$$

(b3) Define

$$X_t := I_t N_t - \int_0^t K_s d\langle M, N\rangle_s.$$

We prove that (X_t) is a martingale with initial value 0. By the defining property of the bracket process, $\langle I, N\rangle_t = \int_0^t K_s d\langle M, N\rangle_s$, observing that both vanish at 0.

Let τ be any bounded stopping time. We prove that $\mathbf{E}(X_\tau) = 0$. By Theorem 5.13, $I_t N_t - \langle I, N\rangle_t$ is a uniformly integrable martingale, and

$$\begin{aligned}\mathbf{E} I_\tau N_\tau &= \mathbf{E}\langle I, N\rangle_\tau = \mathbf{E}\langle I, N\rangle_{\tau\wedge\infty} = \mathbf{E}\langle I, N^\tau\rangle_\infty \\ &\overset{by (6.5)}{=} \mathbf{E}\int_0^\infty K_s d\langle M, N^\tau\rangle_s = \mathbf{E}\int_0^\infty \mathbf{1}_{s<\tau} K_s d\langle M, N\rangle_s \\ &= \mathbf{E}\int_0^\tau K_s d\langle M, N\rangle_s.\end{aligned}$$

This shows that $\mathbf{E}X_\tau = \mathbf{E}X_0 = 0$ and (X_t) is a martingale.

(c) We show that the map $K \mapsto I(K, M)$ is a linear isometry. The linearity is clear: it follows from the linearity of the bracket $\langle M, N\rangle$ and the linearity of $K \mapsto \int_0^t K_s d\langle M, N\rangle_s$. Take $K, K' \in L^2(M)$. Then,

$$\begin{aligned}
\langle I(K,M), I(K',M)\rangle_{H_0^2} &= \mathbf{E}\left(\int_0^\infty K_s \; d\langle M, I(K',M)\rangle_s\right) \\
&= \mathbf{E}\int_0^\infty K_s d\left(\int_0^s K'_r d\langle M, M\rangle_r\right) \\
&= \mathbf{E}\int_0^\infty K_s K'_s d\langle M, M\rangle_s \\
&= \langle K, K'\rangle_{L^2(M)}.
\end{aligned}$$

The isometry follows.

□

Proposition 6.9 *Let $K \in \mathcal{E}$ and $M \in L^2(M)$. Prove that the integral $I(K,M)$ defined in Theorem 6.8 agrees with the elementary integral. Consequently, for all $K \in L^2(M)$,*

$$I(K,M) = \int_0^t K_s dM_s,$$

where the latter is defined by Definition 6.4.

Proof Exercise. □

6.4 Integration w.r.t. Semi-martingales (Lecture 21)

The identity (6.2) in Theorem 6.8 leads easily to the following.

Proposition 6.10 *Let $H \in L^2(M), K \in L^2(N)$ where $M, N \in H^2$. Then*

$$\left\langle \int_0^\cdot H_s dM_s, \int_0^\cdot K_s dN_s \right\rangle = \int_0^t H_s K_s d\langle M, N\rangle_s.$$

Proposition 6.11 *Let τ be a stopping time. Then for $(M_t) \in H^2$, $(K_t) \in L^2(M)$,*

$$\int_0^{\tau\wedge t} K_s dM_s = \int_0^t K_s dM_s^\tau = \int_0^t K_s \mathbf{1}_{[0,\tau]}(s) dM_s.$$

Proof Take $N \in H^2$. Then for any $t \in [0, \infty]$,

$$\left\langle \int_0^{\tau \wedge \cdot} K_s \, dM_s, N \right\rangle_t = \left\langle \int_0^{\cdot} K_s \, dM_s, \ N^\tau \right\rangle_t$$
$$= \int_0^t K_s \, d\langle M, N^\tau \rangle_s = \int_0^t K_s \, d\langle M^\tau, N \rangle_s$$
$$= \left\langle \int_0^{\cdot} K_s \, dM_s^\tau, N_s \right\rangle_t .$$

The first required equality follows. We have used the properties of martingale brackets: $\langle M, N^\tau \rangle = \langle M^\tau, N \rangle$. For the second equality note that,

$$\left\langle \int_0^{\cdot} K_s \mathbf{1}_{s \le \tau} dM_s, N. \right\rangle_t = \int_0^t K_s \mathbf{1}_{s \le \tau} d\langle M, N \rangle_s.$$

By properties of Lebesgue-Stieltjes integrals,

$$\int_0^t K_s \mathbf{1}_{s \le \tau} \, d\langle M, N \rangle_s = \int_0^t K_s \, d\langle M, N \rangle_s^\tau$$
$$= \int_0^t K_s \, d\langle M^\tau, N \rangle_s = \left\langle \int_0^{\cdot} K_s dM_s^\tau, N \right\rangle_t .$$

The second required inequality follows. □

Corollary 6.12 *Let $S \le T$ be stopping times. Let τ be a stopping time. On $\tau \le S \wedge T$, $\int_0^\tau K_s dM_s^S = \int_0^\tau K_s dM_s^T$. In particular if $S \le T$, then on $\{t \le S\}$*

$$\int_0^t K_s dM_s^S = \int_0^t K_s dM_s^T.$$

Proof Just note that if S is a stopping time, then for all $t > 0$,

$$\int_0^t K_s dM_s^S = \int_0^{t \wedge S} K_s dM_s.$$

Then

$$\mathbf{1}_{\{\tau \le S \wedge T\}} \int_0^\tau K_s dM_s^S = \mathbf{1}_{\{\tau \le S \wedge T\}} \int_0^{\tau \wedge S} K_s dM_s = \mathbf{1}_{\{\tau \le S \wedge T\}} \int_0^\tau K_s dM_s.$$

Similarly

$$\mathbf{1}_{\{\tau \le S \wedge T\}} \int_0^\tau K_s dM_s^T = \mathbf{1}_{\{\tau \le S \wedge T\}} \int_0^\tau K_s dM_s.$$

Thus

$$\mathbf{1}_{\{\tau \le S\wedge T\}} \int_0^\tau K_s dM_s^S = \mathbf{1}_{\{\tau \le S\wedge T\}} \int_0^\tau K_s dM_s^T.$$

□

Definition 6.5 Let M be a continuous local martingale. Let $L^2_{loc}(M)$ be the space of progressively measurable stochastic processes, (K_t), for which there is a sequence of stopping times T_n increasing to infinity such that

$$\mathbf{E}\int_0^{T_n} K_s^2 d\langle M, M\rangle_s < \infty.$$

The class $L^2_{loc}(M)$ consists of all progressively measurable K such that

$$\int_0^t K_s^2 d\langle M, M\rangle_s < \infty, \quad \forall t.$$

Recall that

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

Proposition 6.13 *Let M be a continuous local martingale with $M_0 = 0$. If $H \in L^2_{loc}(M)$, there exists a unique local martingale $I(H, M)$ with $I(H, M)(0) = 0$ and*

$$\langle I(H, M), N\rangle_t = H\cdot\langle M, N\rangle_t, \qquad \forall t \ge 0$$

for all continuous local martingales (N_t).

Proof Step 1: Uniqueness. Let I_1, I_2 be continuous local martingales vanishing at 0 and satisfying

$$\langle I_1, N\rangle_t = H\cdot\langle M, N\rangle_t, \qquad \langle I_2, N\rangle_t = H\cdot\langle M, N\rangle_t \qquad \forall t \ge 0.$$

Let $\{T_n\}$ be a sequence of stopping times that reduces both I_1 and I_2 such that both stochastic processes are bounded. Then for all $N \in H^2_0$,

$$\langle I, H\rangle_t^{T_n} = \int_0^{t\wedge T_n} H_s d\langle M, N\rangle_s = \int_0^t H_s\, d\langle M^{T_n}, N\rangle_s$$

$$\langle I_1^{T_n}, N\rangle = \langle I_1, N\rangle^{T_n}, \qquad \langle I_2^{T_n}, N\rangle = \langle I_2, N\rangle^{T_n}.$$

It follows that $\langle I_1^{T_n}, N\rangle = \langle I_2^{T_n}, N\rangle$ for all $N \in H^2_0$. Since $I_1^{T_n}, I_2^{T_n}$ are L^2 bounded martingales, $I_1^{T_n} = I_2^{T_n}$. Since $T_n \to \infty$, $I_1 = I_2$ a.s.

Step 2. Existence. Let

$$T_n = \inf\left\{t \geq 0 : \int_0^t (1+H_s^2) d\langle M, M\rangle_s \geq n\right\}.$$

Then T_n is an increasing sequence of stopping times increasing to infinity such that M^{T_n} is in H^2 and $\|H\mathbf{1}_{s\leq T_n}\|_{L^2(M^{T_n})}$ is bounded by n. For $n < m$, on $t < T_n$,

$$\int_0^t H_s dM_s^{T_n} = \int_0^t H_s dM_s^{T_m}.$$

Define

$$I(H,M)(t) \equiv \int_0^t H_s dM_s(\omega) = \int_0^t \mathbf{1}_{s\leq T_n} H_s dM_s^{T_n}(\omega), \qquad \omega \in \{t < T_n\}.$$

Since

$$I(H,M)_t^{T_n} = \int_0^t \mathbf{1}_{s\leq T_n} H_s dM_s^{T_n}(\omega)$$

are martingales, $I(H,M)$ is a local martingale.

Let N be a continuous local martingale with reducing stopping times S_n such that N^{S_n} is bounded. Then

$$\begin{aligned}
\langle I(H,M), N\rangle_t^{T_n\wedge S_n} &= \left\langle I(H,M)^{T_n}, N^{S_n}\right\rangle_t = \left\langle \int_0^{\cdot} \mathbf{1}_{s\leq T_n} H_s dM_s^{T_n}, N^{S_n}\right\rangle_t \\
&= \int_0^t \mathbf{1}_{s\leq T_n} H_s d\langle M^{T_n}, N^{S_n}\rangle_s \\
&= \int_0^{t\wedge T_n \wedge S_n} H_s d\langle M, N\rangle_s.
\end{aligned}$$

Taking $n \to \infty$ to see that $\langle I(H,M), N\rangle_t = \int_0^t H_s d\langle M, N\rangle_s$. □

Definition 6.6 A progressively measurable process f is locally bounded if there is an increasing sequence of stopping times $\{T_n\}$, with $\lim_{n\to\infty} T_n = \infty$, such that there are constants C_n with $|f^{T_n}| \leq C_n$ for all n.

Both continuous functions and convex functions are locally bounded. Any locally bounded functions are in $L^2_{loc}(M)$ for any continuous local martingale M.

6.5 Stochastic Integration w.r.t. Semi-Martingales (Lecture 22)

Definition 6.7 If $X_t = M_t + A_t$ is a continuous semi-martingale and f is a progressively measurable locally bounded stochastic process, we define

$$\int_0^t f_s dX_s = \int_0^t f_s dM_s + \int_0^t f_s dA_s.$$

Proposition 6.14 *Let X, Y be continuous semi-martingales. Let f, g, K be locally bounded and progressively measurable. Let $a, b \in \mathbf{R}$.*

1. $\int_0^t (af_s + bg_s) dX_s = a \int_0^t f_s dX_s + b \int_0^t g_s dX_s$.
2. $\int_0^t f_s d(aX_s + bY_s) = a \int_0^t f_s dX_s + b \int_0^t f_s dY_s$.
3. $$\int_0^t f_s d\left(\int_0^s g_r dX_r\right) = \int_0^t f_s g_s dX_s.$$
4. *For any stopping time τ,*
$$\int_0^\tau K_s dX_s = \int_0^\infty \mathbf{1}_{s \le \tau} K_s dX_s = \int_0^\infty K_s dX_s^\tau.$$
5. *If X_s is of bounded total variation on $[0, t]$ so is the integral $\int_0^\cdot K_s dX_s$; and if X_s is a local martingale so is $\int K_s dX_s$. In particular for a semi-martingale X_t this gives the Doob-Meyer decomposition of $\int_0^\cdot K_s dX_s$, c.f. Definition 5.2.*

Definition 6.8 Let X, Y be continuous time semi-martingales. The Stratonovich integral is defined as:

$$\int_0^t Y_s \circ dX_s = \int_0^t Y_s dX_s + \frac{1}{2}\langle X, Y\rangle_t.$$

Also note that for continuous processes, Riemann sums corresponding to a sequence of partitions whose modulus goes to zero converges to the stochastic integral **in probability**. Note that this convergence does not help with computation. Although there are sub-sequences that converges a.s. we do not know which subsequence of the partition would work and this subsequence would be likely to differ for different integrands and different times.

Proposition 6.15 *If (K_t) is left continuous and $\Delta^n : 0 = t_0^n < t_1^n < \cdots < t_{N_n}^n = t$ is a sequence of partition of $[0,t]$ such that their modulus goes to zero, then*

$$\int_0^t K_s dX_s = \lim_{n\to\infty} \sum_{i=1}^{N_n} K_{t_i^n}(X_{t_{i+1}^n} - X_{t_i^n}).$$

The sum converges in probability.

The proof is a consequence of Proposition 6.16 below.

6.5.1 Appendix

Proposition 6.16 (Dominated Convergence Theorem) *Let (X_t) be a continuous semi-martingale. Let $\{(K_t^n)\}$ be a sequence of locally bounded progressively measurable stochastic processes, converging to (K_t). Suppose that there is a locally bounded progressively measurable stochastic process (F_t) such that $|K_t^n| \le F_t$ for every n and $t \ge 0$. Then*

$$\lim_{n\to\infty} \sup_{s\le t} \left| \int_0^s K_r^n dX_r - \int_0^s K_r dX_r \right| \to 0.$$

The convergence is in probability. That is, $(\int_0^s K_r^n dX_r, s \ge 0)$ converges in u.c.p.

If $(K_t^n) \to (K_t)$ in $L^2(M)$, this follows from Itô Isometry. Otherwise a localisation procedure will lead to the conclusion. See Theorem 2.12 in Revuz-Yor [24].

6.6 Itô's Formula (Lecture 22-23)

We define

$$\int_s^t H_s dX_s = \int_0^t H_s dX_s - \int_0^s H_s dX_s.$$

A $\mathbf{R}^n$-valued stochastic process $(X_t = (X_t^1, \dots, X_t^n)$ is respectively a martingale, local martingale, semi-martingale or a Brownian motion if each (X_t^i) is such a process. If $(X_t, t \ge 0)$ is a continuous semi-martingale, we denote by $\langle X, X\rangle_t$ the matrix whose entries are $\langle X^i, X^j\rangle_t$.

Theorem 6.17 (Itô's Formula) *Let $X_t = (X_t^1, \dots, X_t^n)$ be a $\mathbf{R}^n$-valued sample continuous semi-martingale and f a C^2 real valued function on $\mathbf{R}^n$ then for $s < t$,*

$$f(X_t) = f(X_s) + \sum_{i=1}^n \int_s^t \frac{\partial f}{\partial x_i}(X_r) dX_r^i + \frac{1}{2} \sum_{i,j=1}^n \int_s^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_r) d\langle X^i, X^j\rangle_r.$$

In short hand,

$$f(X_t) = f(X_s) + \int_s^t (Df)(X_r) dX_r + \frac{1}{2}\int_s^t (D^2 f)(X_r) d\langle X, X\rangle_r.$$

If T is a stopping time, apply Itô's formula to $Y_t = X_{T\wedge t}$ to see that

$$f(X_{T\wedge t}) = f(X_0) + \int_0^{T\wedge t} (Df)_{X_r} dX_r + \frac{1}{2}\int_0^{T\wedge t} (D^2 f)_{x_r} d\langle X, X\rangle_r.$$

Proof We give a sketch proof for $n = 1$. By the localization technique, it is sufficient to prove Itô's formula for a continuous semi-martingale which takes values in $[-C, C]$ for some $C > 0$.

If $f : \mathbf{R} \to \mathbf{R}$ is a C^2 function, let us recall Taylor's Theorem with an integral remainder term:

$$\begin{aligned} f(y) &= f(y_0) + f'(y_0)(y - y_0) + \int_{y_0}^{y} (y - z) f''(z) dz \\ &= f(y_0) + f'(y_0)(y - y_0) + \frac{1}{2} f''(y_0)(y - y_0)^2 + \int_{y_0}^{y} (y - z)(f''(z) - f''(y_0)) dz. \end{aligned}$$

Let R, which depends on y, y_0 and f, denote the remainder term. Let $C > 0$, since f'' is uniformly continuous over the compact interval $[-C, C]$, for $y, y_0 \in [-C, C]$, as $y \to y_0$,

$$\sup_{z\in[y_0,y]} |f''(z) - f''(y_0)| \to 0.$$

The rate of convergence depending only on the distance $|y - y_0|$, not on the their individual positions. Thus

$$\begin{aligned} |R(y_0, y)| &= \left|\int_{y_0}^{y} (y - z)\Big(f''(z) - f''(y_0)\Big) dz\right| \le |y - y_0| \int_{y_0}^{y} |f''(z) - f''(y_0)| dz \\ &\le (y - y_0)^2 \sup_{z\in[y_0,y]} |f''(z) - f''(y_0)|. \end{aligned}$$

Let us take a partition $\Delta^n : s = t_0^n < t_1^n < t_2^n < \cdots < t_{N(n)}^n = t$ and suppose that (X_t) is a continuous real valued semi-martingale bounded by K.

$$\begin{aligned} f(X_t) - f(X_s) &= \sum_{i=0}^{N(n)-1} \Big(f(X_{t_{i+1}^n}) - f(X_{t_i^n})\Big) \\ &= \sum_{i=0}^{N(n)-1} f'(X_{t_i^n}) \Big(X_{t_{i+1}^n} - X_{t_i^n}\Big) + \frac{1}{2}\sum_{i=0}^{N(n)-1} f''(X_{t_i^n}) \Big(X_{t_{i+1}^n} - X_{t_i^n}\Big)^2 \\ &\quad + \sum_{i=0}^{N(n)-1} \Big(R(X_{t_{i+1}^n}, X_{t_i^n})\Big). \end{aligned}$$

The first term, on the right hand side, converges in probability to $\int_0^t f'(X_s)dX_s$, see Proposition 6.15. The second term converges to $\frac{1}{2}\int_0^t f''(X_s)d\langle X,X\rangle_s$ in probability. (If X_t is continuous and has finite total variation, the second term converges to 0.)

The remainder term, denoted by $\tilde{R}$, is bounded by

$$\tilde{R} \le \sum_{i=0}^{N(n)-1} \sup_{t\in[t_i^n,.t_{i+1}^n]} \left|f(X_{t_{i+1}^n}) - f(X_{t_i^n})\right| \left(X_{t_{i+1}^n} - X_{t_i^n}\right)^2.$$

For any $\epsilon > 0$, if the partition size is sufficiently small, then

$$\tilde{R} \le \epsilon \sum_{i=0}^{N(n)-1} \left(X_{t_{i+1}^n} - X_{t_i^n}\right)^2.$$

Since $\sum_{i=0}^{N(n)-1} \left(X_{t_{i+1}^n} - X_{t_i^n}\right)^2$ converges in probability, as $\max_{0\le i\le N(n)-1}(t_{i+1}^n - t_i^n) \to 0$, to $\langle X,X\rangle_t$, this concludes the formula for real valued semi-martingales with values in a compact set. □

Proposition 6.18 (The product formula) *If X_t and Y_t are real valued semi-martingales,*

$$X_tY_t = X_0Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X,Y\rangle_t$$

This also provides an understanding, even serve as an definition, for the bracket process,

$$\langle X,Y\rangle_t = X_tY_t - X_0Y_0 - \int_0^t X_s dY_s - \int_0^t Y_s dX_s$$

Example 6.1 If $B_t = (B_t^2, \dots, B_t^n)$ is an n-dimensional BM, then $\langle B^i, B^j\rangle_t = \delta_{ij}t$, $|B_t|^2 = \sum_i |B_t^i|^2$, and

$$|B_t|^2 = 2\sum_{i=1}^n \int_0^t B_s^i dB_s^i + nt.$$

Example 6.2 Let (M_t) be a continuous semi-martingale. Then $X_t = e^{M_t - \frac{1}{2}\langle M,M\rangle_t}$ satisfies the equation:

$$X_t = e^{M_0} + \int_0^t X_s dM_s.$$

Let $Y_t := M_t - \frac{1}{2}\langle M, M\rangle_t$, then $\langle Y, Y\rangle_t = \langle M, M\rangle_t$, and $X_t = e_y^Y$. Let $f(x) = e^x$ and apply Itô's formula to the function f and the process (Y_t),

$$\begin{aligned}
X_t = e^{Y_t} =& e^{Y_0} + \int_0^t e^{Y_s} dY_s - \frac{1}{2}\int_0^t e^{Y_s} d\langle Y, Y\rangle_s \\
&= e^{M_0} + \int_0^t e^{Y_s} dM_s + \frac{1}{2}\int_0^t e^{Y_s} d\langle M, M\rangle_s - \frac{1}{2}\int_0^t e^{Y_s} d\langle Y, Y\rangle_s \\
=& e^{M_0} + \int_0^t X_s dM_s.
\end{aligned}$$

Definition 6.9 If (M_t) is a continuous local martingale, $e^{M_t - \frac{1}{2}\langle M,M\rangle_t}$ is a continuous local martingale and is called the exponential martingale of M_t.

Theorem 6.19 *1. Let (X_t) be a continuous semi-martingale. Assume that $\frac{\partial}{\partial t}F(t,x)$ and $\frac{\partial^2}{\partial x_i \partial x_j}F(t,x)$, $i, j = 1, \dots, d$, exist and are continuous functions. Then*

$$\begin{aligned}
F(t, X_t) =& F(0, X_0) + \int_0^t \frac{\partial F}{\partial s}(s, X_s) ds + \int_0^t DF(s, X_s) dX_s \\
&+ \frac{1}{2}\int_0^t D^2 F(s, X_s) d\langle X_s, X_s\rangle.
\end{aligned}$$

2. Itô's formula holds for complex valued semi-martingale $X_t + \sqrt{-1}Y_t$.

Appendix

Theorem 6.20 *Let (M_t) be a continuous local martingale. The exponential martingale $N_t := e^{M_t - \frac{1}{2}\langle M,M\rangle_t}$ is a martingale if and only if $\mathbf{E}(N_t) = 1$ for all t.*

Proof If (N_t) is a martingale, the statement that its expectation is constant in t follows from the definition. We prove the converse. Since (N_t) is a continuous local martingale, it is a super-martingale. Indeed for a reducing sequence of stopping times T_n and any pair of real numbers $0 \le s \le t$, we apply Fatou's lemma:

$$\mathbf{E}(N_t|\mathcal{F}_s) \le \lim_{n\to\infty} \mathbf{E}(N_t^{T_n}|\mathcal{F}_s) = \lim_{n\to\infty} N_s^{T_n} = N_s.$$

Let T be a stopping time bounded by a positive number K. By the optional stopping theorem,

$$\mathbf{E}(N_T) \ge \mathbf{E}(N_K) = 1, \qquad \mathbf{E}(N_T) \le \mathbf{E}(N_0) = 1.$$

Thus $\mathbf{E}(N_T) = 1$ and (N_t) is a martingale. $\square$

This can be generalized to stochastic processes that is not positive valued. Let (M_t) be a continuous local martingale with $\mathbf{E}|M_0| < \infty$. Suppose that the family $\{M_T^-, T$ bounded stopping times $\}$ is uniformly integrable. Then (M_t) is a super martingale. It is a martingale if and only if $\mathbf{E}M_t = \mathbf{E}M_0$, see Prop. 2.2 in [6]

Chapter 7

Stochastic Differential Equations

7.1 Stochastic processes defined up to a random time (Lecture 24)

Let (B_t) be a real valued Brownian motion. The stochastic process $X_t(\omega) := \frac{1}{2-B_t(\omega)}$ is defined up to the first time $B_t(\omega)$ reaches 2. We denote this time by τ:

$$\tau(\omega) = \inf_{t\geq 0}\{B_t(\omega) \geq 2\}.$$

For any given time t, no matter how small it is, there is a set of path of positive probability (measured with respect to the Wiener measure on $C([0,t];\mathbf{R}^d)$) which will have reached 2 by time t:

$$P(\tau \leq t) = P(\sup_{s\leq t} B_s \geq 2) = 2P(B_t \geq 2) = \sqrt{\frac{2}{\pi}}\int_{\frac{2}{\sqrt{t}}}^{\infty} e^{-\frac{y^2}{2}}dy > 0.$$

This probability converges to zero as $t \to 0$. We say that (X_t) is defined up to τ and τ is called its life time or explosion time.

Exercise 7.1 (1) Using Itô's formula to prove that on $\{t < \tau(\omega)\}$, the following holds almost surely

$$X_t = \frac{1}{2} + \int_0^t (X_s)^2 dB_s + \int_0^t (X_s)^3 ds.$$

Let $c \neq 0$ and define

$$\tau_c(\omega) = \inf_{t\geq 0}\{B_t(\omega) \geq c\}.$$

Prove that for $c \leq 2$, on $\{t < \tau_c(\omega)\}$, the above identity holds a.s. The process $(X_t, t < \tau_2)$ is the maximal solution to the above integral equation.

(2) Replace (B_t) be another Brownian motion (W_t). Find a stochastic process (X_t) that solves

$$X_t = \frac{1}{2} + \int_0^t (X_s)^2 dW_s + \int_0^t (X_s)^3 ds.$$

(3) Let $a \in \mathbf{R}$, find a stochastic process (X_t) that solves

$$X_t = a + \int_0^t (X_s)^2 dB_s + \int_0^t (X_s)^3 ds.$$

Prove your claim.

Let $\mathbf{R}^d \cup \{\Delta\}$ be the one point compactification of $\mathbf{R}^d$, which is a topological space whose open sets are open sets of $\mathbf{R}^d$ plus set of the form $(\mathbf{R}^d \setminus K) \cup \{\Delta\}$ where K denotes a compact set. Given a process $(X_t, t < \tau)$ on $\mathbf{R}^d$ we define a process $(\hat{X}_t, t \geq 0)$ on $\mathbf{R}^d \cup \{\Delta\}$:

$$\hat{X}_t(\omega) = \left\{ \begin{array}{ll} X_t(\omega), & \text{if } t < \tau(\omega) \\ \Delta, & \text{if } t \geq \tau(\omega). \end{array} \right\}.$$

If $(X_t, t < \tau)$ is a continuous process on $\mathbf{R}^d$ then $(\hat{X}_t)$ is a continuous process on $\mathbf{R}^d \cup \Delta$. Define $\hat{W}(\mathbf{R}^d) \equiv C([0,T]; \mathbf{R}^d \cup \Delta)$ whose elements satisfy that: if $Y_s = \Delta$ then $Y_t(\omega) = \Delta$ for all $t \geq s$. The last condition means that once a process enters the coffin state it does not return.

7.2 Concepts

For $i = 1, 2, \dots, m$, let $\sigma_i, b : \mathbf{R} \times \mathbf{R}^d \to \mathbf{R}^d$ be Lipschitz continuous functions, more generally Borel measurable functions. Let $B_t = (B_t^1, \dots, B_t^m)$ be an $\mathbf{R}^m$ valued Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. We study first the integral equation of Markovian type:

$$x_t = x_0 + \sum_{k=1}^m \int_0^t \sigma_k(s, x_s) dB_s^k + \int_0^t b(s, x_s) ds. \tag{7.1}$$

Once we are familiar with the basic estimates, we will proceed to the concept of stochastic differential equations.

A stochastic process $(x_t(\omega))$ is a solution to the above integral equation if for all t the identity (7.1) holds almost surely. If the functions $\sigma_i(t, x)$ and $b(t, x)$ do not depend on t, the SDE is said to be time homogeneous.

We concern ourselves, mostly, with time homogeneous SDEs (of Markovian type), letting $\sigma_i, b : \mathbf{R}^d \to \mathbf{R}$ be Borel measurable locally bounded functions:

$$E(\sigma, b) \qquad\qquad dx_t = \sum_{i=1}^{m} \sigma_i(x_t) dB_t^i + b(x_t)dt. \tag{7.2}$$

Example 7.1 The following SDE on $\mathbf{R}^d$ is not Markovian:

$$dx_t = \left(\int_0^t x_r dr \right) dB_t,$$

equivalently, $x_t = x_0 + \int_0^t \left(\int_0^s x_r dr \right) dB_s$. Let us define $\sigma : W^d \to \mathbf{R}$, where W^d is the Wiener space, by $\sigma(w) := \left(\int_0^t w_r dr \right)$. Note that $\sigma(x.)$ depends on the value of the whole path of the solution.

7.3 Stochastic Integral Equations (Lectures 22-26)

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered probability space satisfying the standard assumptions Let (B_t) be a real value Brownian motion. Letting $f_s = (f_s^1, \ldots, f_s^d)$ be a progressively measurable stochastic process with values in $\mathbf{R}^d$, we define

$$\int_0^t f_s dB_s = \left(\int_0^t f_s^1 dB_s, \ldots, \int_0^t f_s^d dB_s \right).$$

Denote $|x| = \sqrt{\sum_{i=1}^d (x_i)^2}$ the Euclidean norm of $x = (x_1, \ldots, x_d)$ in $\mathbf{R}^d$.

Lemma 7.1 *Let $T > 0$ and $t \le T$. Let $(f_s, s \le T)$ be progressively measurable $\mathbf{R}^d$-valued stochastic processes.*

(1) There exists a universal constant C such that

$$\mathbf{E} \sup_{t \le T} \left| \int_0^t f_s dB_s \right|^2 \le C \int_0^T \mathbf{E}|f_s|^2 ds.$$

(2)

$$\mathbf{E} \left(\sup_{t \le T} \left| \int_0^t f_s ds \right|^2 \right) \le T \int_0^T \mathbf{E}|f_s|^2 ds.$$

Proof For (1). By Burkholder-Davis-Gundy inequality, Theorem 5.7, there exists a universal constant C s.t.

$$\mathbf{E}\sup_{t\le T}\left|\int_0^t f_s dB_s\right|^2 \le \sum_{i=1}^d \mathbf{E}\left(\sup_{t\le T}\int_0^t f_s^i dB_s\right)^2$$
$$\le C\mathbf{E}\sum_{i=1}^d \int_0^T (f_s^i)^2 ds = C\int_0^T \mathbf{E}|f_s|^2 ds.$$

(2) For any $0 \le t \le T$,

$$\mathbf{E}\sup_{t\le T}\left|\int_0^t f_s ds\right|^2 \le \sum_{i=1}^d \mathbf{E}\sup_{t\le T}\left(\int_0^t f_s^i ds\right)^2$$
$$\le T\mathbf{E}\int_0^T |f_s|^2 ds \le T\int_0^T \mathbf{E}|f_s|^2 ds.$$

□

In particular, if $\{(B_t^k), k = 1, \dots, m\}$ are independent Brownian motions, $f_k(s)$ are progressively measurable $\mathbf{R}^d$-valued stochastic processes, then from the triangle inequality for norms, and (1) in Lemma 7.1,

$$\mathbf{E}\left|\sup_{t\le T}\sum_{k=1}^m \int_0^t f_k(s)dB_s^k\right|^2 \le \mathbf{E}\left(\sup_{t\le T}\sum_{k=1}^m \left|\int_0^t f_k(s)dB_s^k\right|\right)^2$$
$$\le 2^{m-1}\sum_{k=1}^m \mathbf{E}\left|\sup_{t\le T}\int_0^t f_k(s)dB_s^k\right|^2$$
$$\le 2^{m-1}C\sum_{k=1}^m \int_0^T \mathbf{E}|f_k(s)|^2 ds.$$

Definition 7.1 A function $f : \mathbf{R}^d \to \mathbf{R}$ is said to growth at most linearly, if there exists a constant C such that

$$|f(x)| \le C(1 + |x|)$$

for all $x \in \mathbf{R}^d$.

Lemma 7.2 *If $f : \mathbf{R}^d \to \mathbf{R}$ is Lipschitz continuous, then f grows at most linearly at infinity.*

Proof For any $x \in \mathbf{R}^d$,

$$\begin{aligned}|f(x)| &\le |f(x) - f(0)| + |f(0)| \le |f|_{\text{Lip}}|x| + |f(0)| \\ &\le C(1 + |x|),\end{aligned}$$

where $C = \max(|f(0)|, |f|_{\text{Lip}})$. □

Theorem 7.3 *For $i = 1, 2, \ldots, m$, let $\sigma_i, b : \mathbf{R}^d \to \mathbf{R}^d$ be Lipschitz continuous functions. Let $B_t = (B_t^1, \ldots, B_t^m)$ be an $\mathbf{R}^m$ valued Brownian motion on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. Then for each $x_0 \in \mathbf{R}^d$ there exists a unique continuous $\mathcal{F}_t^B$-adapted stochastic process such that*

$$x_t = x_0 + \sum_{k=1}^m \int_0^t \sigma_k(x_s) dB_s^k + \int_0^t b(x_s) ds$$

for all t a.s. Furthermore for each t, x_t is $\mathcal{F}_t^B$-measurable.

Proof Fix $T > 1$. Define, for all $t \in [0, T]$,

$$\begin{aligned}x_t^{(0)} &= x_0, \\ x_t^{(1)} &= x_0 + \sum_{k=1}^m \int_0^t \sigma_k(x_0) dB_r^k + \int_0^t b(x_0) dr \\ &\ldots \\ x_t^{(n)} &= x_0 + \sum_{k=1}^m \int_0^t \sigma_k(x_r^{(n-1)}) dB_r^k + \int_0^t b(x_r^{(n-1)}) dr \\ x_t^{(n+1)} &= x_0 + \sum_{k=1}^m \int_0^t \sigma_k(x_r^{(n)}) dB_r^k + \int_0^t b(x_r^{(n)}) dr.\end{aligned}$$

$$\begin{aligned}\mathbf{E} \sup_{t \le u} |x_t^{(1)} - x_0|^2 &\le 2\mathbf{E} \sup_{t \le u} \left| \sum_{k=1}^m \int_0^t \sigma_k(x_0) B_t^k \right|^2 + 2u\mathbf{E}|b(x_0)|^2 \\ &\le 2^m \sum_{k=1}^m \mathbf{E} \left| \sigma_k(x_0) B_u^k \right|^2 + 2u|b(x_0)|^2 \\ &\le 2^m \sum_{k=1}^m \tilde{C}^2 (1 + |x_0|)^2 \mathbf{E}(B_u^k)^2 + 2u\tilde{C}^2(1 + |x_0|^2) \\ &= (2^m mu + 2u)(\tilde{C})^2 (1 + |x_0|)^2 = C_0,\end{aligned}$$

where $\tilde{C}$ is the common linear growth constants for σ_k and b. By induction and analogous estimation, $\mathbf{E}\sup_{t\le u}|x_t^{(n)}|^2$ is finite and the stochastic integrals make sense. By construction each $(x_t^{(n)})$ is sample continuous and is adapted to the filtration of (B_t).

We estimate the differences between iterations:

$$\begin{aligned}
&\mathbf{E}\sup_{s\le t}|x_s^{(n+1)}-x_s^{(n)}|^2\\
&=\mathbf{E}\sup_{s\le t}\left|\sum_{k=1}^m\int_0^s\Big(\sigma_k(x_r^{(n)})-\sigma(x_r^{(n-1)}\Big)dB_r^k+\int_0^s\Big(b(x_r^{(n)})-b(x_r^{(n-1)}\Big)dr\right|^2.\\
&\le 2\mathbf{E}\sup_{s\le t}\left(\sum_{k=1}^m\left|\int_0^s\Big(\sigma_k(x_r^{(n)})-\sigma_k(x_r^{(n-1)}\Big)dB_r^k\right|\right)^2+2\mathbf{E}\sup_{s\le t}\left|\int_0^s\Big(b(x_r^{(n)})-b(x_r^{(n-1)})\Big)dr\right|^2\\
&\le 2^m\sum_{k=1}^m\mathbf{E}\sup_{s\le t}\left|\int_0^s\Big(\sigma_k(x_r^{(n)})-\sigma_k(x_r^{(n-1)}\Big)dB_r^k\right|^2+2\mathbf{E}\sup_{s\le t}\left|\int_0^s\Big(b(x_r^{(n)})-b(x_r^{(n-1)})\Big)dr\right|^2.
\end{aligned}$$

Let K be the common Lipschitz constant for σ_k and b. By Lemma 7.1,

$$\begin{aligned}
&\mathbf{E}\sup_{s\le t}|x_s^{(n+1)}-x_s^{(n)}|^2\\
&\le 2^mC\sum_{k=1}^m\mathbf{E}\left(\int_0^t\left|\sigma_k(x_r^{(n)})-\sigma_k(x_r^{(n-1)})\right|^2dr\right)+2T\mathbf{E}\int_0^t\left|b(x_r^{(n)})-b(x_r^{(n-1)})\right|^2dr.\\
&\le 2^mC\sum_{k=1}^mK^2\int_0^t\mathbf{E}\left|x_r^{(n)}-x_r^{(n-1)}\right|^2dr+2TK^2\int_0^t\mathbf{E}\left|x_r^{(n)}-x_r^{(n-1)}\right|^2dr.
\end{aligned}$$

Let

$$D=2^mCmK^2+2TK^2,$$

Then

$$\begin{aligned}
\mathbf{E}\sup_{s\le t}|x_s^{(n+1)}-x_s^{(n)}|^2&\le D\int_0^t\mathbf{E}\left|x_r^{(n)}-x_r^{(n-1)}\right|^2dr\\
&\le D\int_0^t\mathbf{E}\sup_{r\le s_1}\left|x_r^{(n)}-x_r^{(n-1)}\right|^2ds_1\\
&\le D^2\int_0^t\int_0^{s_1}\mathbf{E}\sup_{r\le s_2}\left|x_r^{(n-1)}-x_r^{(n-2)}\right|^2ds_2ds_1\\
&\le D^n\int_0^t\int_0^{s_1}\dots\int_0^{s_{n-1}}\mathbf{E}\sup_{r\le s_n}\left|x_r^{(1)}-x_r^{(0)}\right|^2ds_n\dots ds_2ds_1.
\end{aligned}$$

By induction we see that

$$\mathbf{E}\sup_{s\le t}|x_s^{(n+1)}-x_s^{(n)}|^2\le C_1\frac{D^nT^n}{n!}.$$

whee $C_1=T\mathbf{E}\sup_{t\le T}|x_t^{(1)}-x_0|^2\le TC_0$. By Minkowski inequality,

$$\sqrt{\mathbf{E}\left(\sum_{k=1}^\infty\sup_{s\le t}|x_s^{(k+1)}-x_s^{(k)}|\right)^2}\le\sum_{k=1}^\infty\left(\mathbf{E}\sup_{s\le t}|x_s^{(k+1)}-x_s^{(k)}|^2\right)^{\frac{1}{2}}<\infty.$$

By Fatou's lemma,

$$\begin{aligned}&\sum_{k=1}^\infty\left(\mathbf{E}\sup_{s\le t}|x_s^{(k+1)}-x_s^{(k)}|^2\right)^{\frac{1}{2}}\\ &\le\sum_{k=1}^\infty\left(\mathbf{E}\sup_{s\le t}|x_s^{(k+1)}-x_s^{(k)}|^2\right)^{\frac{1}{2}}\le C_1\sum_{k=1}^\infty\sqrt{\frac{D^kT^k}{k!}}<\infty.\end{aligned}$$

In particular for almost surely all ω,

$$\sum_{k=1}^\infty\sup_{s\le t}|x_s^{(k+1)}(\omega)-x_s^{(k)}(\omega)|<\infty.$$

For such ω, $\{x_s^{(n)}(\omega)\}$ is a Cauchy sequence in $C([0,t];\mathbf{R}^d)$. Let $x_t(\omega)=\lim_{n\to\infty}x_t^{(n)}(\omega)$. The process is continuous in time by the uniform convergence.

We take $n\to\infty$ in

$$x_t^{(n)}=x_0+\sum_{k=1}^m\int_0^t\sigma_k(x_r^{(n-1)})dB_r^k+\int_0^t b(x_r^{(n-1)})dr.$$

As $n\to\infty$, $\int_0^t\sigma_k(x_s^{(n)})dB_s^k\to\int_0^t\sigma(x_s)dB_s$ in probability. There will be an almost sure convergent subsequence and this proves that

$$x_t=x_0+\sum_{k=1}^m\int_0^t\sigma_k(x_s)dB_s^k+\int_0^t b(x_s)ds.$$

Since each $(x_t^{(n)})$ is a adapted to the filtration of (B_t), so is its limit.

(2) Uniqueness. Let (x_t) and (y_t) be two solutions with $x_0 = y_0$ a.s. Let C^* be a constant.

$$
\begin{aligned}
&\mathbf{E} \sup_{s\le t} |x_s - y_s|^2 \\
&= \mathbf{E} \sup_{s\le t} \left| \sum_{k=1}^m \int_0^s \Big(\sigma_k(x_r) - \sigma(y_r)\Big) dB_r^k + \int_0^s \Big(b(x_r) - b(y_r)\Big) dr \right|^2 \\
&\le \mathbf{E} \sup_{s\le t} \left(\sum_{k=1}^m \left| \int_0^s \Big(\sigma_k(x_r) - \sigma_k(y_r)\Big) dB_r^k \right| + \left| \int_0^s \Big(b(x_r) - b(y_r)\Big) dr \right| \right)^2 \\
&\le 2^m \sum_{k=1}^m \mathbf{E} \left(\sup_{s\le t} \left| \int_0^s \Big(\sigma_k(x_r) - \sigma_k(y_r)\Big) dB_r^k \right| \right)^2 + 2\mathbf{E} \left(\sum_{k=1}^m \sup_{s\le t} \left| \int_0^s \Big(b(x_r) - b(y_r)\Big) dr \right| \right)^2 \\
&\le 2^m C^* \sum_{k=1}^m \mathbf{E} \left(\int_0^t |\sigma_k(x_r) - \sigma_k(y_r)|^2 \, dr \right) + 2T\mathbf{E} \int_0^t |b(x_r) - b(y_r)|^2 \, dr \\
&\le 2^m C^* \sum_{k=1}^m K^2 \int_0^t \mathbf{E} \, |x_r - y_r|^2 \, dr + 2TK^2 \int_0^t \mathbf{E} \, |x_r - y_r|^2 \, dr \\
&\le (2^m m C^* K^2 T + 2TK^2) \int_0^t \mathbf{E} \left(\sup_{r\le s} |x_r - y_r|^2 \right) ds,
\end{aligned}
$$

By Grownall's inequality,

$$\mathbf{E} \sup_{s\le t} |x_s - y_s|^2 = 0.$$

In particular, $\sup_{s\le t} |x_s - y_s|^2 = 0$ almost surely. □

Lemma 7.4 (Grownall's Inequality/Gronwall's Lemma) *Let $T > 0$. Suppose that $f : [0,T] \to \mathbf{R}_+$ is a locally bounded Borel function such that there are two real numbers C and K such that for all $0 \le t$,*

$$f(t) \le C + K \int_0^t f(s) ds.$$

Then

$$f(t) \le Ce^{Kt}, \quad t \le T$$

In particular if $C = 0$, $f(t) = 0$ for all $t \le T$.

Definition 7.2 A solution $(x_t, t < \tau)$ of an SDE is a maximal solution if $(y_t, t < \bar{\tau})$ is any other solution on the same probability space with the same driving noise and with $x_0 = y_0$ a.s., then $\tau \geq \bar{\tau}$ a.s.. We say that τ is the explosion time or the life time of (x_t).

By localisation, or cut off the functions σ_k we have the following theorem:

Theorem 7.5 *Suppose that for $k = 1, \dots, m$, $\sigma_k : \mathbf{R}^d \to \mathbf{R}^d$ and $b : \mathbf{R}^d \to \mathbf{R}^d$ are locally Lipschitz continuous , i.e. for each $N \in \mathcal{N}$, there exists a number K_N such that for all x, y with $|x| \leq N, |y| \leq N$,*

$$|\sigma_k(x) - \sigma_k(y)| \leq K_N|x - y|, \qquad |b(x) - b(y)| \leq K_N|x - y|.$$

Then there is a maximal solution $(x_t, t < \tau)$. If $(x_t, t < \tau)$ and $(y_t, t < \zeta)$ be two maximal solutions with the same initial value $x \in \mathbf{R}^d$, then $\tau = \zeta$ a.s. and (x_t) and (y_t) are indistinguishable.

7.4 Examples

Example 7.2 Consider $\dot{x}(t) = ax(t)$ on $\mathbf{R}$ where $a \in \mathbf{R}$. Let $x_0 \in \mathbf{R}$. Then $x(t) = x_0 e^{at}$ is a solution with initial value x_0. It is defined for all $t \geq 0$.

Let $\phi_t(x_0) = x_0 e^{at}$. Then $(t, x) \mapsto \phi_t(x)$ is continuous and $\phi_{t+s}(x_0) = \phi_t(\phi_s(x))$.

Example 7.3 Linear Equation. let $a, b \in \mathbf{R}$. Let $d = m = 1$. Then

$$x(t) = x_0 e^{aB_t - \frac{a^2}{2}t + bt}$$

solves

$$dx_t = a\, x_t\, dB_t + b\, x_t\, dt, \quad x(0) = x_0.$$

The solution exists for all time.

Is this solution unique? The answer is yes. Let y_t be a solution starting from the same point, we could compute and prove that $\mathbf{E}|x_t - y_t|^2 = 0$ for all t, which implies that $x_t = y_t$ a.s. for all t. However we do not prove it here, see Theorem 7.19.

Example 7.4 Additive noise. Consider a particle of mass 1, subject to a force which is proportional to its own speed, is subject to $\dot{v}_t = -kv_t$. Its random perturbation equation is the Langevin equation:

$$dv_t(\omega) = -kv_t(\omega)dt + dB_t(\omega).$$

For each realisation of the noise (that means for each ω), the solution is an Ornstein-Uhlenbeck process,

$$v_t(\omega) = v_0 e^{-kt} + \int_0^t e^{-k(t-r)} dB_r(\omega).$$

We check that the above equation satisfies $-k\int_0^t v_s ds = v_t - v_0 - B_t$.

$$\begin{aligned}
-k\int_0^t v_s ds &= -kv_0 \int_0^t e^{-ks} ds - \int_0^t \left(ke^{-ks} \int_0^s e^{kr} dB_r \right) ds \\
&= -v_0 + v_0 e^{-kt} + \int_0^t \left(\int_0^s e^{kr} dB_r \right) d(e^{-ks}) \\
&= -v_0 + v_0 e^{-kt} + \left(e^{-ks} \int_0^s e^{kr} dB_r \right) \Big|_{s=0}^{s=t} - \int_0^t e^{-ks} d \left(\int_0^s e^{kr} dB_r \right) \\
&= -v_0 + v_0 e^{-kt} + e^{-kt} \int_0^t e^{kr} dB_r - B_t \\
&= -v_0 + v_t - B_t.
\end{aligned}$$

This proved that the Ornstein-Uhlenbeck process is solution to the Langevin equation, with life time ∞.

Example 7.5 (1) Small Perturbation. Let $\epsilon > 0$ be a small number,

$$x_t^\epsilon = x_0 + \int_0^t b(x_s^\epsilon) ds + \epsilon B_t.$$

As $\epsilon \to 0$, $x_t^\epsilon \to x_t$. (Exercise)

(2) Let $y_t^\epsilon = y_0 + \epsilon \int_0^t b(y_s^\epsilon) ds + \sqrt{\epsilon} W_t$. Assume that b are bounded, as $\epsilon \to 0$, y_t^ϵ on any finite time interval converges uniformly in time on any finite time interval $[0,t]$, $\mathbf{E} \sup_{0\le s\le t}(y_s^\epsilon - y_0) \to 0$.

7.5 Notions of Solutions (Lectures 26-27)

For $i = 1, 2, \dots, m$, let $\sigma_i, b : \mathbf{R}_+ \times \mathbf{R}^d \to \mathbf{R}^d$ be Borel measurable locally bounded functions. Let $B_t = (B_t^1, \dots, B_t^m)$ be an $\mathbf{R}^m$ valued Brownian motion.

Definition 7.3 A d-dimensional stochastic process $(x_t, t < \tau)$, where $\tau \le \infty$, on a probability space $(\Omega, \mathcal{G}, P)$ is a solution to the SDE (of Markovian type)

$$dx_t = \sigma(t, x_t) dB_t + b(t, x_t) dt. \tag{7.3}$$

If there exists a filtration $(\mathcal{F}_t)$ such that

(1) x_t is adapted to $\mathcal{F}_t$,

(2) a $\mathcal{F}_t$ Brownian motion $B_t = (B_t^1, \dots, B_t^m)$ with $B_0 = 0$;

(3) for all stopping times $T < \tau$, the following makes sense and holds almost surely

$$x_T = x_0 + \sum_{k=1}^{m} \int_0^T \sigma_k(s, x_s) dB_s^k + \int_0^T b(s, x_s) ds.$$

We may replace (3) by (3')

(3') an adapted continuous stochastic process $x.$ in $C([0,\infty); \mathbf{R}^d \cup \{\Delta\})$, s.t. for all $t \geq 0$,

$$x_t = x_0 + \sum_{k=1}^{m} \int_0^t \sigma_k(s, x_s) dB_s^k + \int_0^t b(s, x_s) ds, a.s.$$

In essence the SDE holds on $\{t < \tau(\omega)\}$. The maximal time τ, up to which a solution is defined is the explosion time, the solution $(x_t, t < \tau)$ is the maximal solution.

Definition 7.4 A solution is a global solution is its life time infinite. We say that the SDE does not explode from x_0 if its solution from x_0 is global. We say that the SDE does not explode if all of its solutions are global.

Example 7.6 (Tanaka's SDE) Consider the equation on $\mathbf{R}$,

$$dx_t = \text{sign}(x_t) dB_t$$

where

$$\text{sign}(x) = \begin{cases} -1, & \text{if } x \leq 0 \\ \ \ 1, & \text{if } x > 0. \end{cases}$$

Let (W_t) be a Brownian motion on any probability space with $B_0 = 0$. Let $x_0 \in \mathbf{R}^d$. Define

$$B_t = \int_0^t \text{sign}(x + W_s)\ dW_s.$$

This is a local martingale with quadratic variation t and hence a Brownian motion. Furthermore

$$\int_0^t \text{sign}(x + W_s) dB_s = \int_0^t dW_s = W_t.$$

Thus

$$x + W_t = x + \int_0^t \text{sign}(x + W_s) dB_s,$$

and $x + W_t$ solves Tanaka's equation.

Example 7.7 Consdier $dx_t = x_t dB_t$ on $\mathbf{R}$. Let $d \geq 2$ and $\Omega = W^d = C([0, \infty); \mathbf{R}^d)$ with the Borel σ-algebra generated

$$d(\omega_1, \omega_2) = \sum_{n=1}^{\infty} \frac{\sup_{t \leq n} |\omega_1(t) - \omega_2(t)| \wedge 1}{2^n}$$

and filtration generated by evaluation maps $(\pi_s : s \leq t)$. Let $B_t = \pi_t^1$, projection to the first component, and evaluated at time t. Then $x_t^1 = xe^{\pi_t^1 - \frac{t}{2}}$ is a solution with initial value x. Let $B_t^2 = \pi_t^2$ then $x_t^2 = xe^{\pi_t^2 - \frac{t}{2}}$ is a solution with initial value x. The stochastic processes x_t^1 and x_t^2 are independent Brownian motions! However there is a related universal function $F_t(x, \sigma) = xe^{\pi_t(\sigma) - \frac{t}{2}}$.

Let (x_t) be a solution with initial value x_0. Suppose that the solution $(x_0, \omega) \mapsto (x_s(\omega), s \leq t)$ is adapted to $\mathcal{B}(\mathbf{R}^d) \times \mathcal{F}_t^B$. Then there exists a Borel measurable function F_t on $\mathbf{R}^d \times C([0, t]; \mathbf{R}^d)$ such that

$$x_t(\omega) = F_t(x_0, (B_s, 0 \leq s \leq t)).$$

This raises the following questions.

Questions.

1. Given any $y_0 \in \mathbf{R}^d$, is $F_t(y_0, B.)$ a solution with initial value y_0?
2. Given another probability space $(\Omega', \mathcal{F}', \mathcal{F}_t', P')$ and a Brownian motion (W_t), is $y_t = F_t(y_0, (W_s, 0 \leq s \leq t))$ a solution to

$$y_t = y_0 + \sum_{k=1}^{m} \int_0^t \sigma_k(y_s) dW_s^k + \int_0^t \sigma(y_s) ds?$$

Definition 7.5 A solution (x_t, B_t) on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is said to be a strong solution, if x_t is adapted to the filtration of B_t for each t. By a weak solution we mean one which is not strong.

If (x_t, B_t) is a solution on the canonical probability space so $B_t(\omega) = \omega(t)$. If x_t is measurable w.r.t. to the natural filtration of (B_t) then there exists a Borel measurable function $F_t(x_0, \cdot) : W_0^m \to \hat{W}^d$ such that

$$x_t(\omega) = F_t(x_0, (B_s(\omega), s \leq t)) = F_t(x, (\omega.s \leq t)).$$

7.6 Notions of uniqueness (Lecture 27)

Example 7.8 (Tanaka's SDE) Consider the equation on $\mathbf{R}$,

$$dx_t = \text{sign}(x_t)dB_t$$

where

$$\text{sign}(x) = \begin{cases} -1, & \text{if } x \le 0 \\ 1, & \text{if } x > 0. \end{cases}$$

If (x_t) is a solution, then $x_t - x_0 = \int_0^t x_s dB_s$ is a Brownian motion. The stochastic integral $\int_0^t x_s dB_s$ is a martingale with quadratic variation t. By Lévy Characterisation Theorem, the distribution of $(x_s - x_0, s \le t)$ is the Wiener measure on $C([0,t];\mathbf{R}^d)$. In another word, the probability distribution of (x_t) is that of a Brownian motion with initial value x_0.

Definition 7.6 If, whenever (x_t) and $(\tilde{x}_t)$ are two solutions with $x_0 = \tilde{x}_0$ almost surely, the probability distribution of $\{x_t : t \ge 0\}$ is the same as the probability distribution of $\{\tilde{x}_t, t \ge 0\}$, we say that uniqueness in law holds.

Uniqueness in law holds for Tanaka's equation. Uniqueness in law implies the following stronger conclusion: whenever x_0 and $\tilde{x}_0$ have the same distribution, the corresponding solutions have the same law.

Example 7.9 (Tanaka's SDE) If (x_t) solves Tanaka's equation $x_t = \int_0^t x_s dB_s$ (initial value 0), then so does $(-x_s)$.

Definition 7.7 We say pathwise uniqueness of solution holds for an SDE, If whenever (x_t) and $(\tilde{x}_t)$ are two solutions for the SDE on the same probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ with the same driving Brownian motion (B_t) and same initial data ($x_0 = \tilde{x}_0$ a.s.), then $x_t = \tilde{x}_t$ for all $t \ge 0$ almost surely.

Parthwise uniqueness fails for Tanaka's equation.

Example 7.10 ODE $\dot{x}_t = (x_t)^\alpha dt$, $\alpha < 1$, which has two solutions from zero: the trivial solution 0 and $x_t = (1-\alpha)^{\frac{1}{1-\alpha}}\ t^{\frac{1}{1-\alpha}}$. Both uniqueness fails.

Example 7.11 Dimension $d = 1$. Consider $dx_t = \sigma(x_t)dW_t$. Suppose that σ is Hölder continuous of order α, $|\sigma(x) - \sigma(y)| \le c|x-y|^\alpha$ for all x, y. If $\alpha \ge 1/2$ then pathwise uniqueness holds for $dx_t = \sigma(x_t)dW_t$. If $\alpha < 1/2$ uniqueness no longer holds. For $\alpha > 1/2$ this goes back to Skorohod (62-65) and Tanaka(64). The $\alpha = 1/2$ case is credited to Yamada-Watanabe.

7.6.1 The Yamada-Watanabe Theorem

The following beautiful, and somewhat surprising, theorem of Yamada and Watanabe, states that the existence of a weak solution for any initial distribution together with pathwise uniqueness implies the existence of a unique strong solution.

Proposition 7.6 *If pathwise uniqueness holds then any solution is a strong solution and uniqueness in law holds.*

For the precise meaning of 'universally measurable' see P163 of Ikeda-Watanabe's book [13].

Theorem 7.7 (The Yamada-Watanabe Theorem) *If for each initial probability distribution there is a weak solution to the SDE and suppose that pathwise uniqueness holds then there exists a unique strong solution. By this we meant that there is a progressively measurable map: $F : \mathbf{R}^d \times W_0^m \to \hat{W}^d$, where the σ-algebras are 'universally complete', such that*

1. *for any probability measure μ on $\mathbf{R}^d$ there exists $\tilde{F}$ that is measurable w.r.t. $\mathcal{B}(\mathbf{R}^d \times W_0^m)^{\mu \times P}$ s.t. $F(x, \omega) = \tilde{F}(x, \omega)$ a.s.. If $\xi_0 \in \mathcal{F}_0$ we set $F(\xi_0, B) = \tilde{F}(\xi_0, B)$.*
2. *For any BM (B_t) on a probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, and any $\xi_0 \in \mathcal{F}_0$, $x_t = \tilde{F}_t(\xi_0, B_t)$ is a solution to the SDE with driving noise (B_t) and initial value ξ_0.*
3. *If x_t is a solution to the equation with driving noise (B_t), then $x_t = F_t(x_0, B)$ a.s.*

In another word, for any B_t, and $x_0 \in \mathbf{R}^d$, $F_t(x_0, B)$ is a solution with the driving noise B_t. If x_t is a solution on a filtered probability space with driving noise B_t, then $x_t = F_t(x, B)$ a.s.

We do not prove this theorem, but refer to Ikeda-Watanabe and Revuz-Yor. The following observation is important for the proof of the Yamada-Watanabe Theorem. Given two solutions on two probability space we could build them on the same probability space: $\hat{W}^d \times \hat{W}^d \times W^m$.

Lemma 7.8 *Let f, g be locally bounded predictable processes (measurable with respect to the filtration generated by left continuous processes), and B, W continuous semi-martingales. If $(f, B) = (g, W)$ in distribution then*

$$(f, B, \int_0^t f_s dB_s) \stackrel{law}{=} (g, W, \int_0^t g_s dW_s),$$

i.e. they have the same probability distribution.

See exercise 5.16 Revuz-Yor.

7.7 Markov process and Transition function (Lecture 26)

Definition 7.8 A family of $\mathcal{F}_t$ adapted stochastic process X_t is a Markov process if for all real valued bounded Borel measurable function f and for all $0 \le s \le t$,

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

It is strong Markov if for all stopping time τ and $t \ge 0$, $\mathbf{E}\{f(X_{\tau+t})|\mathcal{F}_\tau\} = \mathbf{E}\{f(X_{\tau+t})|X_\tau\}$.

Definition 7.9 Let $(E, \mathcal{B}(E))$ be a measurable space. A family or probability measures, $\{P(s, x, \cdot), 0 \le s < \infty, x \in E\}$, on E is a Markov transition function for a time homogeneous Markov process, if

1. For all $0 \le s$ and $x \in E$, $A \mapsto P(s, x, A)$ is a probability measure on $\mathcal{B}(E)$.
2. For all $x \in E$, $P(0, x, \cdot) = \delta_x$, the $delta$ measure at x.
3. For all $A \in \mathcal{B}(E)$, $(s, x) \mapsto P(s, x, A) : [0, \infty) \times E \to \mathbf{R}$ is bounded and Borel measurable.
4. For all $0 \le s \le t \le u$, all $x \in E$, and $A \in \mathcal{B}(E)$,

   $$P(s + t, x, A) = \int_{y \in E} P(s, x, dy) P(t, y, A).$$

   This equation is the Chapman-Kolmogorov equation.

If f a bounded integrable function, we define

$$P_t f(x) = \int f(y) P(t, x, dy). \tag{7.4}$$

Definition 7.10 Let B be a Banach space of functions. A family of bounded linear operators P_t is a semigroup if P_0 is the identity map, $P_{s+t} = P_s P_t$ for any $0 \le s, t$.

Let (P_t) be defined by formula (7.4). Then it is a semigroup of bounded linear operators on the space of bounded measurable functions, c.f. condition (2) and the Chapman-Kolmogorov equation.

Definition 7.11 An $\mathcal{F}_t$ adapted stochastic process (X_t) is a time homogeneous Markov process w.r.t. $\mathcal{F}_t$ and with Markov transition probabilities $P_{s,t}$, if for $0 \le s \le t$ and for all $f : (E, \mathcal{B}(E)) \to \mathbf{R}$ bounded measurable

$$\mathbf{E}\{f(X_{t+s})|\mathcal{F}_s\} = P_t f(X_s) \equiv \int f(y) P(t, X_s, dy). \tag{7.5}$$

It is strong Markov if for any stopping time τ,

$$\mathbf{E}\{f(X_{s+\tau})|\mathcal{F}_\tau\} = P_s f(X_\tau).$$

The probability measure $\nu(\cdot) = P(X_0 \in \cdot)$ is the initial distribution of the Markov process.

Exercise 7.2 Prove that (7.5) implies the the Chapman-Kolmogorov equation.

Exercise 7.3 Let (X_t) be a Markov process with probability transition function $P(t, x, \cdot)$ and initial distribution ν. Prove that for $A_0, A_1, \dots, A_k$ Borel measurable sets of E and $0 \le t_1 < \dots < t_k$,

$$P\left(X_0 \in A_0, X_{t_1} \in A_1, \dots, X_{t_k} \in A_k\right)$$
$$\int_A \int_{A_1} \dots \int_{A_K} P(t_k - t_{k-1}, x_{k-1}, dx_k) \dots P(t_2 - t_1, x_1, dx_2) P(t_1, x_0, dx_1) \nu(dx_0). \tag{7.6}$$

If $E = \mathbf{R}^n$, there exists a Markov process (X_t) whose finite dimensional distribution is determined by (7.6), (7.6) is equivalent to the following: for any finite family of bounded measurable functions $f_i : \mathbf{R}^d \to \mathbf{R}$,

$$\begin{aligned} &\mathbf{E}\Pi_{i=0}^k f_i(X_{t_i}) \\ &= \int_{\mathbf{R}^d} f_0(x_0)\nu(dx_0) \int_{\mathbf{R}^d} f_1(x_1) P(t_1, x_0, dx_1) \dots \int_{\mathbf{R}^d} f_k(x_k) P(t_k - t_{k-1} x_{k-1}, dx_k) \\ &= \int_{\mathbf{R}^d} f_0(x_0) P_{t_1}\left[f_1(x_0) \dots P_{t_k - t_{k-1}} f_k(x_{k-1})\right] \nu(dx_0). \end{aligned}$$

Remark 7.1 If (X_t) is a Markov process we may be tempted to define: $\{T_{s,t}, t \ge s \ge 0\}$ on Bounded measurable functions by

$$T_{s,t} f(x) = \mathbf{E}\{f(X_t)|X_s = x\}.$$

Indeed this is a good proposal needing some thoughts. The conditional expectation $\mathbf{E}\{f(X_t)|X_s\}$ is defined up to a set of measure zero and this set of measure zero may differ when the function f is changed. For the object $\mathbf{E}\{f(X_t)|X_s = x\}$ to be well defined we only need to consider matters related to regular conditional probabilities, such considerations are in general quite messy. However if we begin with a transition function, we have no more problem.

7.7.1 Semigroup and Generators

Definition 7.12 A semigroup is strongly continuous, if $\lim_{t\to 0} T_t f = f$ for every $f \in B$. It is a contraction semigroup if $\|P_t\| \leq 1$.

Definition 7.13 The infinitesimal generator of a semigroup T_t of bounded linear operates, defined on a Banach space B, is the linear operator $\mathcal{L}$ given by the formula:

$$\mathcal{L}f = \lim_{t\to 0} \frac{P_t f - f}{t}.$$

The domain of $\mathcal{L}$ is the set of $f \in B$ such that the above limit exists.

Proposition 7.9 *Let P_t be a strongly continuous semigroup on a Banach space B with infinitesimal generator $\mathcal{L}$. Let $t \geq 0$. The following holds*

1. *If $f \in B$, then $\int_0^t P_s f ds \in \mathrm{Dom}(\mathcal{L})$ and*

$$P_t f - f = \mathcal{L}\left(\int_0^t P_s f ds\right).$$

2. *If $f \in \mathrm{Dom}(\mathcal{L})$ then $P_t f \in \mathrm{Dom}(\mathcal{L})$, then*

$$P_t f - f = \int_0^t \mathcal{L}(P_s f) ds, \qquad \frac{d}{dt} P_t f = \mathcal{L}(P_t f) = P_t(\mathcal{L} f).$$

We will review a number of properties concerning the Markov transition function. Let $E = \mathbf{R}$ for simplicity. If $P_{s,t}(x, dy)$ is absolutely continuous with respect to dy its density is denoted by $p(s, x, t, y)$.

Theorem 7.10 *If the transition density $p(s, x, t, y)$ of a diffusion process is measurable in all its arguments, then for each $s < u < t, x$ and for almost all y,*

$$\int p(s, x, u, z) p(u, z, t, y) dz = p(s, x, t, y).$$

Theorem 7.11 *Assume that the transition density $p(s, x, t, y)$ of a diffusion process satisfies:*

1. *For $0 \leq s < t$, $t - s > \delta > 0$, $p(s, x, t, y)$ is continuous and bounded in s, t, x.*

2. *p is twice differentiable in x and once differentiable in t. Then for $0 < s < t$, $p(s, x, t, y)$ satisfies the* **backward Kolmogorov equation***:*

$$\frac{\partial p}{\partial t}(s, x, t, y) = -b(s, x)\frac{\partial p}{\partial x}(s, x, t, y) - \frac{1}{2}\sigma(s, x)\frac{\partial^2 p}{\partial x^2}(s, x, t, y).$$

In the backward Kolmogorov equation we differentiate the x-variable. An integration by parts of the formula gives,

Theorem 7.12 *Assume that the transition density $p(s,x,t,y)$ of a diffusion process is such that the partial derivatives $\frac{\partial p}{\partial s}(s,x,t,y)$, $\frac{\partial p}{\partial y}(s,x,t,y)$, $\frac{\partial^2 p}{\partial y^2}(s,x,t,y)$ exist. Assume that $\frac{\partial \sigma}{\partial x}(s,x)$ and $\frac{\partial b}{\partial x}(s,x)$ exist. Then the Kolomogrov's equation (***Fokker-Plank equation***) holds for $s<t$:*

$$\frac{\partial p}{\partial t}(s,x,t,y) = -\frac{\partial}{\partial y}\left(b(s,x)p(s,x,t,y)\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(\sigma(s,x)p(s,x,t,y)\right).$$

7.7.2 Solutions of SDE as Markov process

We consider the time homogeneous SDE $E(\sigma,b)$.

Theorem 7.13 *If uniqueness in law holds, the solution is a Markov process.*

Definition 7.14 Let $(F_t(x), t<\tau(x))$ be a solution to the SDE with initial value x. For $f:\mathbf{R}\to\mathbf{R}$ bounded measurable we define

$$P_t f(x) = \mathbf{E} f(F_t(x))\mathbf{1}_{t<\tau(x)}.$$

We say P_t is the probability semigroup for the SDE.

In the rest of the section we assume that for each x_0, there is a unique global solution $F_t(x_0)$ to $E(\sigma,b)$. The solution is also denoted by (x_t).

Let $f:\mathbf{R}^d\to\mathbf{R}$ be C^2, we define a linear operator $\mathcal{L}: C^2(\mathbf{R}^d)\to C(\mathbf{R}^d)$ by

$$\mathcal{L}f(x) = \frac{1}{2}\sum_{i,j=1}^d \left(\sum_{k=1}^m \sigma_k^i(x)\sigma_k^j(x)\right)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{j=1}^d \frac{\partial f}{\partial x_j}(x) b^j(x).$$

The linear operator is the infinitesimal generator of the SDE $E(\sigma,b)$. We do not discuss its domain.

Exercise 7.4 If $f:\mathbf{R}^d\to\mathbf{R}$ is C^2 prove that

$$f(x_t) = f(x_0) + \sum_{j=1}^d\sum_{k=1}^m \int_0^t \frac{\partial f}{\partial x_j}(x_s)(\sigma_k^j(x_s))dB_s^k + \int_0^t \mathcal{L}f(x_s)ds. \qquad (7.7)$$

This lead to :

Lemma 7.14 • *If $f \in C^2(\mathbf{R}^d, \mathbf{R})$,*

$$M_t^f := f(x_t) - f(x_0) - \int_0^t \mathcal{L}f(x_s)ds$$

is a local martingale.

• *If (W_t) is the canonical process and $\hat{P} - (x.)_* P$. Then for any $f \in C^2(\mathbf{R}^d, \mathbf{R})$,*

$$f(W_t) - f(W_0) - \int_0^t \mathcal{L}f(W_s)ds$$

is a local martingale.

Remark 7.2 If $f \in C_K^\infty(\mathbf{R}^d, \mathbf{R})$, C^∞ smooth functions with compact supports, σ and b are continuous bounded (and Lipschitz continuous), then

$$P_t f(x) = f(x) + \int_0^t P_s(\mathcal{L}f)(x)ds.$$

Proposition 7.15 *Let u_t be a bounded regular solution to the Cauchy problem for the parabolic equation*

$$\frac{\partial}{\partial t} u_t = \mathcal{L}u_t, \quad u_0 = f.$$

Then $u_t(x) = \mathbf{E}u_0(x_t)$.

Proof We apply Itô's formula to $(t, x) \mapsto u_{T-t}(x)$.

$$u_{T-t}(x_t) = u_T(x) + \int_0^t (\frac{\partial}{\partial s}(u_{T-s}) + \mathcal{L}u_{T-s})(x_s)ds + \sum_{k=1}^m \sum_{i=1}^d \int_0^t u_{T-s}(x_s)(\sigma_k^i(x_s))dB_s^k.$$

The last term is a true martingale and vanish after taking the expectation. Note that $\frac{\partial}{\partial s}(u_{T-s}) = -\frac{\partial}{\partial r} u_r|_{r=T-s}$ and the penultimate term vanishes also. Take $t \to T$ we see that

$$u_T(x) = \mathbf{E}u_0(X_T) = P_T f(x).$$

□

The equation

$$\frac{\partial}{\partial t} U_t + \mathcal{L}u_t = 0, \ 0 \le t \le T, \quad U_T = g$$

is called Kolmogorov's forward equation. If $P_{T-t}g$ is reasonably smooth so we can apply Itô's formula, It is clear that it is the solution to Kolmogorov's forward equation.

Lemma 7.16 *1. If $f \in C^2(\mathbf{R}^d, \mathbf{R})$,*

$$M_t^f := f(x_t) - f(x_0) - \int_0^t \mathcal{L}f(x_s)ds$$

is a local martingale.

2. If (W_t) is the canonical process on the canonical probability space, let $\hat{P}_x = (x.)_*(P)$. Then for any $f \in C^2(\mathbf{R}^d, \mathbf{R})$,

$$f(W_t) - f(W_0) - \int_0^t \mathcal{L}f(W_s)ds$$

is a local martingale.

This can be checked by Itô's formula.

7.8 Existence of Solutions and The Martingale Problem

Definition 7.15 A probability measure μ on the canonical space $(W^d, \mathcal{B}(W^d)$ solves the martingale problem for $\mathcal{L}$ if for any $f \in C_K^2$,

$$f(W_t) - f(W_0) - \int_0^t \mathcal{L}f(W_s)ds$$

is a local martingale.

The Martingale problem method has been used and developed by Stroock and Varadhan.

Theorem 7.17 (Ikedat-Watanabe, pp169) *The existence of solutions to equation $E(\sigma, b)$ is equivalent to the marginal problem associated to $\mathcal{L}$ is solvable.*

The idea of the proof is to construct Brownian motions B_t^i. Let $f(x) = x_i$ and $f(x) = x_i x_j$. Compute $\mathcal{L}f$. Then $M_t^i = x_t^i - x_0^i - \int_0^t b(x_s)ds$ is a local martingale, it is essentially

$$\sum_k \int_0^t \sigma_k^i(x_s)dB_s^k.$$

One can compute the brackets $\langle M^i, M^j \rangle_t$, which is $\int_0^t \sum \sigma_k^i \sigma_k^j(x_s)ds$.

7.9 Localisation

Theorem 7.18 *If σ_k and b are locally Lipschitz continuous and grows at most linearly there is a unique global strong solution to $E(\sigma, b)$.*

Proof Let $\tau_n = \inf_t\{|x_t| \ge n\}$. Let $f(x) = |x|^2$. Then $|x_t|^2 - |x_0|^2 - \int_0^t \mathcal{L}f(x_s)ds$ is a local martingale. Check that there exists C s.t. $\mathcal{L}f(x) \le C(1 + |x|^2)$. This gives,

$$\begin{aligned}\mathbf{E}|x_{t\wedge\tau_n}|^2 &\le \mathbf{E}|x_0|^2 + \mathbf{E}\int_0^{t\wedge\tau_n} C(1+|x_s|^2)ds \\ &\le \mathbf{E}|x_0|^2 + Ct + C\mathbf{E}\int_0^t |x_{s\wedge\tau_n}|^2 ds.\end{aligned}$$

By Grownall,

$$\mathbf{E}|x_{t\wedge\tau_n}|^2 \le C(\mathbf{E}|x_0|^2 + Ct)e^{CT}.$$

By Fatou, $\mathbf{E}|x_t|^2 < \infty$. So $|x_t| < \infty$ almsot surely. □

Theorem 7.19 *Suppose that the coefficients of the SDE are locally Lipschitz continuous. For each initial value x_0 there is a unique strong solution $(x_t, t < \tau)$. Furthermore $\lim_{t\uparrow\tau} |x_t| = \infty$ on $\{\omega : \tau(\omega) < \infty\}$.*

Proof Write $b = (b_1, \dots, b_d)$ in components. For each N, let b_j^N, $j = 1, \dots, d$, be a globally Lipschitz continuous function with $b_j^N = b^N$ if $|x| \le N$ and $b_j^N = 0$ if $|x| > N + 1$. Let $b^N = (b_1^N, \dots, b_d^N)$. We also define a sequence of σ^N in the same way. Let x_t^N be the unique strong global solution to the SDE:

$$dx_t^N = \sigma^N(x_t^N)dB_t + b^N(x_t^N)dt.$$

Let τ^N be the first time that $|x_t^N|$ is greater or equal to N. Then x_t^N agrees with x_t^{N+1} for $t < \tau^N$ and τ^N increases with N. Define

$$x_t(\omega) = x_t^N(\omega), \qquad \text{on } \{\omega : t < \tau^N(\omega)\}.$$

Then x_t is defined up to $t < \tau$ where

$$\tau = \sup_N \tau_N.$$

Note that the exit time from B_N of x_t is τ^N and $\lim_{t\uparrow\tau(\omega)} x_t(\omega) = \infty$ on $\{\tau(\omega) < \infty\}$. The sets $\Omega_N = \{\omega : t < \tau^N(\omega)\}$ increase with N. Let $\bar{\Omega} = \cup_N\{\omega : t < \tau^N(\omega)\}$.

$$P(\bar{\Omega}) = \lim_{N\to\infty} P(\{\omega : t < \tau^N(\omega)\}).$$

On $\{t < \tau(x_0)\} = \cup_N \Omega_N$, we have patched up a solution x_t.

It is now clear that x_t is a maximal solution for $E(\sigma, b)$. It is a strong solution.

□

Theorem 7.20 *Suppose that for* $k = 1, \dots, m$, $\sigma_k : \mathbf{R}^d \to \mathbf{R}^d$ *and* $b : \mathbf{R}^d \to \mathbf{R}^d$ *are locally Lipschitz continuous , i.e. for each* $N \in \mathcal{N}$, *there exists a number* K_N *such that for all* x, y *with* $|x| \le N, |y| \le N$,

$$|\sigma_k(x) - \sigma_k(y)| \le K_N|x - y|, \qquad |b(x) - b(y)| \le K_N|x - y|.$$

Then pathwise uniqueness holds.

Proof Let us prove the case of $d = 1, m = 1$, and let $(x_t, t < \tau)$ and $(y_t, t < \tilde{\tau})$ be two solutions to

$$x_t = x_0 + \int_0^t \sigma(x_s) dB_s + \int_0^t b(x_s) ds$$

with the same initial value and the same Brownian motion, and life times τ and $\tilde{\tau}$. Let

$$\tau_N = \inf_{t>0}\{|x_t| \ge N\}, \tau_N = \inf_{t>0}\{|y_t| \ge N\}.$$

Then

$$x_{t\wedge\tau^N} = x_0 + \sum_{k=1}^m \int_0^{t\wedge\tau^N} \sigma_k(x_s) dB_s^k + \int_0^{t\wedge\tau^N} b(x_s) ds.$$

Furthermore $\tau = \sup_N \tau^N$. On $\tau < \infty$, $\lim_{N\to\infty} |x_{\tau_N}| = \infty$.

Let $\zeta_N = \min\{\tau_N, \tilde{\tau}_N\}$ and let $T > 0$. For $t \le T$,

$$\begin{aligned}
|x_{t\wedge\zeta_N} - y_{t\wedge\zeta_N}|^2 &= \left(\int_0^{t\wedge\zeta_N} (\sigma(x_s) - \sigma(y_s)) dB_s + \int_0^{t\wedge\zeta_N} (b(x_s) - b(y_s)) ds\right)^2 \\
&\le 2\left(\int_0^{t\wedge\zeta_N} (\sigma(x_s) - \sigma(y_s)) dB_s\right)^2 + 2\left(\int_0^{t\wedge\zeta_N} (b(x_s) - b(y_s)) ds\right)^2 \\
&\le 2\left(\int_0^t (\sigma(x_{s\wedge\zeta_N}) - \sigma(y_{s\wedge\zeta_N})) dB_s\right)^2 + 2(t\wedge\zeta_N)\int_0^{t\wedge\zeta_N} \Big(b(x_{s\wedge\zeta_N}) - b(y_{s\wedge\zeta_N})\Big)^2 ds.
\end{aligned}$$

In the last step we applied Hölder's inequality,

$$\int_0^t |f(s)g(s)| ds \le \left(\int_0^t |f(s)|^p ds\right)^{\frac{1}{p}} \left(\int_0^t |g(s)|^q ds\right)^{\frac{1}{q}}.$$

with $p, q = 2$. Apply Burkholder-Davies-Gundy inequality to see that, for some constant C,

$$\begin{aligned}
&\mathbf{E}\left(\int_0^t \Big(\sigma(x_{s\wedge\zeta_N}) - \sigma(y_{s\wedge\zeta_N})\Big)dB_s\right)^2 \\
\le C\mathbf{E}&\left\langle \int_0^\cdot \Big(\sigma(x_{s\wedge\zeta_N}) - \sigma(y_{s\wedge\zeta_N})\Big)dB_s, \int_0^\cdot \Big(\sigma(x_{s\wedge\zeta_N}) - \sigma(y_{s\wedge\zeta_N})\Big)dB_s \right\rangle_t \\
&= C\mathbf{E}\int_0^t \Big(\sigma(x_{s\wedge\zeta_N}) - \sigma(y_{s\wedge\zeta_N})\Big)^2 ds.
\end{aligned}$$

$$\mathbf{E}|x_{t\wedge\zeta_N} - y_{t\wedge\zeta_N}|^2 \le 2C\mathbf{E}\left(\int_0^t (\sigma(x_{s\wedge\zeta_N}) - \sigma(y_{s\wedge\zeta_N}))^2 ds\right) + 2t\int_0^t \mathbf{E}\Big(b(x_{s\wedge\zeta_N}) - b(y_{s\wedge\zeta_N})\Big)^2 ds.$$

By local Lipschitz continuity,

$$\mathbf{E}|x_{t\wedge\zeta_N} - y_{t\wedge\zeta_N}|^2 \le \Big(2C(K_N)^2 + 2(K_N)^2 t\Big)\int_0^t \mathbf{E}|x_{s\wedge\zeta_N} - y_{s\wedge\zeta_N}|^2 ds.$$

By Grownall's inequality, for all $t \le T$, $\mathbf{E}|x_{t\wedge\zeta_N} - y_{t\wedge\zeta_N}|^2 = 0$. Since T is arbitrary, $\mathbf{E}|x_{t\wedge\zeta_N} - y_{t\wedge\zeta_N}|^2 = 0$ for all t. This implies that $x_t = y_t$ on $t < \tau_N \wedge \tilde{\tau}_N$. By the sample continuity of (X_t) and (y_t), we see that $\tau_N = \tilde{\tau}_N$ and and $x_t = y_t$ on $t < \tau_N \wedge \tilde{\tau}_N$. This $\tau = \tilde{\tau}$ and $x_t = y_t$ for all $t < \tau$.

□

Chapter 8

Girsanov Transform

Let $(\Omega, \mathcal{F})$ be a measurable space. Let Q, P be probability measures with Q absolutely continuous with respect to P (This is denoted by $Q << P$). Then there exists a random variable $D \in L^1(\Omega, \mathcal{F}, P)$ such that for any $X : \Omega \to \mathbf{R}$ bounded measurable,

$$\int_\omega X dQ = \int_\Omega X D dP.$$

If there is risk of confusion, we denote by $\mathbf{E}^P$ taking expectation with respect to P. Note that $\mathbf{E}^P(D) = 1$. If, furthermore, $\frac{1}{D} \in L^1(\Omega, \mathcal{F}, Q)$, then Q is equivalent to P. Indeed, for any $A \in \mathcal{F}$

$$P(A) = \int_A \left(\frac{1}{D} D\right) dP = \int_A \frac{1}{D} dQ.$$

If A has measure $Q(A) = 0$, the integral on the right hand side vanishes, and $P(A) = 0$. Conversely if P is equivalent to Q, then $\frac{dP}{dQ} = \frac{1}{D}$ and $\frac{1}{D} \in L^1(\Omega, \mathcal{F}, Q)$.

8.1 Girsanov Theorem For Martingales (Lecture 28)

Let P and Q be equivalent probability measures on $(\Omega, \mathcal{F}, P)$. Let $f = \frac{dQ}{dP}$. Then $f \ge 0$ and $\mathbf{E} f = 1$. Let $f_t = \mathbf{E}(f|\mathcal{F}_t)$. Then (f_t) is a strictly positive integrable martingale. Let $(\mathcal{F}_t)$ be a complete and right continuous filtration. Let $D \in L^1(\Omega, \mathcal{F}, P)$. We define

$$D_t = \mathbf{E}\{D|\mathcal{F}_t\}.$$

We take D_t to be a cadlag version. Then $(M_t, t < \infty)$ is a closed martingale. See Proposition 4.15.

Proposition 8.1 *Let (f_t) be a strictly positive continuous local martingale. There is a continuous local martingale N_t s.t.*

$$f_t = e^{N_t - \frac{1}{2}\langle N,N\rangle_t}.$$

Proof By Itô's formula,

$$\log f_t = \log f_0 + \int_0^t \frac{df_s}{f_s} - \frac{1}{2}\int_0^t \frac{1}{(f_s)^2} d\langle f, f\rangle_s.$$

Let $N_t = \log f_0 + \int_0^t \frac{1}{f_s} df_s$, then

$$\langle N, N\rangle_t = \frac{1}{2}\int_0^t \frac{1}{(f_s)^2} d\langle f, f\rangle_s,$$

and $\log(f_t) = N_t - \frac{1}{2}\langle N, N\rangle_t$. □

Let (N_t) be a continuous local martingale. We define Q on $\mathcal{F}_t$ such that $\frac{dQ}{dP} = e^{N_t - \frac{1}{2}\langle N,N\rangle_t}$. If $\mathbf{E}(e^{N_t - \frac{1}{2}\langle N,N\rangle_t}) = 1$ then Q is a probability measure on $(\mathcal{F}_t)$. This is equivalent to $(e^{N_t - \frac{1}{2}\langle N,N\rangle_t}, s \leq t)$ is a true martingale.

Theorem 8.2 (Novikov criterion) *Let (N_t) be a continuous local martingale. The exponential martingale $e^{N_t - \frac{1}{2}\langle N,N\rangle_t}$ is a true martingale if $\mathbf{E}\left(e^{\frac{1}{2}\langle N,N\rangle_t}\right) < \infty$ for all $t \geq 0$.*

8.2 Girsanov for Martingales

Let P and Q be equivalent probability measures on $(\Omega, \mathcal{F}, P)$ and $(\mathcal{F}_t)$ a standard filtration. Then

$$\frac{dQ}{dP} = e^{N_t - \frac{1}{2}\langle N,N\rangle_t}$$

for a continuous local martingale (N_t).

Proposition 8.3 *Let $f_t = e^{N_t - \frac{1}{2}\langle N,N\rangle_t}$ where (N_t) is a continuous local martingale. Let (M_t) be a continuous local martingale prove that*

$$\int_0^t \frac{1}{f_s} d\langle f, M\rangle_s = \langle M, N\rangle_t.$$

Proof If (M_t) is a sample continuous local martingale, then

$$\langle f, M \rangle_t = \int_0^t f_s d\langle M, N \rangle_s.$$

This follows from $df_t = f_t dN_t$. In particular,

$$\int_0^t \frac{1}{f_s} d\langle f, M \rangle_s = \langle M, N \rangle_t.$$

□

Theorem 8.4 (Girsanov Theorem) *Let (M_t) be a continuous $(\mathcal{F}_t)$-martingale w.r.t. P. Then*

$$\tilde{M}_t = M_t - \langle M, N \rangle_t$$

ia a Q-martingale.

Proof We only need to prove the case of $f_0 = 1$ in which case $N_0 = 0$. Let $s < t$ and $A \in \mathcal{F}_s$, we prove that

$$\int_A \tilde{M}_t dQ = \int_A \tilde{M}_s dQ.$$

Equivalently,

$$\int_A \tilde{M}_t f_s t dP = \int_A \tilde{M}_s f_s dP.$$

It is sufficient to prove that $\tilde{M}_t f_t = M_t f_t - \langle M, N \rangle_t f_t$ is a P-martingale. We apply Itô's formula,

$$\begin{aligned}\langle M, N \rangle_t f_t &= \int_0^t f_s d\langle M, N \rangle_s + \int_0^t \langle M, N \rangle_s df_s \\ &= \langle f, M \rangle_t + \int_0^t \langle M, N \rangle_s df_s.\end{aligned}$$

Thus

$$\begin{aligned}\tilde{M}_t f_t &= M_t f_t - \langle M, N \rangle_t f_t \\ &= M_t f_t - \langle f, M \rangle_t - \int_0^t \langle M, N \rangle_s df_s.\end{aligned}$$

Both $M_t f_t - \langle f, M \rangle_t$ and $\int_0^t \langle M, N \rangle_s df_s$, the latter is a stochastic integral wi.r.t. to a martingale, are martingales w.r.t. P. This completes the proof. □

Corollary 8.5 *Let (B_t) be a Brownian motion with respect to P then $\tilde{B}_t = B_t - \langle B, N\rangle_t$ is a Q-Brownian motion.*

Proof This is clear by Lévy's characterisation theorem: $(\tilde{B}_t)$ is a BM with martingale with $\langle \tilde{B}, \tilde{B}\rangle_t = t$.

□

Example 8.1 Let $\sigma, b : \mathbf{R} \to \mathbf{R}$ be Borel measurable and such that there is a unique (in law) solution to the SDE

$$dx_t = \sigma(x_t)dB_t + b(x_t)dt. \tag{8.1}$$

Then for any Borel measurable set A,

$$P(X_t \in A) = \mathbf{E}\left(\mathbf{1}_{B_t \in A} e^{\int_0^t b(x_0+B_s)dB_s - \frac{1}{2}\int_0^t b^2(x_0+B_s)ds}\right).$$

Proof Let $N_t = \int_0^t b(y_s)dB_s$. Let Q be defined by the formula:

$$\frac{dQ}{dP} = e^{\int_0^t b(x_0+B_s)dB_s - \frac{1}{2}\int_0^t b^2(x_0+B_s)ds}.$$

Then

$$\langle N, B\rangle_t = \int_0^t b(x_0 + B_s)ds.$$

Let

$$\tilde{B}_t = B_t - \int_0^t b(y_s)ds.$$

Let $y_t = x_0 + B_t$. Then

$$dy_t = dB_t = d\tilde{B}_t + b(y_t)dt.$$

By uniqueness in law (y_t) under Q has the same distribution as (x_t) under Q. That is

$$P(X_t \in A) = Q(y_t \in A) = \mathbf{E}\mathbf{1}_{\{x_0+B_t \in A\}} e^{\int_0^t b(x_0+B_s)dB_s - \frac{1}{2}\int_0^t b^2(x_0+B_s)ds}.$$

□

Chapter 9

Appendix

9.1 Lyapunov Function Test

Definition 9.1 A C^2 function $V : \mathbf{R}^d \to \mathbf{R}_+$ is a Lyapunov function, for the explosion problem associated to an infinitesimal generator $\mathcal{L}$, if

(1) $V \geq 0$,

(2) $\lim_{|x|\to\infty} |V(x)| = \infty$ and

(3) $\mathcal{L}V \leq cV + K$ for some number c.

We will use below Fatou's Lemma, $\liminf_{n\to\infty} \int f d\mu \leq \int \liminf_{n\to\infty} f_n d\mu$,

Proposition 9.1 *Assume that σ_k's are continuous and the SDE $E(\sigma, b)$ has a solution x_t. Suppose that there is a Lyapunov function V for the generator $\mathcal{L}$ the SDE does not explode.*

Proof Let τ be the life time of (x_t) and $\zeta < \tau$ a stopping time. By assumption (2), $\mathcal{L}V$ is locally bounded. Apply Itô's formula to V and $x_{t\wedge\zeta}$ Then

$$\begin{aligned} V(x_{t\wedge\zeta}) &= V(x_0) + \int_0^{t\wedge\zeta} (\mathcal{L}V)(x_s)\, ds + \sum_{k=1}^m \int_0^{t\wedge\zeta} \frac{\partial V}{\partial x_j}(x_s)\sigma_k(x_s) dB_s^k \\ &\leq V(x_0) + \int_0^{t\wedge\zeta} \Big(cV(x_s) + K\Big)\, ds + \sum_{k=1}^m \int_0^{t\wedge\zeta} \frac{\partial V}{\partial x_j}(x_s)\sigma_k(x_s) dB_s^k \\ &\leq V(x_0) + Kt + c\int_0^t \Big(V(x_{s\wedge\zeta})\Big)\, ds + \sum_{k=1}^m \int_0^{t\wedge\zeta} \frac{\partial V}{\partial x_j}(x_s)\sigma_k(x_s) dB_s^k \end{aligned}$$

Let $\tau_N = \inf\{|x_t| \geq N\}$ and take $\zeta = \tau_N$ in the above. Since dV and σ_k are continuous and therefore bounded on the ball of radius N, the local martingale is a martingale. Note that V and $\mathcal{L}V(x) \leq CV + k$ are also bounded on B_N. Taking expectation to see that

$$\mathbf{E}V(x_{t\wedge\tau_N}) \leq V(x_0) + Kt + c\int_0^t \mathbf{E}\Big(V(x_{s\wedge\tau_N})\Big)ds.$$

By Grownall' lemma,

$$\mathbf{E}V(x_{t\wedge\tau_N}) \leq [V(x_0) + Kt]e^{ct}.$$

Since

$$\begin{aligned}\mathbf{E}V(x_{t\wedge\tau^N}) &= \mathbf{E}V(x_t)\mathbf{1}_{t<\tau^N} + \mathbf{E}V(x_{t\wedge\tau_N})\mathbf{1}_{t\geq\tau^N})\\ &= \mathbf{E}V(x_t)\mathbf{1}_{t<\tau^N} + V(N)P(t \geq \tau^N),\end{aligned}$$

$$V(N)P(t \geq \tau_N) \leq \Big(V(x_0) + Kt\Big)e^{Ct}$$

By Fatou's lemma for non-negative functions,

$$P(\tau \leq t) = \mathbf{E}\mathbf{1}_{\tau\leq t} \leq \lim_{N\to\infty} P(t \geq \tau_N).$$

Since $\lim_{N\to\infty}|V(N)| = \infty$,

$$\lim_{N\to\infty} P(t \geq \tau_N) \leq \lim_{N\to\infty} \frac{1}{V(N)}[V(x_0) + kt]e^{Ct} = 0.$$

So $\tau \geq t$ for any t and there is no explosion. □

Example 9.1 1. Assume that $\sum_{k=1}^m |\sigma_k(x)| \leq c(1 + |x|^2)$ and $\langle b(x), x\rangle_{\mathbf{R}^d} \leq c(1 + |x|^2)$. Then $1 + |x|^2$ is a Lyapunov function:

$$\begin{aligned}\mathcal{L}(|x|^2 + 1) &= \sum_{k=1}^m \sum_{i=1}^d (\sigma_k^i(x))^2 + 2\sum_{l=1}^d b_l(x)x_l\\ &= \sum_{k=1}^m |\sigma_k(x)|^2 + 2\langle b(x), x\rangle \leq 2c(1 + |x|^2).\end{aligned}$$

2. Show that the SDE below does not explode,

$$\begin{aligned}dx_t &= (y_t^2 - x_t^2)dB_t^1 + 2x_t y_t dB_t^2\\ dy_t &= -2x_t y_t dB_t^1 + (x_t^2 - y_t^2)dB_t^2.\end{aligned}$$

3. Let $r(x) = |(x^1, \ldots, x^d)| = \sqrt{\sum x_i^2}$ be the radius function. Let ϕ be a harmonic function, so

$$\begin{aligned}
\phi(x) &= c_1|x| + c_2, \qquad \text{dimension} = 1\\
\phi(x) &= c_1 \log|x| + c_2, \qquad \text{dimension} = 2\\
\phi(x) &= \frac{c_1}{|x|^{n-2}} + c_2, \qquad \text{dimension} > 2
\end{aligned}$$

For dimension 1 and 2 harmonic functions can be used to build Lyapunov functions for generators of the form $C\Delta$, where C can be a function. Modify the function inside the ball of radius one so that the function is smooth. Harmonic functions in dimension 3 or greater are not useful for explosion problems.

9.1.1 Strong Completeness, flow

Suppose that path wise uniqueness holds and the SDE does not explode.

Definition 9.2 If for each point x there is a solution $F_t(x, \omega)$ to

$$dx_t = \sum_i \sigma_i(x_t) \circ dB_t^i + b(x_t)dt,$$

and there is a version of $F_t(x, \omega)$ such that $(t, x) \mapsto F_t(x)$ is continuous, we say that the SDE is strongly complete.

Definition 9.3 Let ξ be $\mathcal{F}_s$-measurable. We denote by $F_{s,t}(\xi)$ the solution to:

$$x_t = \xi + \sum_i \int_s^t \sigma_i(x_r)dB_r^i + \int_s^t b(x_r)dr.$$

For simplification let $F_t = F_{0,t}$.

Definition 9.4 Let S be a stopping time we define the shift operator: $\theta_S B = B_{S+\cdot} - B_S$.

The process $(\theta_S B)_t := B_{S+t} - B_S$ is an $\mathcal{F}_{\cdot+S}$ BM. If (B_t) is the canonical process on the Wiener space, this is $\theta_S(\omega)(t) = \omega(S+t) - \omega(S)$.

Theorem 9.2 *Given $0 \leq S \leq T$ be stopping times, assume there is a unique global strong solution $F_t(\cdot, B), t \geq 0)$ to the SDE. Then the flow property holds:*

$$F_{S,T}(F_S(x_0, B), \omega) = F_T(x, B). \tag{9.1}$$

And the Cocycle property holds:

$$F_{T-S}(F_S(x, \omega), \theta_S(\omega)) = F_T(x, \omega). \tag{9.2}$$

Proof The flow property follows from the pathwise uniqueness of the solution. □

Remark. We do not prove this. Given the existence of a strong solution and path wise uniqueness, we have the Cocycle property which implies the Markov property and the Cocycle property with stopping times implies the strong Markov property.

$$\begin{aligned}\mathbf{E}f(F_{S+t}(x, B)) &= \mathbf{EE}\left\{f(F_t(F_S(x), \theta_S(B)))|\mathcal{F}_S(x)\right\} \\ &= \mathbf{E}f(F_t(-, \theta_S(B)))(F_S(x)) = P_t f(F_S(x)).\end{aligned}$$

Bibliography

[1] Ludwig Arnold. *Random dynamical systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[2] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[3] V. I. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007.

[4] K. D. Elworthy. *Stochastic differential equations on manifolds*, volume 70 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1982.

[5] K. D. Elworthy, Y. Le Jan, and Xue-Mei Li. *On the geometry of diffusion operators and stochastic flows*, volume 1720 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1999.

[6] K. D. Elworthy, Xue-Mei Li, and M. Yor. The importance of strictly local martingales; applications to radial Ornstein-Uhlenbeck processes. *Probab. Theory Related Fields*, 115(3):325–355, 1999.

[7] K. David Elworthy, Yves Le Jan, and Xue-Mei Li. *The geometry of filtering*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2010.

[8] Gerald B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.

[9] Avner Friedman. *Stochastic differential equations and applications. Vol. 1*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975. Probability and Mathematical Statistics, Vol. 28.

[10] Avner Friedman. *Stochastic differential equations and applications. Vol. 2*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976. Probability and Mathematical Statistics, Vol. 28.

[11] Ĭ. Ī. Gīhman and A. V. Skorohod. *Stochastic differential equations*. Springer-Verlag, New York, 1972. Translated from the Russian by Kenneth Wickwire, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 72.

[12] R. Z. Has′minskiĭ. *Stochastic stability of differential equations*, volume 7 of *Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis*. Sijthoff & Noordhoff, Alphen aan den Rijn, 1980. Translated from the Russian by D. Louvish.

[13] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1989.

[14] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[15] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[16] H. Kunita. *Lectures on stochastic flows and applications*, volume 78 of *Tata Institute of Fundamental Research Lectures on Mathematics and Physics*. Published for the Tata Institute of Fundamental Research, Bombay, 1986.

[17] Hiroshi Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.

[18] Roger Mansuy and Marc Yor. *Aspects of Brownian motion*. Universitext. Springer-Verlag, Berlin, 2008.

[19] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

[20] Bernt Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.

[21] K. R. Parthasarathy. *Probability measures on metric spaces*. AMS Chelsea Publishing, Providence, RI, 2005. Reprint of the 1967 original.

[22] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.

[23] Michael Reed and Barry Simon. *Methods of modern mathematical physics. I*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, second edition, 1980. Functional analysis.

[24] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[25] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1987. Itô calculus.

[26] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 1*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Ltd., Chichester, second edition, 1994. Foundations.

[27] Daniel W. Stroock. *An introduction to Markov processes*, volume 230 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 2005.

[28] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979.

[29] S. R. S. Varadhan. *Stochastic processes*, volume 16 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York, 2007.

[30] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

[31] Marc Yor. *Some aspects of Brownian motion. Part II*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1997. Some recent martingale problems.