B8.2 Continuous martingales and stochastic calculus

March 8, 2026

Contents

The structure of this course and the notes are based on the lecture notes written by Ben Hambly and the previous lecturers. Useful references for further readings include:

1) N. I. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*. 2nd Edition, North-Holland, Kodansha (1989). This is the main reference I have used for preparing this course. [Chapter 1 (Sections 5, 6 and 7), and Chapter 2 (Sections 1, 2, 4 and 5)]

2) I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*, Springer (2nd ed.), 1991, Chapters 1-3.

3) B. Oksendal, *Stochastic Differential Equations: An introduction with applications*, 6th edition, Springer (Universitext), 2007. Chapters 1 - 3.

4) S.N. Cohen and R.J. Elliott, *Stochastic Calculus and Applications*, Birkhöuser, 2015, Chapters 1-5, 8-12

5) C. Dellacheris and P. A. Meyer, Probabilités et Potentiels, 2e édition, chapitres V-VIII. Hermann, Paris (1980). [Of course, there is no better place to learn the theory of martingales and stochastic integration than the work of P. A. Meyer].

To revise material from B8.1, you might want to look at

i. D. Williams, *Probability with Martingales*, Cambridge, 1991.

ii. *Lectures Notes* for B8.1.

1 Introduction

Our topic is part of the huge field devoted to the study of *stochastic processes*. Since first year, you have had the notion of a *random variable*. In this course, we want to think of random processes, that is random variables that evolve in time.

When we model deterministic quantitities that evolve with (continuous) time, such as particles moving under gravity or some other force, we often appeal to *dynamical systems* (in discrete time or in continuous time) as models. In this course we develop a *calculus* necessary to develop an analogous theory of *random dynamical systems* in continuous time, i.e. a theory of *stochastic (ordinary) differential equations* (SDEs). These can be used to model quantities such as the prices of assets or particles moving in turbulent fluids, where the randomness is a key part of the evolution.

An ordinary differential equation might take the form

$$\frac{dX(t)}{dt} = b(t, X(t)),$$

for a suitably nice function a. Thinking of this as an infinitesimal evolution equation we could write

$$dX(t) = b(t, X(t))dt,$$

to describe how the quantity X changes in an infinitesimal amount of time dt. A stochastic differential equation is often formally written as

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dB_t,$$

where the second term on the right models 'noise' or fluctuations, through a function σ and a noise dB which we can think of as the infinitesimal change in a random process B, that is an 'infinitesimal normal distribution' with mean 0 and variance dt. In order to try to make sense of this we can write it equivalently in an integral form:

$$X(t) = X(0) + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s, X(s))dB_s.$$

Here the random process $(B_t)_{t\geq 0}$ is a fundamental object that we call Brownian motion. The punchline of the course is that we can give a rigorous meaning to the last term in the equation and although we will consider what appear to be more general driving noises, under rather general conditions they can all be built from Brownian motion. Indeed if we added possible (random) 'jumps' in $X(t)$, we would capture essentially the most general theory. We are not going to allow jumps, so we will be thinking of settings in which our stochastic equation has a continuous solution $t \mapsto X_t$.

2 Stochastic processes

In this course we shall handle random variables, which means we have a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and by a random variable we mean a mapping $X : \Omega \mapsto E$, where E is called a state space, which is equipped a σ-algebra $\mathscr{E}$. It is required that X is a *measurable* mapping with respect to $\mathscr{F}$ and $\mathscr{E}$, in the sense that for each $U \in \mathscr{E}$, $X^{-1}(U) \in \mathscr{F}$. Hence we can assign a probability to the event that $X \in U$. Often

$(E, \mathscr{E})$ is just $(\mathbb{R}, \mathscr{B}(\mathbb{R})$ (where $\mathscr{B}(\mathbb{R})$ is the Borel σ-algebra on $\mathbb{R}$, the σ-algebra of all Borel measurable subsets) and this just says that for each $x \in \mathbb{R}$ we can assign a probability to the event $\{X \leq x\}$. By definition, the distribution of X is the probability measure μ_X on $(E, \mathscr{E})$ defined by $\mu_X(A) = \mathbb{P}[X \in A]$ for every $A \in \mathscr{E}$.

A *stochastic process*, indexed by some set $\mathbb{T}$[1], is a collection of random variables $(X_t)_{t \in \mathbb{T}}$ defined on a common probability space $(\Omega, \mathscr{F}, \mathbb{P})$, taking values in a common *state space* E (equipped with a σ-algebra $\mathscr{E}$). We also use $\{X_t : t \in \mathbb{T}\}$ or X_t for $t \in \mathbb{T}$, or simply by X (if the time parameter is clear) to mean a process $(X_t)_{t \in \mathbb{T}}$. For us, $\mathbb{T}$ will generally be either $[0, \infty)$ or $[0, T]$ and we think of X_t as the state of the random evolution at instant $t \in \mathbb{T}$. For each sample point $\omega \in \Omega$, the mapping $t \mapsto X_t(\omega)$, represents a *realisation* of our stochastic process[2], called a *sample path* or *trajectory*.

We will work with real-valued or $\mathbb{R}^d$-valued processes, i.e. the state space $E = \mathbb{R}^d$ with the σ-algebra $\mathscr{E} = \mathscr{B}(\mathbb{R}^d)$, the Borel σ-algebra on $\mathbb{R}^d$, which is the smallest σ-algebra containing all open sets.

An $\mathbb{R}^d$-valued process $X = (X_t)_{t \geq 0}$ is continuous (resp. right continuous; resp. left continuous) if for almost all $\omega \in \Omega$, the sample paths $t \mapsto X_t(\omega)$ is continuous at any $t \geq 0$ (resp. right continuous at all $t \geq 0$; resp. left continuous at every $t > 0$)[3].

In discrete time, we often made statements which held 'almost surely', that is, up to a set of measure zero. In continuous time, we need to be more careful with what this means:

Definition 2.1. *Let $X = (X_t)_{t \geq 0}, Y = (Y_t)_{t \geq 0}$ be two stochastic processes defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.*

[1] In this course, $\mathbb{T} \subset \mathbb{R}$ which is a ordered one dimensional subset. If $\mathbb{T}$is finite (studied in the elementary probability theory, Prelims and Part A Probability) or countable infinite, then a stochastic process is also called a random sequence (paper B8.1).

[2] A realization of a stochastic process $X = (X_t)_{t \geq 0}$, is therefore an element of the space of paths in E. Let us use $E^{\mathbb{T}}$ to denote the collection of all mappings (paths in E), $w : \mathbb{T} \mapsto E$, then $X = (X_t)_{t \geq 0}$ can be considered as a mapping from Ω to $E^{\mathbb{T}}$ by sending each sample point ω to the trajectory $t \in \mathbb{T} \mapsto X_t(\omega)$. In the discrete time case, i.e. $\mathbb{T}$ is finite or countable infinite, $E^{\mathbb{T}}$ can be equipped with the product σ-algebra $\mathscr{E} \otimes \cdots \otimes \mathscr{E} \otimes \cdots$, and X can be regarded as an $E^{\mathbb{T}}$-valued random variable. In the case where $\mathbb{T} = [0, \infty)$ or any interval of $\mathbb{R}$, $E^{\mathbb{T}}$ is a huge infinite dimensional space (and in general too big to be useful), and we shall later on define a proper σ-algebra on $E^{\mathbb{T}}$, but in general X is no longer an $E^{\mathbb{T}}$-valued random variable. This is the complication in the uncountable time parameter setting.

[3] Let $C([0, \infty); \mathbb{R}^d)$ denote the space of all continuous functions $w : [0, \infty) \mapsto \mathbb{R}^d$, called the continuous path space (or just path space). Then each trajectory of a continuous process $X = (X_t)_{t \geq 0}$ is an element of the path space $C([0, \infty); \mathbb{R}^d)$, i.e. an $\mathbb{R}^d$-valued function on $[0, \infty)$, sending t to $X_t(\omega)$ for each sample point $\omega \in \Omega$. This point-views applies to right- resp. left-continuous processes with obvious modifications. Therefore a continuous process can be regarded as a random function taking values in the continuous path space $C([0, \infty); \mathbb{R}^d)$. There is a substantial advantage of the continuous path space $C([0, \infty); \mathbb{R}^d)$ over the (total) path space $E^{[0,\infty)}$ (where $E = \mathbb{R}^d$ with its Borel σ-algebra), that is, there is a natural σ-algebra on $C([0, \infty); \mathbb{R}^d)$. Define a metric $d(w, w') = \sum_{n=1}^{\infty} 2^{-n} (\sup_{t \leq n} |w(t) - w'(t)| \wedge 1)$ as we did in Paper A2, which is the metric defining uniform convergence over any compact subset, and equip $C([0, \infty); \mathbb{R}^d)$ with the Borel σ-algebra defined by the metric.

1) We say that X is a modification *of Y if, for all $t \geq 0$, we have $X_t = Y_t$ a.s.*[4]*;*
2) We say that X and Y are indistinguishable if

$$\mathbb{P}[X_t = Y_t \text{ for all } 0 \leq t < \infty] = 1,$$

or[5] *equivalently, there is $\mathscr{N} \in \mathscr{F}$ with $\mathbb{P}(\mathscr{N}) = 0$, such that $X_t(\omega) = Y_t(\omega)$ for all t and for all $\omega \notin \mathscr{N}$.*

By definition, if $X = (X_t)_{t\in\mathbb{T}}$ and $Y = (Y_t)_{t\in\mathbb{T}}$ are modifications of one another then, in particular, they have the same finite dimensional distributions[6],

$$\mathbb{P}[(X_{t_1}, \ldots, X_{t_n}) \in A] = \mathbb{P}[(Y_{t_1}, \ldots, Y_{t_n}) \in A]$$

for any finite collection of times $\{t_1, t_2, \ldots, t_n\} \subset \mathbb{T}$, and all measurable sets A, but indistinguishability is a much stronger property.

Example 2.2. *Let $T \sim U([0,1])$ be a uniform random variable, and take the random process $X_t = \mathbf{1}_{\{T=t\}}$. Then $Y_t := 0$ is a modification of X_t, as $Y_t = X_t$ a.s. for each t. However, Y and X are not indistinguishable, as $X \neq Y$ for* some *t with positive probability (in fact, with probability 1).*

3 Theory of martingales in continuous time

The good news is that, there is essentially no new martingale theory in continuous time[7], what you have learned in the paper B8.1 is still valid for continuous parameter martingales with proper modifications. Doob's maximal inequality, Doob's up-crossing lemma and etc. can be generalized to continuous time setting without any additional work. The only new result[8] is the *regularity of sample paths*

[4]Note that $\{X_t = Y_t\}$ is measurable for each t. In general for any countable subset $D \subset [0,\infty)$, $\{X_t = Y_t \text{ for } t \in D\}$ is measurable.

[5]A difficulty arises here: in general $G := \{X_t = Y_t \text{ for all } t \in [0,\infty)\}$ is *not measurable*. Some care is needed here: that X and Y are indistinguishable really meant that G^c is a null set with respect to the outer measure associated with $\mathbb{P}$, in other words, for every $\varepsilon > 0$, there is a sequence $A_n \in \mathscr{F}$ such that $\bigcup_{n=1}^{\infty} A_n \supset G^c$ and $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \varepsilon$. If the underlying probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is complete (see the Appendix), then $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are indistinguishable if $G := \{X_t = Y_t \text{ for all } t \in [0,\infty)\}$ is measurable and $\mathbb{P}[G] = 1$.

[6]This allows us to introduce the following concept. If $X = (X_t)_{t\in\mathbb{T}}$ is a process on $(\Omega, \mathscr{F}, \mathbb{P})$ valued in E, and $Y = (Y_t)_{t\in\mathbb{T}}$ is another process on maybe a *different* probability space $(\widetilde{\Omega}, \widetilde{\mathscr{F}}, \widetilde{\mathbb{P}})$ but valued in the same state space E. Then we say X and Y have the same finite dimensional distributions if $\mathbb{P}[(X_{t_1}, \ldots, X_{t_n}) \in A] = \widetilde{\mathbb{P}}[(Y_{t_1}, \ldots, Y_{t_n}) \in A]$ for any n, $A \in \mathscr{E}^{\otimes n}$ and any finite many distinct indices $t_i \in \mathbb{T}$.

[7]Good references include: 1) Ikeda and Watanabe (Chapter 1, Section 6), Karatzas and Shreve (Chapter 1), see Appendix for a summary.

[8]For continuous time $\mathbb{T} = [0,\infty)$ we can speak of the possible limits of a stochastic process $X = (X_t)_{t\geq 0}$ at any finite time $t_0 \in [0,\infty)$, in addition to the limit of X_t as $t \to \infty$. While in discrete time setting for a random sequence $(X_n)_{n\geq 0}$ we can only talk about the possible limit of X_n as $n \to \infty$. The "convergence theorem" of martingales (or super-martingales) as $t \to t_0$ at finite t_0 (rather than $t \to \infty$) leads to "the regularity theorem" of super-martingales, which is somehow new material in this course, as it is not needed in discrete-setting, Paper B8.1.

of martingales, which can be considered as a continuous time version of Doob's martingale convergence theorem. On the other hand the sample path regularity of martingales follows from Doob's crossing lemma.

The results in this section will to a large extent mirror what you proved in B8.1 for discrete parameter martingales (and we use those results repeatedly in our proofs).

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. A family $(\mathscr{F}_t)_{t\geq 0}$ of σ-algebras of sets in $\mathscr{F}$ is a *filtration* if $\mathscr{F}_s \subseteq \mathscr{F}_t$ for any $s \leq t$. (Intuitively, $\mathscr{F}_t$ corresponds to the information known to an observer at time t.) $(\Omega, \mathscr{F}, \mathbb{P})$ equipped with a filtration $(\mathscr{F}_t)_{t\geq 0}$ is called a *filtered probability space*.

A stochastic process $X = (X_t)_{t\geq 0}$ is adapted (to filtration $(\mathscr{F}_t)_{t\geq 0}$) if X_t is $\mathscr{F}_t$-measurable for every $t \geq 0$.

For a stochastic process process $X = (X_t)_{t\geq 0}$ we define $\mathscr{F}_t^X = \sigma\{X(s) : s \leq t\}$ (that is $\mathscr{F}_t^X$ is the information obtained by observing X up to time t) to be *the natural filtration*[9] associated with the process X.

Given a filtration $(\mathscr{F}_t)_{t\geq 0}$, we set $\mathscr{F}_{t+} = \cap_{s>t}\mathscr{F}_s$ for any $t \geq 0$, $\mathscr{F}_{0-} = \mathscr{F}_0$ and $\mathscr{F}_{t-} = \sigma\{\mathscr{F}_s : s < t\}$ for $t > 0$. The filtration $(\mathscr{F}_t)_{t\geq 0}$ is right continuous if $\mathscr{F}_t = \mathscr{F}_{t+}$ for $t \geq 0$. As usual $\mathscr{F}_\infty = \sigma\{\mathscr{F}_t : t \geq 0\}$.

We assume throughout that a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ is given, unless otherwise specified.

Definition 3.1. *An adapted*[10] *(real valued) stochastic process $(X_t)_{t\geq 0}$ such that $X_t \in L^1(\mathbb{P})$ (i.e. $\mathbb{E}[|X_t|] < \infty$) for any $t \geq 0$, is called*

1) a martingale if $\mathbb{E}[X_t|\mathscr{F}_s] = X_s$ for all $0 \leq s \leq t$,

2) a super-martingale if $\mathbb{E}[X_t|\mathscr{F}_s] \leq X_s$ for all $0 \leq s \leq t$,

3) a sub-martingale if $\mathbb{E}[X_t|\mathscr{F}_s] \geq X_s$ for all $0 \leq s \leq t$.

Remark. 1) If $X = (X_t)_{t\geq 0}$ is super-martingale, then $t \mapsto \mathbb{E}[X_t]$ is non-increasing, and it is sub-martingale, then $t \mapsto \mathbb{E}[X_t]$ is non-decreasing, like in the discrete-setting.

2) If $X = (X_t)_{t\geq 0}$ is a super-martingale (resp. sub-martingale; resp. martingale), and $D \subset [0, \infty)$ is a finite subset so that the elements of D can be ordered as $t_0 < t_1 < \cdots < t_N$ for some N. Define $\mathscr{G}_n = \mathscr{F}_{t_n}$ and $M_n = X_{t_n}$ (for $n = 0, \cdots, N$), then $M = (M_n)_{n\geq 0}$ (for $n > N$, we can set $\mathscr{G}_n = \mathscr{G}_N$ and $M_n = M_N$ trivially) is a $(\mathscr{G}_n)$-super-martingale (resp. sub-martingale; resp. martingale) in discrete-setting.

Exercise 3.2. *Suppose $\{Z_t : t \geq 0\}$ is an adapted process with independent increments, i.e. for all $0 = t_0 < t_1 < \cdots < t_n$, $Z_{t_i} - Z_{t_{i-1}}$ $(i = 0, 1, \cdots, n)$ are independent, and $Z_0 = 0$. The following give us examples of martingales:*

1)if $\forall t \geq 0$, $Z_t \in L^1$, then $\tilde{Z}_t := Z_t - \mathbb{E}[Z_t]$ is a martingale (with respect to its natural filtration),

[9] Various notations may be used: if $X = (X_t)_{t\geq 0}$ is an E-valued stochastic process (where E is equipped with a σ-algebra $\mathscr{E}$), then the natural filtration may be also denoted by $(\mathscr{F}_t^0)_{t\geq 0}$, or just by $(\mathscr{F}_t)_{t\geq 0}$ if no confusion may arise.

[10] If no filtration is specified a priori, then the natural filtration is applied.

2) if $\forall t \geq 0$, $Z_t \in L^2$, then $\tilde{Z}_t^2 - \mathbb{E}[\tilde{Z}_t^2]$ is a martingale,
3) if for some $\theta \in \mathbb{R}$, and $\forall t \geq 0$, $\mathbb{E}[e^{\theta Z_t}] < \infty$, then $\frac{e^{\theta Z_t}}{\mathbb{E}[e^{\theta Z_t}]}$ is a martingale.

Warning: It is important to remember that a process is a martingale *with respect to a filtration* – giving yourself more information (enlarging the filtration) may destroy the martingale property. For us, even when we don't explicitly mention it, there is a filtration implicitly assumed (usually the natural filtration associated with the process, augmented to satisfy the usual conditions).

Given a martingale (or sub-martingale) it is easy to generate many more.

Proposition 3.3. *Let $(X_t)_{t\geq 0}$ be a martingale (respectively sub-martingale) and $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex (respectively convex and increasing) such that $\mathbb{E}[|\varphi(X_t)|] < \infty$ for any $t \geq 0$. Then $(\varphi(X_t))_{t\geq 0}$ is a sub-martingale.*

Proof. Apply the conditional Jensen inequality, cf. Appendix. □

In particular, if $(X_t)_{t\geq 0}$ is martingale with $\mathbb{E}[|X_t|^p] < \infty$, for some $p \geq 1$ and all $t \geq 0$, then $|X_t|^p$ is a sub-martingale (and consequently, $t \mapsto \mathbb{E}[|X_t|^p]$ is non-decreasing).

3.1 Martingale inequalities

In this section we formulate the continuous parameter versions of Doob's martingale inequalities, established in Paper B8.1. The generalization shall be done without much additional work indeed, except for the sample path regularity of martingales.

Theorem 3.4. *If $(X_t)_{t\geq 0}$ is a martingale (or a non-negative sub-martingale) on a filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ and D is a finite or countable subset of $[0,\infty)$, then for and $T > 0$, $p \geq 1$ and $\lambda > 0$,*

$$\mathbb{P}\left[\sup_{t\in D\cap[0,T]} |X_t| \geq \lambda\right] \leq \frac{1}{\lambda^p}\mathbb{E}\left[|X_T|^p\right] \tag{1}$$

(called Doob's maximal inequality), and for any $p > 1$

$$\mathbb{E}\left[\sup_{t\in D\cap[0,T]} |X_t|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|X_T|^p\right]. \tag{2}$$

This is called Doob's L^p-inequality.

Proof. If X_t is a martingale, then $|X_t|^p$ (for every $p \geq 1$) is a non-negative sub-martingale. Let D_k ($k = 1, 2, \cdots$) be an increasing sequence of finite subsets of D such that $D_k \uparrow D$. List $D_k \cap [0,T] = \{0 = t_0 < t_1 < \cdots < t_m \leq T\}$ for some integer m. Then (X_{t_n}) is a martingale (or non-negative sub-martingale in discrete-time.

Applying Doob's maximal and L^p-inequality (in the discrete-parameter setting) to (X_{t_n}) with terminal m to obtain that

$$\mathbb{P}\left[\sup_{t\in D_k\cap[0,T]}|X_t|\geq\lambda\right]\leq\frac{1}{\lambda^p}\mathbb{E}\left[|X_{t_m}|^p\right]\leq\frac{1}{\lambda^p}\mathbb{E}\left[|X_T|^p\right]$$

and for any $p>1$

$$\mathbb{E}\left[\sup_{t\in D_k\cap[0,T]}|X_t|^p\right]\leq\left(\frac{p}{p-1}\right)^p\mathbb{E}\left[|X_{t_m}|^p\right]\leq\left(\frac{p}{p-1}\right)^p\mathbb{E}\left[|X_T|^p\right].$$

Letting $k\uparrow\infty$ and applying MCT (Monotone Convergence Theorem) to deduce (1) and (2). In fact, since $D_k\uparrow D$, $\sup_{t\in D_k\cap[0,T]}|X_t|\uparrow\sup_{t\in D\cap[0,T]}|X_t|$, which implies that $1_{\left\{\sup_{t\in D_k\cap[0,T]}|X_t|\geq\lambda\right\}}\uparrow 1_{\left\{\sup_{t\in D\cap[0,T]}|X_t|\geq\lambda\right\}}$ as $k\to\infty$, thus, according to MCT

$$\mathbb{P}\left[\sup_{t\in D\cap[0,T]}|X_t|\geq\lambda\right]=\lim_{k\to\infty}\mathbb{P}\left[\sup_{t\in D_k\cap[0,T]}|X_t|\geq\lambda\right]\leq\frac{1}{\lambda^p}\mathbb{E}\left[|X_T|^p\right].$$

Similarly we prove the second inequality. □

Exercise 3.5. *If $X=(X_t)_{t\geq0}$ is a right-continuous martingale (or a right-continuous and non-negative sub-martingale), then*

$$\mathbb{P}\left[\sup_{t\in[0,T]}|X_t|\geq\lambda\right]\leq\frac{1}{\lambda^p}\mathbb{E}\left[|X_T|^p\right]$$

for any $p\geq1$, $\lambda>0$ and $T>0$, and

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

and

$$\mathbb{E}\left[\sup_{t\in[0,T]}|X_t|\right]\leq\frac{e}{e-1}\left(1+\mathbb{E}\left[|X_T|\ln^+|X_T|\right]\right)$$

*for any $p>1$ and $T>0$. [*Hint: Let $T>0$ be any but fixed. Apply Doob's inequalities above with $D=\mathbb{Q}\cup\{T\}$*].*

Exercise. Suppose $X=(X_t)_{t\geq0}$ is a super-martingale, then

$$\mathbb{P}\left[\sup_{D\cap[0,t]}|X_t|\geq\lambda\right]\leq\frac{1}{\lambda}\left(\mathbb{E}\left[X_0\right]+2\mathbb{E}\left[X_t^-\right]\right)$$

for any $\lambda>0$ and for any countable subset $D\subset[0,\infty)$.

Hint: This is the super-martingale version of Doob's maximal inequality.

Next we state *Doob's up-crossing lemma*[11], which is again a consequence of the corresponding up-crossing lemma for discrete time setting we did in the paper B8.1. Nevertheless let us state and prove it carefully.

Let us recall the number of up-crossings. If $\{x_0, x_1, \cdots, x_N\}$ is a ordered finite subset of $\mathbb{R}$, and $a < b$. Let $T_0 = \inf\{i : x_i < a\}$ with the convention that $\inf \emptyset = \infty$, $T_1 = \inf\{i > T_0 : x_i > b\}$, then inductively define $T_{2j} = \inf\{i > T_{2j-1} : x_i < a\}$ and $T_{2j+1} = \inf\{i > T_{2j} : x_i . b\}$.. Then the number of up-crossings of (x_i) through $[a,b]$ is denoted by $U_a^b((x_i))$, and

$$U_a^b((x_i)_{i\leq N}) = \max\left\{j : T_{2j-1} \leq N\right\}.$$

Suppose $(x_i)_{i=0,1,\dots}$ is a sequence of real numbers, then clearly $N \mapsto U_a^b((x_i)_{i\leq N})$ is increasing, and $U_a^b((x_i)_{i\geq 0}) = \lim_{N\to\infty} U_a^b((x_i)_{i\leq N})$ exists (which may be ∞).

Suppose $f : [0,\infty) \mapsto \mathbb{R}$ is a real-valued function, and A is any finite subset of $[0,\infty)$. List $A = \{0 \leq t_1 < t_2 < \cdots < t_m\}$. Let $a < b$. The up-crossing number of the sequence $f(t_1)$, $f(t_2)$, $\cdots$, $f(t_m)$ through the interval $[a,b]$ is denoted by $U_a^b(f,A)$. If $D \subset [0,\infty)$ is any subset, then we define the up-crossing number of the real function f on $[0,\infty)$ through $[a,b]$ by

$$U_a^b(f,D) = \sup_{A\subset D, A \text{ is finite}} U_a^b(f,A)$$

which of course may be infinite.

If $X = (X_t)_{t\geq 0}$ is a process, then

$$U_a^b(X,D)(\omega) = U_a^b((X_t(\omega))_{t\geq 0}, D) \quad \text{for every } \omega \in \Omega.$$

Exercise 3.6. *If D is countable, D_k is an increasing sequence of finite subsets such that $D_k \uparrow D$, then $U_a^b(f,D) = \lim_{k\to\infty} U_a^b(f,D_k)$. If $X = (X_t)_{t\geq 0}$ is a real valued stochastic process, then $U_a^b(X,D)$ is non-negative (but may be ∞) random variable, and*

$$\mathbb{E}\left[U_a^b(X,D)\right] = \lim_{k\to\infty} \mathbb{E}\left[U_a^b(X,D_k)\right]$$

*[*Hint: use the fact that $k \mapsto U_a^b(f,D_k)$ is increasing. Use MCT for the second part*]*.

We are now in a position to state a version of Doob's up-crossing lemma.

Theorem 3.7. *(*Doob's up-crossing lemma*) Let $X = (X_t)_{t\geq 0}$ be a super-martingale, $D \subset [0,\infty)$ be a countable dense subset, and $a < b$. Then for any $t > s \geq 0$*

$$\mathbb{E}\left[U_a^b(X, D\cap[s,t])\right] \leq \frac{1}{b-a}\mathbb{E}\left[(X_t - a)^-\right].$$

[11] The notion of up-crossing numbers is used to characterize the convergence of real sequences. More precisely, a real sequence (a_n) has the same upper limit $\limsup a_n$ and lower limit $\liminf a_n$ (including the case the lower / upper limit may be ∞, and $-\infty$) if and only if the up-crossing number $U_a^b((a_n)) < \infty$ for any $a < b$ (where a, b are rationals). This statement should be a theorem (or an exercise) in Prelims Analysis I, but unfortunately in Oxford, the upper limits and lower limits of sequences are not covered in Prelims Analysis ! Of course, the up-crossing numbers can be replaced by the *down-crossin*g numbers of a sequence, and the same statement is still true. We count up-crossings only as a convention.

In particular[12], $U_a^b(X, D\cap[s,t]) < \infty$ *almost surely for any* $a < b$ *and* $t > s \geq 0$.

Proof. Let $0 \leq s < t$. Let D_k be an increasing sequence of finite subsets such that $D_k \uparrow D$. List

$$D_k \cup \cap[s,t] = \{s = t_0 < t_1 < \cdots < t_m \leq t\}$$

and applying Doob's up-crossing lemma for discrete case (Paper B8.1) to the super-martingale $(X_{t_n})_{n\geq 0}$, we obtain that

$$\mathbb{E}\left[U_a^b(X, D_k \cup \cap[s,t]\right] \leq \frac{1}{b-a}\mathbb{E}\left[(X_{t_m}-a)^-\right] \leq \frac{1}{b-a}\mathbb{E}\left[(X_t-a)^-\right]$$

(the last inequality is due to the fact that $(X_t - a)^-$ is a sub-martingale), letting $k \to \infty$, by MCT we have

$$\mathbb{E}\left[U_a^b(X, D \cup \cap[s,t]\right] \leq \frac{1}{b-a}\mathbb{E}\left[(X_t-a)^-\right]$$

and therefore

$$\begin{aligned}\mathbb{E}\left[U_a^b(X, D\cap[s,t])\right] &\leq \mathbb{E}\left[U_a^b(X, D \cup \cap[s,t]\right] \\ &\leq \frac{1}{b-a}\mathbb{E}\left[(X_t-a)^-\right].\end{aligned}$$

The proof is complete. ☐

Exercise 3.8. *Suppose* $X = (X_t)_{t\geq 0}$ *is a right-continuous super-martingale*[13]*, then for any* $0 \leq s < t$ *and* $a < b$ *we have*

$$\mathbb{E}\left[U_a^b(X, [s,t])\right] \leq \frac{1}{b-a}\mathbb{E}\left[(X_t-a)^-\right].$$

*[*Hint: Use the fact that, for right continuous path X, for any $a < b$ we have $U_a^b(X, [s,t]) = U_a^b(X, \mathbb{Q}\cap[s,t])$. *]*

As a consequence we have the following regularity of super-martingales.

Lemma 3.9. *Let* $X = (X_t)_{t\geq 0}$ *be a super-martingale on* $(\Omega, \mathscr{F}, \mathbb{P})$ *and* $D \subset [0,\infty)$ *be a dense and countable subset. Then for almost all* $\omega \in \Omega$, $\lim_{s\in D, s\downarrow t} X_s(\omega)$ *for all* $t \geq 0$, *and* $\lim_{s\in D, s\uparrow t} X_s(\omega)$ *for all* $t > 0$, *exist*[14].

If in addition X *is right-continuous, then for almost all* $\omega \in \Omega$, *the left-limit* $X_{t-}(\omega) = \lim_{s\uparrow t} X_s(\omega)$ *exist for all* $t > 0$.

[12] You may recall that in the discrete-time setting, the almost sure boundedness of the up-crossing number of a super-martingale is ensured if the super-martingale is bounded in L^1. Here we do not require a similar condition as $[s,t]$ is bounded.

[13] A process $X = (X_t)_{t\geq 0}$ is right-continuous, if for all most all $\omega \in \Omega$, $\lim_{s>t, s\downarrow t} X_s(\omega) = X_t(\omega)$ for all $t \geq 0$. Note that here we demand that there is a single null set works for all $t \geq 0$. More precisely, there is a null set $\mathscr{N} \subset \Omega$, such that $\lim_{s>t, s\downarrow t} X_s(\omega) = X_t(\omega)$ for all $t \geq 0$ and for all $\omega \in \Omega \setminus \mathscr{N}$. Similarly we define the notion of continuous processes, left-continuous processes. We shall return to this notion later on in lectures.

[14] This is indeed the continuous-time version of the convergence theorem for super-, sub- and martingales: martingale convergence theorem not at ∞ but at a finite time. You should appreciate that here the boundedness condition in L^1 is automatically true.

Proof. The proof is quite similar to the proof of the martingale convergence theorem in Paper B8.1. For $t > 0$ and $a < b$, let

$$W_{t,a<b} = \left\{ \omega \in \Omega : \sup_{s \in D \cap [0,t]} |X_s(\omega)| = \infty \text{ or } U_a^b(X(\omega), D \cap [0,t]) = \infty \right\}.$$

Since X_t is super-martingale, so that (see Theorem 7.6, Appendix)

$$\mathbb{P}\left[\sup_{s \in D \cap [0,t]} |X_s(\omega)| \geq \lambda \right] \leq \frac{1}{\lambda} \left(\mathbb{E}[X_0] + 2\mathbb{E}\left[X_t^-\right] \right)$$

for every $\lambda > 0$, which implies that $\sup_{s \in D \cap [0,t]} |X_s(\omega)| < \infty$ almost surely, and, according to Doob's up-crossing lemma (above), $\mathbb{P}(W_{t,a<b}) = 0$. Clearly $W_{t,a<b} \in \mathscr{F}_t$, Let $W_t = \bigcup_{a<b; a,b \in \mathbb{Q}} W_{t,a<b}$. Then $\mathbb{P}(W_t) = 0$. Now we notice that $t \mapsto W_t$ is increasing so that $W := \bigcup_{t \geq 0} W_t = \bigcup_{n=1}^{\infty} W_n$ is measurable and $\mathbb{P}(W) = 0$. For $\omega \in \Omega \setminus W$, the limits $\lim_{s \in D, s \downarrow t} X_s(\omega)$ for all $t \geq 0$, and $\lim_{s \in D, s \uparrow t} X_s(\omega)$ for all $t > 0$, exist.

We shall now prove the second part. Since $t \mapsto X_t$ is a right continuous martingale, then, according to Doob's up-crossing lemma,

$$\tilde{W}_{t,a<b} = \left\{ \omega \in \Omega : \sup_{s \in [0,t]} |X_s(\omega)| = \infty \text{ or } U_a^b(X(\omega), [0,t]) = \infty \right\}$$

is a null set, so that $\tilde{W} := \bigcup_{t \geq 0} \tilde{W}_t$ is null too. For every $\omega \in \Omega \setminus \tilde{W}$, the left limits $\lim_{s \in D, s \uparrow t} X_s(\omega)$ for all $t > 0$ exist. □

We are now in a position to prove the regularity result about martingales. To this end, we need the following fact.

Lemma 3.10. *Let $X = (X_t)_{t \geq 0}$ be a super-martingale. Let $t \geq 0$ and a sequence $r_n > t$ and $r_n \to t$ as $n \to \infty$. Then $\{X_{r_n} : n = 1, 2, \cdots\}$ is uniformly integrable.*

Proof. X is a super-martingale, $s \mapsto \mathbb{E}[X_s]$ is non-increasing and bounded below by $\mathbb{E}[X_t]$ (for $s \geq t$). Hence $\lim_{s \downarrow t} \mathbb{E}[X_s] = l$ exists. For every $\varepsilon > 0$, there is $s_0 > t$ such that $0 \leq l - \mathbb{E}[X_s] < \frac{1}{2}\varepsilon$ for any $s \in (t, s_0]$. Since $r_n \downarrow t$, there is N, $t < r_n \leq s_0$ for all $n \geq 1$. For every $C > 0$ and $n \geq N$ we have

$$\begin{aligned}
\mathbb{E}[|X_{r_n}| : |X_{r_n}| > C] &= \mathbb{E}[X_{r_n} : X_{r_n} > C] - \mathbb{E}[X_{r_n} : X_{r_n} < -C] \\
&= \mathbb{E}[X_{r_n}] - \mathbb{E}[X_{r_n} : X_{r_n} < C] - \mathbb{E}[X_{r_n} : X_{r_n} < -C] \\
&\leq \mathbb{E}[X_{r_n}] - \mathbb{E}[X_{s_0} : X_{r_n} < C] - \mathbb{E}[X_{s_0} : X_{r_n} < -C] \\
&\leq \mathbb{E}[X_{r_n}] - \mathbb{E}[X_{s_0}] + \mathbb{E}[X_{s_0} : X_{r_n} > C] - \mathbb{E}[X_{s_0} : X_{r_n} < -C] \\
&\leq \frac{\varepsilon}{2} + \mathbb{E}[|X_{s_0}| : |X_{r_n}| > C]
\end{aligned}$$

for all $n \geq N$, here the first inequality due to the super-martingale property: since $r_n \leq s_0$, $\mathbb{E}[X_{r_n} : A] \geq \mathbb{E}[X_{s_0} : A]$ for any $A \in \mathscr{F}_{r_n}$. Moreover

$$\begin{aligned}\mathbb{P}[|X_{r_n}| > C] &\leq \frac{1}{C}\mathbb{E}[|X_{r_n}|] = \frac{1}{C}\mathbb{E}\left[X_{r_n}^+ + X_{r_n}^-\right] \\ &= \frac{1}{C}\mathbb{E}\left[X_{r_n} + 2X_{r_n}^-\right] \\ &\leq \frac{1}{C}l + \frac{1}{C}\mathbb{E}\left[2X_{s_0}^-\right]\end{aligned}$$

(note that X^- is a sub-martingale), so that $\mathbb{P}[|X_{r_n}| > C] \to 0$ as $C \to \infty$, uniformly in n. Since X_{s_0} is integrable, there is $C_0 > 0$ such that

$$\mathbb{E}[|X_{s_0}| : |X_{r_n}| > C] \leq \frac{\varepsilon}{2} \quad \text{for any } C \geq C_0$$

and for all $n \geq N$, and therefore

$$\mathbb{E}[|X_{r_n}| : |X_{r_n}| > C] \leq \varepsilon \quad \text{for any } C \geq C_0$$

for all $n \geq N$. Since $\{X_{r_l} : l = 1, \cdots, N-1\}$ is uniformly integrable, so that there is $C_1 > 0$ such that

$$\mathbb{E}[|X_{r_n}| : |X_{r_n}| > C] \leq \varepsilon \quad \text{for any } C \geq C_1$$

for $n = 1, \ldots, N-1$. Therefore we may conclude that

$$\lim_{C \to \infty} \sup_n \mathbb{E}[|X_{r_n}| : |X_{r_n}| > C] = 0.$$

Thus $\{X_{r_n} : n \geq 1\}$ is uniformly integrable. □

Theorem 3.11. [15] *Let $X = (X_t)_{t \geq 0}$ be a super-martingale (resp. martingale) on a filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$, and $D \subset [0, \infty)$ be a countable dense subset. Then there is a process*[16] *$Y = (Y_t)_{t \geq 0}$ which possesses the following properties:*

1) Y is adapted to the filtration $(\mathscr{F}_{t+})$, i.e. Y_t is $\mathscr{F}_{t+}$-measurable for each $t \geq 0$, the sample paths of Y, i.e. $t \mapsto Y_t$ are right-continuous, and for almost all $\omega \in \Omega$, $Y_t(\omega) = \lim_{s \in D, s \downarrow t} X_s(\omega)$ for all $t \geq 0$;

2) for almost all $\omega \in \Omega$, the left-limit

$$Y_{t-}(\omega) = \lim_{s \in D, s \uparrow t} Y_s(\omega) = \lim_{s \in D, s \uparrow t} X_s(\omega)$$

exists for all $t > 0$, so that $Y = (Y_t)$ is right-continuous with left limits;

3) for every $t \geq 0$, $X_t \geq \mathbb{E}[Y_t | \mathscr{F}_t]$ almost surely (resp. $X_t = \mathbb{E}[Y_t | \mathscr{F}_t]$ almost surely);

4) $Y = (Y_t)_{t \geq 0}$ is a super-martingale (resp. martingale) on $(\Omega, \mathscr{F}, \mathscr{F}_{t+}, \mathbb{P})$.

[15] This is a classical result, and this version (without assumption of "usual conditions") was proved by H. Föllmer: The exit measure of a supermartingale, *Z. f. W-theorie*, 21 (1972).

[16] The "right limit" process Y of a super-martingale X constructed in this theorem depends on the countable set D.

Proof. [*The proof is not examinable*.] We shall use the notations established in the proof of the previous lemma. Let $W_{t+} = \bigcap_{s>t} W_s$. Then $W_{t+} \in \mathscr{F}_{t+}$ for any $t \geq 0$ and $\mathbb{P}(W_{t+}) = 0$. Suppose $\omega \in \Omega \setminus W_{t+}$, then there is $r > t$ such that $\omega \in \Omega \setminus W_r$, so that the limit $\lim_{s\in D, s\downarrow t} X_s(\omega)$ exists. Set

$$Y_t(\omega) = \begin{cases} \lim_{s\in D, s\downarrow t} X_s(\omega) & \text{if } \omega \in \Omega \setminus W_{t+} \\ 0, & \omega \in W_{t+} \end{cases}$$

which is therefore $\mathscr{F}_{t+}$-measurable. Let $t \geq 0$. By definition $s \mapsto W_{s+}$ is non-decreasing, which in particular implies that $Y_s(\omega) = 0$ for any $s \geq t$ and $\omega \in W_{t+}$. Therefore $Y(\omega)$ is right-continuous at t for every $\omega \in W_{t+}$. Suppose $\omega \notin W_{t+}$. Since $W_{t+} = \bigcap_{r>t} W_r = \bigcap_{r>t} W_{r+}$, there is $r_0 > t$ such that $\omega \notin W_{r_0+}$ and therefore $\omega \notin W_{r+}$ for $r \in (t, r_0]$. For every $\varepsilon > 0$, choose $0 < \delta < r_0 - t$, such that for any $s \in D$, $0 < s - t < \delta$, we have $|Y_t(\omega) - X_s(\omega)| < \varepsilon$. Therefore for any r such that $0 < r - t < \delta$ we have

$$|Y_t(\omega) - Y_r(\omega)| = \lim_{s\in D, s\downarrow r} |Y_t(\omega) - X_s(\omega)| \leq \varepsilon,$$

by definition $Y_t(\omega) = \lim_{s\downarrow t} Y_s(\omega)$ for every $\omega \in \Omega \setminus W_{t+}$. Thus Y is right continuous at $t \geq 0$. Similarly we may prove 2).

Proof of 3) for a super-martingale X. Let $t \geq 0$. Choose a sequence $r_n \in D$ such that $r_n > t$ and $r_n \to t$. Then for every $A \in \mathscr{F}_t$

$$\mathbb{E}[X_t : A] \geq \mathbb{E}[X_{r_n} : A]$$

for every n. Since (X_{r_n}) is uniformly integrable (Lemma 3.10), by Doob's convergence theorem, so that $X_{r_n} \to Y_t$ in L^1. Letting $n \to \infty$,

$$\mathbb{E}[X_t : A] \geq \mathbb{E}[Y_t : A].$$

The proof of martingale case is similar.

Proof of 4) for a super-martingale X. Let $0 \leq s < t$. Choose two sequences $s_n, t_n \in D$ such that $s < s_n < t < t_n$ such that $s_n \downarrow s$ and $t_n \downarrow t$. Then for every $A \in \mathscr{F}_{s+}$ we have

$$\mathbb{E}[X_{s_n} : A] \geq \mathbb{E}[X_{t_n} : A].$$

As in the proof of 3), both sequences (X_{s_n}) and (X_{t_n}) are uniformly integrable (Lemma 3.10), so that $X_{s_n} \to Y_s$ and $X_{t_n} \to Y_t$ in L^1. Letting $n \to \infty$ we obtain

$$\mathbb{E}[Y_s : A] \geq \mathbb{E}[Y_t : A]$$

for every $A \in \mathscr{F}_{s+}$. Hence by definition, $Y = (Y_t)_{t\geq 0}$ is a super-martingale with respect to the filtration $(\mathscr{F}_{t+})_{t\geq 0}$. □

Remark 3.12. *In the proof we have used the following fact which was proved in B8.1: Suppose $\xi_k \to \xi$ in probability, then $\xi_k \to \xi$ in L^1 if and only if the family $\{\xi_k\}$ is uniformly integrable.*

As a consequence we have the following corollary.

Corollary 3.13. *If* $X = (X_t)_{t\geq 0}$ *is a* right-continuous *super-martingale (resp. martingale with right-continuous sample paths) with respect to the filtration* $(\mathscr{F}_t)_{t\geq 0}$*, then it is also a super-martingale (resp. martingale) with respect to the filtration* $(\mathscr{F}_{t+})_{t\geq 0}$.

Proof. Since $t \mapsto X_t$ is right continuous, so that X and Y, constructed in the previous theorem, are indistinguishable. Thus the conclusion follows 4) in the previous theorem. □

Theorem 3.14. *Let* $(\mathscr{F}_t)_{t\geq 0}$ *be a* right continuous *filtration on a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$*, and* $X = (X_t)_{t\geq 0}$ *be a super-martingale with respect to* $(\mathscr{F}_t)$*. Then there is a right-continuous and adapted modification of* $X = (X_t)_{t\geq 0}$ *if and only if* $t \mapsto \mathbb{E}[X_t]$ *is right-continuous on* $[0, \infty)$.

Proof. Choose a countable dense subset D in $[0, \infty)$. Let Y be constructed in Theorem 3.11. Since $\mathscr{F}_{t+} = \mathscr{F}_t$ for every $t \geq 0$, Y is a super-martingale with respect to $(\mathscr{F}_t)_{t\geq 0}$. By 3) in Theorem 3.11, we have $X_t \geq Y_t$ almost surely, for each $t \geq 0$. Therefore $X_t = Y_t$ almost surely if and only if $\mathbb{E}(X_t - Y_t) = 0$, i.e. if and only if $\mathbb{E}(X_t) = \mathbb{E}(Y_t)$. On the other hand, for $t \geq 0$, choose a sequence $r_n \in D$, $r_n > t$ and $r_n \downarrow t$. Since (X_{r_n}) is uniformly integrable, so that $X_{r_n} \to Y_t$ in L^1, therefore $\mathbb{E}(Y_t) = \lim_{n\to\infty} \mathbb{E}(X_{r_n})$. Therefore the condition that $\mathbb{E}(X_t) = \mathbb{E}(Y_t)$ if and only if $\mathbb{E}(X_t) = \lim_{n\to\infty} \mathbb{E}(X_{r_n})$, that is, if and only if $t \mapsto \mathbb{E}(X_t)$ is right continuous. □

Corollary 3.15. *Let* $(\mathscr{F}_t)_{t\geq 0}$ *be a right continuous filtration on a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$*, and* $X = (X_t)_{t\geq 0}$ *be a martingale with respect to* $(\mathscr{F}_t)_{t\geq 0}$*. Then there is a right continuous and adapted modification of* $X = (X_t)_{t\geq 0}$.

Remark 3.16. *Let's make some comments on the assumptions. The assumption that the filtration is right continuous is necessary. For example, let* $\Omega = \{-1, +1\}$ *and* $\mathbb{P}[\{1\}] = \mathbb{P}[\{-1\}] = 1/2$*. We set*

$$X_t(\omega) = \begin{cases} 0, & 0 \leq t \leq 1, \\ \omega, & t > 1. \end{cases}$$

Then X *is a martingale with respect to its natural filtration (which is complete since there are no nonempty negligible sets), but no modification of* X *can be right continuous at* $t = 1$*. Similarly, take* $X_t = f(t)$*, where* $f(t)$ *is deterministic, non-increasing and not right continuous. Then no modification can have right continuous sample paths.*

3.2 Martingale convergence theorem

As another consequence of Doob's up-crossing lemma, we can establish a continuous parameter version of Doob's convergence theorem (as $t \to \infty$), whose proof is completely similar to that of the discrete parameter version we have done in Paper B8.1.

Theorem 3.17. *Let X be a super-martingale with right continuous sample paths. Assume that $(X_t)_{t\geq 0}$ is bounded in L^1, i.e.* $\sup_{t\geq 0}\mathbb{E}[|X_t|] < \infty$ *(or more equivalently* $\sup_{t\geq 0}\mathbb{E}[X_t^-] < \infty$*). Then there exists $X_\infty \in L^1$ such that* $\lim_{t\to\infty} X_t = X_\infty$ *almost surely.*

Under the assumptions of this theorem, X_t may not converge to X_∞ in L^1.

Exercise 3.18. *Let $X = (X_t)_{t\geq 0}$ be a non-negative super-martingale. Prove that $\{X_t : t \geq 0\}$ is bounded in L^1 and X_t converges almost surely to an integrable random variable X_∞ as $t \to \infty$. Prove that $X = (X_t)_{t\in[0,\infty]}$ is a super-martingale.*

The next result gives, for martingales, necessary and sufficient conditions for L^1-convergence.

Theorem 3.19 (Martingale Convergence Theorem)**.** *Let $X = (X_t)$ be a uniformly integrable super-martingale (resp. a martingale) with right continuous sample paths. Then there is an integrable random variable X_∞, such that*

1) $X_t \to X_\infty$ almost surely and in L^1 as $t \to \infty$,

2) X is closed *in the sense that $(X_t)_{t\in[0,\infty]}$ is a super-martingale (resp: a martingale):* $\mathbb{E}[X_t|\mathscr{F}_s] \leq X_s$ *(resp.* $\mathbb{E}[X_t|\mathscr{F}_s] = X_s$*) for all* $0 \leq s < t \leq \infty$ *. In particular* $\mathbb{E}[X_\infty|\mathscr{F}_t] \leq X_t$ *(resp.* $X_t = \mathbb{E}[X_\infty|\mathscr{F}_t]$*) for every* $t \geq 0$*.*

Proof. Since $\{X_t : t \geq 0\}$ is uniformly integrable, $\{X_t : t \geq 0\}$ is bounded in L^1, hence $X_t \to X_\infty$ almost surely as $t \to \infty$. Due to again the uniform integrability of $\{X_t : t \in [0,\infty]\}$, $\mathbb{E}[|X_t - X_\infty|] \to 0$ as $t \to \infty$ which proves 1). Since X is a super-martingale, $\mathbb{E}[X_t|\mathscr{F}_s] \leq X_s$ for any $t > s \geq 0$, by letting $t \to \infty$ to deduce that $\mathbb{E}[X_\infty|\mathscr{F}_s] \leq X_s$, which completes the proof of 2). □

Theorem 3.20. *Let $X = (X_t)_{t\geq 0}$ be a martingale (or a non-negative sub-martingale) with right continuous sample paths. Suppose* $\sup_{t\geq 0}\mathbb{E}[|X_t|^p] < \infty$ *for some constant* $p > 1$*. Then $(X_t)_{t\geq 0}$ is uniformly integrable, and $X_t \to X_\infty$ almost surely and in L^p as $t \to \infty$. Moreover* $\|X_\infty\|_p = \sup_{t\geq 0}\|X_t\|_p$*.*

Exercise 3.21. *Let $(\mathscr{F}_t)_{t\geq 0}$ be a right continuous filtration on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and ξ be a real random variable. Let $X = (X_t)_{t\geq 0}$ be the right continuous modification of the martingale $\mathbb{E}[\xi|\mathscr{F}_t]$, $t \geq 0$ (Corollary 3.15). Prove that $X_t \to \mathbb{E}[\xi|\mathscr{F}_\infty]$ almost surely and in L^1 as $t \to \infty$, where $\mathscr{F}_\infty = \sigma\{\mathscr{F}_t : t \geq 0\}$. [*Hint: Note that the martingale $\{\mathbb{E}[\xi|\mathscr{F}_t] : t \geq 0\}$ is uniformly integrable*].*

3.3 Stopping times

We now wish to establish a continuous parameter version of *Doob's optional stopping time theorem*. To this end we need to generalize a few notions about stopping times. The notion of stopping times can be generalized to the continuous setting.

Definition 3.22. *Let $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ be a filtered probability space. A random variable $T : \Omega \mapsto [0,\infty]$ is called a* stopping time *(relative to filtration $(\mathscr{F}_t)_{t\geq 0}$) if $\{T \leq t\} \in \mathscr{F}_t$ for every $t \geq 0$.*

Stopping times are sometimes called *optional*[17] or *Markov times*[18] (for example, in the *optional* stopping theorem).

Exercise 3.23. *A random variable $T : \Omega \mapsto [0, \infty]$ is an $\{\mathscr{F}_{t+} : t \geq 0\}$ stopping time if and only if $\{T < t\} \in \mathscr{F}_t$ for every $t > 0$.*

*[*Hint*: Use the following relations*

$$\{T < t\} = \bigcup_{n=1}^{\infty} \left\{ T \leq T - \frac{1}{n} \right\}$$

for any $t > 0$, and

$$\{T \leq t\} = \bigcap_{n=1}^{\infty} \left\{ T < t + \frac{1}{n} \right\}$$

for any $t \geq 0$].

With a stopping time we can associate the information known at time T.

Definition 3.24. *Given a stopping time T relative to $(\mathscr{F}_t)_{t \geq 0}$ we define σ-algebra at T*

$$\mathscr{F}_T := \{A \in \mathscr{F}_\infty : A \cap \{T \leq t\} \in \mathscr{F}_t \text{ for any } t \geq 0\}.$$

Applying the definition to the filtration $(\mathscr{F}_{t+})_{t \geq 0}$ to define

$$\mathscr{F}_{T+} := \{A \in \mathscr{F}_\infty : A \cap \{T \leq t\} \in \mathscr{F}_{t+} \text{ for any } t \geq 0\}.$$

One can verify that $\mathscr{F}_T$ (and therefore $\mathscr{F}_{T+}$ as well) is a σ-algebra [*Exercise*]. Let us state a few important facts about σ-algebras at stopping times.

Proposition 3.25. *Let T be a stopping time. Then*

1) T is $\mathscr{F}_T$-measurable.

2) $\mathscr{F}_T \subseteq \mathscr{F}_{T+}$.

3) If $T = t$ is a deterministic time, then $\mathscr{F}_T = \mathscr{F}_t$.

4) If T and S are stopping times then so are $T \wedge S$, $T \vee S$ and $T + S$, and $\{T \leq S\} \in \mathscr{F}_{T \wedge S}$. Further if $T \leq S$ then $\mathscr{F}_T \subseteq \mathscr{F}_S$.

5) If T is a stopping time and S is a $[0, \infty]$-valued random variable which is $\mathscr{F}_T$-measurable and $S \geq T$, then S is a stopping time. In particular

$$T^{(k)} := \sum_{l=1}^{\infty} \frac{l}{2^k} 1_{\{\frac{l-1}{2^k} \leq T < \frac{l}{2^k}\}} + \infty 1_{\{T = \infty\}} \tag{3}$$

($k = 1, 2, \ldots$) is a sequence of stopping times with $T^{(k)} \downarrow T$ as $k \to \infty$.

[17] For example, K. L. Chung: *Lectures from Markov Processes to Brownian Motion*, Springer (1982). This book has a new and merged version by K. L. Chung and J. B. Walsh: *Markov Processes, Brownian Motion, and Time Symmetry* (Second Edition), Springer (2005).

[18] For example, in K. Itô and H. P. McKean: *Diffusion Processes and their Sample Paths*. Springer-Verlag (1965, 1974). This is a classic written by the father of the subject, known today as Stochastic Analysis.

Proof. Note that $\{S \le t\} = \{S \le t\} \cap \{T \le t\} \in \mathscr{F}_t$ for every $t \ge 0$, since S is $\mathscr{F}_T$-measurable. Hence S is a stopping time. We have $T^{(k)} \downarrow T$ by definition, and clearly $T^{(k)}$ is $\mathscr{F}_T$-measurable since T is $\mathscr{F}_T$-measurable. Hence all $T^{(k)}$ are stopping times. □

Lemma 3.26. *Let S and T be two stopping times. Then*

1) If $S \le T$, then $\mathscr{F}_S \subset \mathscr{F}_T$;

2) $\mathscr{F}_{S\wedge T} = \mathscr{F}_S \cap \mathscr{F}_T$;

3) If $A \in \mathscr{F}_{S\vee T}$, then $A \cap \{S \le T\} \in \mathscr{F}_T$, $A \cap \{S < T\} \in \mathscr{F}_T$ and $A \cap \{S = T\} \in \mathscr{F}_{S\wedge T}$;

4) $\mathscr{F}_{S\vee T} = \sigma\{\mathscr{F}_S, \mathscr{F}_T\}$. Moreover, if $A \in \mathscr{F}_{S\vee T}$, then there are $A_1 \in \mathscr{F}_S$ and $A_2 \in \mathscr{F}_T$ such that $A = A_1 \cup A_2$ and $A_1 \cap A_2 = \emptyset$.

Proof. 1) Suppose $S \le T$, and $A \in \mathscr{F}_S$. Then

$$A \cap \{T \le t\} = (A \cap \{S \le t\}) \cap \{T \le t\} \in \mathscr{F}_t$$

for every $t \ge 0$, so by definition $A \in \mathscr{F}_T$, which proves that $\mathscr{F}_S \subset \mathscr{F}_T$.

2) We only need to show that $\mathscr{F}_S \cap \mathscr{F}_T \subset \mathscr{F}_{S\wedge T}$. Let $A \in \mathscr{F}_S \cap \mathscr{F}_T$. Then

$$A \cap \{S \wedge T \le t\} = (A \cap \{S \le t\}) \cup (A \cap \{T \le t\}) \in \mathscr{F}_t$$

for every $t \ge 0$, which implies that $A \in \mathscr{F}_{S\wedge T}$.

3) Since $A \in \mathscr{F}_{S\vee T}$ and S, T are $\mathscr{F}_{S\vee T}$-measurable, so that $A \cap \{S \le T\} \in \mathscr{F}_{S\vee T}$. It follows that

$$A \cap \{S \le T\} \cap \{T \le t\} = (A \cap \{S \le T\}) \cap \{S \vee T \le t\} \in \mathscr{F}_t$$

for every $t \ge 0$, which implies that $A \cap \{S \le T\} \in \mathscr{F}_T$. Similarly we have $A \cap \{S < T\} \in \mathscr{F}_T$. Therefore $A \cap \{S = T\} \in \mathscr{F}_T$ and by symmetry $A \cap \{S = T\} \in \mathscr{F}_S$. This proves 3).

4) We only need to prove the decomposition for every $A \in \mathscr{F}_{S\vee T}$, and the others will follow immediately. Indeed $A_1 = A \cap \{T < S\}$ and $A_2 = A \cap \{S \le T\}$ satisfy 4). □

Lemma 3.27. *Let S and T be two stopping times.*

1) $\{S \le T\}$, $\{S < T\}$ and $\{S = T\}$ belong to $\mathscr{F}_{S\wedge T}$.

2) If ξ be $\mathscr{F}_{S\vee T}$-measurable. Then $\xi 1_{\{S\le T\}}$, $\xi 1_{\{S<T\}}$ are $\mathscr{F}_T$-measurable and $\xi 1_{\{S=T\}}$ is $\mathscr{F}_{S\wedge T}$-measurable.

3) We have

$$\{S \le T\} \cap \mathscr{F}_{S\vee T} = \{S \le T\} \cap \mathscr{F}_T,$$
$$\{S < T\} \cap \mathscr{F}_{S\vee T} = \{S < T\} \cap \mathscr{F}_T$$

and

$$\{S = T\} \cap \mathscr{F}_{S\vee T} = \{S = T\} \cap \mathscr{F}_{S\wedge T} = \{S = T\} \cap \mathscr{F}_T = \{S = T\} \cap \mathscr{F}_S.$$

4) Similarly we have

$$\{S \le T\} \cap \mathscr{F}_{S\wedge T} = \{S \le T\} \cap \mathscr{F}_S$$

and

$$\{S < T\} \cap \mathscr{F}_{S\wedge T} = \{S < T\} \cap \mathscr{F}_S.$$

Proof. Let us prove 4). We only need to show that $\{S \le T\} \cap \mathscr{F}_S \subset \{S \le T\} \cap \mathscr{F}_{S\wedge T}$. Let $A \in \mathscr{F}_S$. Then $A \cap \{S \le T\} \in \mathscr{F}_S$. Since for every $t \ge 0$

$$A \cap \{S \le T\} \cap \{T \le t\} = (A \cap \{S \le t\}) \cap \{T \le t\} \cap \{S \wedge t \le T \wedge t\} \in \mathscr{F}_t$$

which yields that $A \cap \{S \le T\} \in \mathscr{F}_T$ as well. Therefore $A \cap \{S \le T\} \in \mathscr{F}_{S\wedge T}$. □

Exercise 3.28. *Suppose ξ is an integrable random variable, and S, T are two stopping times, then*

$$1_{\{S\le T\}}\mathbb{E}[\xi|\mathscr{F}_S] = 1_{\{S\le T\}}\mathbb{E}[\xi|\mathscr{F}_{S\wedge T}].$$

*[*Hint*: By the previous lemma $\{S \le T\} \in \mathscr{F}_{S\wedge T}$, and $\mathscr{F}_{S\wedge T} \subset \mathscr{F}_S$, so that*

$$\begin{aligned}\mathbf{1}_{\{S\le T\}}\mathbb{E}[X|\mathscr{F}_{S\wedge T}] &= \mathbb{E}\left[\mathbf{1}_{\{S\le T\}}\mathbb{E}[X|\mathscr{F}_S]|\mathscr{F}_{S\wedge T}\right] \\ &= \mathbf{1}_{\{S\le T\}}\mathbb{E}[X|\mathscr{F}_S]\end{aligned}$$

where the second equality follow from 4) in the previous lemma, $\mathbf{1}_{\{S\le T\}}\mathbb{E}[X|\mathscr{F}_S]$ is $\mathscr{F}_{S\wedge T}$-measurable].

3.4 Optional stopping time theorem

It is often useful to be able to 'stop' a process at a stopping time and know that the result still has nice measurability properties. To this end we need a definition. Recall that adaptedness tells us about measurability in ω at each time t, but nothing about regularity in time. Measurability of a process tells us about regularity in time and space, but not about adaptedness.

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\ge 0}, \mathbb{P})$ be a filtered probability space.

As we have pointed[19], an $\mathbb{R}^d$-valued process $X = (X_t)_{t\ge 0}$ give rise to a mapping $X : [0,\infty) \times \Omega \to \mathbb{R}^d$ which sends $(t, \omega) \in [0,\infty) \times \Omega$ to $X_t(\omega)$ (i.e. $X(t, \omega) = X_t(\omega)$). This consideration leads to several notions about measurability of a process.

First, we say X is a *measurable* process (also called Borel measurable[20]), if X is a measurable mapping from $([0,\infty) \times \Omega, \mathscr{B}([0,\infty)) \otimes \mathscr{F})$ to $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$, that is

[19] Recall that the three view-points when we say a process $X = (X_t)_{t\ge 0}$, first as a parameterized (ordered) family of random variables $\{X_t : t \ge 0\}$; second as a random function $X : \Omega \mapsto E^{[0,\infty)}$ taking values in the path space (where $E = \mathbb{R}^d$); thirdly as a mapping $X : [0,\infty) \times \Omega \mapsto \mathbb{R}^d$ sending (s, ω) to $X_s(\omega)$. The last point-view induces several concepts of measureabilities, where $[0,\infty) \times \Omega$ ie endowed the product σ-algebra $\mathscr{B}([0,\infty)) \otimes \mathscr{F}$, as long as $\mathscr{F}$ is a σ-algebra on the sample space Ω.

[20] cf. K. L. Chung: *Lectures from Markov processes to Brownian motion*.

for every Borel measurable $G \in \mathscr{B}(\mathbb{R}^d)$, $\{(t,\omega) : X_t(\omega)\} \in \mathscr{B}([0,\infty)) \otimes \mathscr{F}$. This measurability is more or less the least condition we shall assume when dealing with stochastic processes.

Second, we say an $\mathbb{R}^d$-valued process $X = (X_t)_{t\geq 0}$ is $(\mathscr{F}_t)_{t\geq 0}$ progressive (or *progressively measurable*) if, for each $t \geq 0$, the mapping X restricted on $[0,t] \times \Omega$, is a measurable mapping from $([0,t] \times \Omega, \mathscr{B}([0,t]) \otimes \mathscr{F}_t)$ to $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. By definition, an $(\mathscr{F}_t)$ progressively measurable process must be adapted to $(\mathscr{F}_t)_{t\geq 0}$.

If X is $(\mathscr{F}_t)_{t\geq 0}$ progressive, then it is $(\mathscr{F}_t)_{t\geq 0}$ adapted, but the converse is not necessarily true. One can show, with difficulty, that any adapted and measurable process has a progressive modification[21]. However, every right continuous $(\mathscr{F}_t)_{t\geq 0}$-adapted process is $(\mathscr{F}_t)_{t\geq 0}$-progressive and since we are interested in continuous processes, we will not need to dwell on these details.

Proposition 3.29. *An adapted process (X_t) whose paths are all right-continuous (or are all left-continuous) is* progressively measurable.

Proof. We present the argument for a right-continuous X. For $t > 0$, $n \geq 1$, $k = 0,1,2\ldots,2^n - 1$ let $X_0^{(n)}(\omega) = X_0(\omega)$ and $X_s^{(n)}(\omega) := X_{\frac{k+1}{2^n}t}(\omega)$ for $\frac{kt}{2^n} < s \leq \frac{k+1}{2^n}t$.

Clearly $(X_s^{(n)} : s \leq t)$ takes finitely many values and is $\mathscr{B}([0,t]) \otimes \mathscr{F}_t$–measurable. Further, by right continuity, $X_s(\omega) = \lim_{n\to\infty} X_s^{(n)}(\omega)$, and hence is also measurable (as a limit of measurable mappings). □

Theorem 3.30. *Let X be an $\mathbb{R}^d$-valued progressively measurable process and T a stopping time. Then $X_T 1_{T<\infty}$ is $\mathscr{F}_T$-measurable. The stopped process $X^T = (X_{T\wedge t})_{t\geq 0}$ is progressively measurable (where $X_T 1_{\{T<\infty\}}(\omega) = X_{T(\omega)}(\omega) 1_{\{T(\omega)<\infty\}}$).*

Proof. Since $X = (X_t)$ is progressively measurable, so that for every $t \geq 0$, the mapping $f_1 : (s,\omega) \mapsto X_s(\omega)$ is a measurable mapping from $([0,t] \times \Omega, \mathscr{B}([0,t]) \otimes \mathscr{F}_t)$ to $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. Since T is a stopping time, so that $f_2 : \omega \mapsto (T(\omega) \wedge t, \omega)$ is a measurable mapping from $(\Omega, \mathscr{F}_t)$ to $([0,t] \times \Omega, \mathscr{B}([0,t]) \otimes \mathscr{F}_t)$. Therefore the composition $f_1 \circ f_2 : \omega \mapsto X_{T(\omega)\wedge t}(\omega)$ is a measurable mapping from $(\Omega, \mathscr{F}_t)$ to $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. Therefore for every $A \in \mathscr{B}(\mathbb{R}^d)$

$$\left\{X_T 1_{\{T<\infty\}} \in A\right\} \cap \{T \leq t\} = \{X_T \in A, T \leq t\} = \{X_{T\wedge t} \in A, T \leq t\} \in \mathscr{F}_t$$

which implies, by definition, that $\left\{X_T 1_{\{T<\infty\}} \in A\right\} \in \mathscr{F}_T$ for every $A \in \mathscr{B}(\mathbb{R}^d)$. Hence $X_T 1_{\{T<\infty\}}$ is $\mathscr{F}_T$-measurable. □

Exercise 3.31. *Let $Z = (Z_t)_{t\geq 0}$ be a stochastic process such that $Z_t \to Z_\infty$ in probability. Then for every stopping time T, we define*

$$Z_T(\omega) = 1_{\{T(\omega)<\infty\}} Z_{T(\omega)}(\omega) + 1_{\{T(\omega)=\infty\}} Z_\infty(\omega)$$

[21] P. A. Meyer: *Probability and poentials*. Blaisdell Publishing Company, Waltham, Mass. (1966).

for every $\omega \in \Omega$. If $Z = (Z_t)_{t\geq 0}$ is progressively measurable, then Z_T is $\mathscr{F}_T$-measurable. The proof is exactly the same, as

$$\{Z_T \in A\} \cap \{T \leq t\} = \left\{Z_T 1_{\{T<\infty\}} \in A\right\} \cap \{T \leq t\}$$

for any $t \geq 0$.

The important examples of stopping times are those associated with adapted processes. That is, those stopping times that '*first time a certain phenomenon occurs*'. Our fundamental examples will be first hitting times of sets. If $X = (X_t)_{t\geq 0}$ is an $\mathbb{R}^d$-valued stochastic process and $\Gamma \in \mathscr{B}(\mathbb{R}^d)$, then

$$H_\Gamma(\omega) := \inf\{t \geq 0 : X_t(\omega) \in \Gamma\} \tag{4}$$

(with the convention that inf of empty set is ∞) is called the first entry time of X in Γ (or called the hitting time). Similarly

$$T_\Gamma(\omega) := \inf\{t > 0 : X_t(\omega) \in \Gamma\}$$

which is called the first hitting time of the process X on Γ.

Exercise 3.32. *Show that*

1) if X is adapted to $(\mathscr{F}_t)_{t\geq 0}$ and has right-continuous paths, then H_Γ, for Γ an open set, is a stopping time relative to $(\mathscr{F}_{t+})_{t\geq 0}$,

2) if X has continuous paths, then H_Γ, for Γ a closed set, is a stopping time relative to $(\mathscr{F}_t)_{t\geq 0}$.

One can show that the hitting time H_Γ of any Borel set Γ of a progressively measurable process $X = (X_t)_{t\geq 0}$ is a stopping time, assuming that $(\mathscr{F}_t)_{t\geq 0}$ satisfies the usual conditions, but this is surprisingly difficult!

We shall now establish conditions under which we have an optional stopping theorem for martingales in continuous time. As usual, our starting point will be the corresponding discrete time result and we shall pass to a suitable limit. Let us first recall the optional stopping time theorem (Paper B8.1).

Theorem 3.33 (Optional stopping for uniformly integrable discrete time martingales)**.** *Let $(Y_n)_{n\geq 0}$ be a uniformly integrable super-martingale (in discrete time) with respect to the filtration $(\mathscr{G}_n)_{n\geq 0}$, and Y_∞ be the a.s. limit of Y_n when $n \to \infty$. Let S, T be two stopping times S and T such that $S \leq T$. Then $Y_T \in L^1$ and*

$$Y_S \geq \mathbb{E}[Y_T | \mathscr{G}_S],$$

where

$$\mathscr{G}_S = \{A \in \mathscr{G}_\infty : A \cap \{S = n\} \in \mathscr{G}_n \text{ for every } n \geq 0\},$$

with the convention that $Y_T = Y_\infty$ on the event $\{T = \infty\}$, and similarly for Y_S.

Theorem 3.34. *(*Optional stopping time theorem*) Let* $(\mathscr{F}_t)_{t\geq 0}$ *be right continuous filtration, and* $X = (X_t)_{t\geq 0}$ *be a* uniformly integrable *super-martingale*[22] *(resp. martingale) with right continuous sample paths on* $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$*. Then* $X_t \to X_\infty$ *almost surely and in* L^1 *as* $t \to \infty$ *(*Doob's martingale convergence theorem*). Suppose* S, T *are two stopping times with* $S \leq T$*, then*

1) both X_S *and* X_T *are integrable,*

2) $\mathbb{E}[X_T|\mathscr{F}_S] \leq X_S$ *(resp.* $\mathbb{E}[X_T|\mathscr{F}_S] = X_S$*).*

Proof. Let us prove the theorem for super-martingale case. For each $k = 1, 2, \cdots$, let $D_k = \left\{ t_n^{(k)} = \frac{n}{2^k} : n = 0, 1, 2, \cdots \right\}$, and read X along the dyadic times $t_n^{(k)}$. That is, for each k, $X^{(k)} = (X_n^{(k)})_{n\geq 0}$ and $\mathscr{F}_n^{(k)} = \mathscr{F}_{t_n^{(k)}}$, where $X_n^{(k)} = X_{t_n^{(k)}}$. Then $X^{(k)} = (X_n^{(k)})_{n\geq 0}$ is a uniformly integrable super-martingale with respect to the filtration $(\mathscr{F}_n^{(k)})_{n\geq 1}$. We shall apply the optional stopping time theorem to each $X^{(k)}$, and let $k \to \infty$. To this end we define

$$S^{(k)} = \sum_{l=1}^{\infty} \frac{l}{2^k} 1_{\left\{\frac{l-1}{2^k} \leq S < \frac{l}{2^k}\right\}} + \infty 1_{\{S=\infty\}}$$

and

$$T^{(k)} = \sum_{l=1}^{\infty} \frac{l}{2^k} 1_{\left\{\frac{l-1}{2^k} \leq T < \frac{l}{2^k}\right\}} + \infty 1_{\{T=\infty\}}.$$

Then $S^{(k)} \downarrow S$ and $T^{(k)} \downarrow T$ as $k \uparrow \infty$, and clearly $S^{(k)} \leq T^{(k)}$ for each k. Since X is right continuous, we also have $X_{S^{(k)}} \to X_S$ and $X_{T^{(k)}} \to X_T$ almost surely when $k \to \infty$. Observe that $2^k S^{(k)}$ and $2^k T^{(k)}$ are $(\mathscr{F}_n^{(k)})$-stopping times and

$$X^{(k)}_{2^k S^{(k)}} = X_{S^{(k)}} \quad \text{and} \quad X^{(k)}_{2^k T^{(k)}} = X_{T^{(k)}}$$

and similarly

$$\mathscr{F}^{(k)}_{2^k S^{(k)}} = \mathscr{F}_{S^{(k)}} \quad \text{and} \quad \mathscr{F}^{(k)}_{2^k T^{(k)}} = \mathscr{F}_{T^{(k)}}$$

for each $k = 1, 2, \cdots$. Applying the (discrete parameter setting) optional stopping time theorem to $X^{(k)}$ with stopping times $2^k S^{(k)} \leq 2^k T^{(k)}$, both $X_{S^{(k)}}$ and $X_{T^{(k)}}$ are integrable and

$$\mathbb{E}[X_{T^{(k)}}|\mathscr{F}_{S^{(k)}}] \leq X_{S^{(k)}}.$$

In particular, for every $A \in \mathscr{F}_S \subset \mathscr{F}_{S^{(k)}}$ (for every k)

$$\mathbb{E}[X_{T^{(k)}} : A] \leq \mathbb{E}[X_{S^{(k)}} : A]. \tag{5}$$

In order to take limit under integration as $k \to \infty$, we show that $\{X_{S^{(k)}} : k = 1, 2, \cdots\}$ is uniformly integrable (and similarly $\{X_{T^{(k)}} : k = 1, 2, \cdots\}$ is uniformly integrable

[22] This theorem holds for right-continuous super-martingale, closed at right-hand side, $\{X_t : t \in [0, \infty]\}$, i.e. for a right-continuous super-martingale X such that $X_t \geq \mathbb{E}[X_\infty|\mathscr{F}_t]$ for some integrable random variable X_∞.

too). By definition $S^{(k)} \geq S^{(k+1)}$, so by applying the optional stopping time theorem to $X^{(k+1)}$ with $S^{(k)} \geq S^{(k+1)}$ we obtain that

$$\mathbb{E}\left[X_{S^{(k)}}|\mathscr{F}_{S^{(k+1)}}\right] \leq X_{S^{(k+1)}}$$

for $k = 1, 2, \cdots$. Let $Y_{-n} = X_{S^{(n)}}$ and $\mathscr{G}_{-n} = \mathscr{F}_{S^{(n)}}$ (where $n = 1, 2, \cdots$). Then (Y_{-n}) is a (backward) super-martingale with respect to $(\mathscr{G}_{-n})$, and $\mathbb{E}[Y_{-n}] = \mathbb{E}[X_{S^{(n)}}] \leq \mathbb{E}[X_0]$, hence (Y_{-n}) is uniformly integrable (see the appendix for a proof). Therefore both $(X_{S^{(k)}})_{k\geq 1}$ and $(X_{T^{(k)}})_{k\geq 1}$ are uniformly integrable. Therefore $X_{S^{(k)}} \to X_S$ and $X_{T^{(k)}} \to X_T$ in L^1. By letting $k \to \infty$ in (5) we deduce that

$$\mathbb{E}[X_T : A] \leq \mathbb{E}[X_S : A]$$

for every $A \in \mathscr{F}_S$. Under our assumptions, X_S is $\mathscr{F}_S$-measurable, we therefore have $\mathbb{E}[X_T|\mathscr{F}_S] \leq X_S$. □

Remark 3.35. *The proof of the martingale case can be simplified. Indeed, if $X = (X_t)_{t\geq 0}$ is a uniformly integrable martingale, then*

$$\mathbb{E}[X_{T^{(k)}}|\mathscr{F}_{S^{(k)}}] = X_{S^{(k)}}$$

and

$$\mathbb{E}[X_{T^{(k)}} : A] = \mathbb{E}[X_{S^{(k)}} : A] \tag{6}$$

for every $A \in \mathscr{F}_S \subset \mathscr{F}_{S^{(k)}}$ (for every k). Since we also have $X_{S^{(k)}} = \mathbb{E}[X_\infty|\mathscr{F}_{S^{(k)}}]$, so that $\{X_{S^{(k)}} : k \geq 1\}$ is uniformly integrable. So for this case we do not need to consider the backward martingale.

Corollary 3.36. *Suppose $X = (X_t)_{t\geq 0}$ is a martingale with right continuous paths with respect to a right continuous filtration $(\mathscr{F}_t)_{t\geq 0}$, and suppose S, T are two* bounded *stopping times with $S \leq T$, then X_S, $X_T \in L^1$ and $X_S = \mathbb{E}[X_T|\mathscr{F}_S]$.*

Proof. Let a be such that $S \leq T \leq a$. The martingale $(X_{t\wedge a})_{t\geq 0}$ is closed by X_a, hence $(X_{t\wedge a})_{t\geq 0}$ is uniformly integrable. We may apply our previous results to $X_{t\wedge a}$. □

Exercise 3.37. *Let $(\mathscr{F}_t)$ be right continuous, and $X = (X_t)_{t\geq 0}$ be a right continuous super-martingale on $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$. Then for every stopping time T*

$$\mathbb{E}\left[|X_T|1_{\{T<\infty\}}\right] \leq 3 \sup_{t\geq 0} \mathbb{E}[|X_t|].$$

*[*Hint. For every $a > 0$, the stopped super-martingale $X^a = (X_{t\wedge a})$ is a uniformly integrable, right continuous super-martingale with $X^a_\infty = X_a$, so that

$$\mathbb{E}\left[|X_{T\wedge a}|1_{\{T<\infty\}}\right] \leq \mathbb{E}[|X_{T\wedge a}|] = \mathbb{E}[X_{T\wedge a}] + 2\mathbb{E}\left[X^-_{T\wedge a}\right]$$

where $X_t^- = (-X_t) \vee 0$ the negative part of X_t for every $t \geq 0$. Then using optional stopping and the fact that $X^- = (X_t^-)_{t\geq 0}$ is a sub-martingale to prove the inequality.*]*

Corollary 3.38. *Suppose that $(X_t)_{t\geq 0}$ is a martingale with right continuous paths with respect to a right continuous filtration, and T is a stopping time.*

1) $X^T = (X_{t\wedge T})_{t\geq 0}$ is a martingale.

2) If in addition $\{X_t : t \geq 0\}$ is uniformly integrable, then $X^T = (X_{t\wedge T})_{t\geq 0}$ is uniformly integrable and for every $t \geq 0$, $X_{t\wedge T} = \mathbb{E}[X_T|\mathscr{F}_t]$.

Proof. We know $X_t^T = X_{t\wedge T} = X_T^t$, and that X_t is integrable. Hence, by the optional stopping theorem applied to the stopped process X^t, we see that X_t^T is integrable for every t. Furthermore, for any $s < t$, as $T \wedge s$ and $T \wedge t$ are bounded stopping times, by the optional stopping theorem

$$\begin{aligned} X_s^T &= X_{T\wedge s} = \mathbb{E}\left[X_{T\wedge t}|\mathscr{F}_{T\wedge s}\right] = 1_{\{T<s\}}X_T + 1_{\{T\geq s\}}\mathbb{E}\left[X_{T\wedge t}|\mathscr{F}_{T\wedge s}\right] \\ &= 1_{\{T<s\}}X_T + 1_{\{T\geq s\}}\mathbb{E}\left[X_{T\wedge t}|\mathscr{F}_s\right] = \mathbb{E}\left[X_{T\wedge t}|\mathscr{F}_s\right]. \end{aligned}$$

Therefore X^T is a martingale. □

Next is a general form of the optional stopping time theorem.

Theorem 3.39. *Let $(\mathscr{F}_t)_{t\geq 0}$ be right continuous. Suppose $X = (X_t)_{t\geq 0}$ is a uniformly integrable $(\mathscr{F}_t)$ super-martingale (resp. $(\mathscr{F}_t)$ martingale) with right continuous sample paths. Then $\mathbb{E}\left[X_T|\mathscr{F}_S\right] \leq X_{T\wedge S}$ (resp. $\mathbb{E}\left[X_T|\mathscr{F}_S\right] = X_{T\wedge S}$) for any stopping times S, T.*

Proof. Since $(\mathscr{F}_t)_{t\geq 0}$ is right continuous and the sample paths of X are right continuous, X_T is $\mathscr{F}_T$-measurable, and by Lemma, $X_T 1_{\{T\leq S\}}$ is $\mathscr{F}_S$-measurable, hence

$$\begin{aligned} \mathbb{E}\left[X_T|\mathscr{F}_S\right] &= \mathbb{E}\left[X_T 1_{\{T\leq S\}}|\mathscr{F}_S\right] + \mathbb{E}\left[X_{T\vee S} 1_{\{T>S\}}|\mathscr{F}_S\right] \\ &\leq X_T 1_{\{T\leq S\}} + X_S 1_{\{T>S\}} = X_{S\wedge T} \end{aligned}$$

which completes the proof. □

A converse result is also possible.

Theorem 3.40. *Suppose M is a right continuous process defined for $t < \infty$, and adapted to a right continuous filtration $(\mathscr{F}_t)_{t\geq 0}$, then M is a martingale if, and only if $|M_T|$ is integrable and $\mathbb{E}\left[M_T\right] = \mathbb{E}\left[M_0\right]$ for every bounded stopping time T.*

Proof. By considering the process $(M_t - M_0)_{t\geq 0}$, we can assume without loss of generality that $M_0 = 0$. If M is a martingale, then M_T is integrable $M_0 = \mathbb{E}\left[M_T|\mathscr{F}_0\right]$ for any bounded stopping time (by the previous optional stopping time theorem). Conversely, for $t > s \geq 0$ and $A \in \mathscr{F}_s$. Let $T = s1_A + t1_{A^c}$. Then T is a bounded stopping time. Hence

$$0 = \mathbb{E}\left[M_T\right] = \mathbb{E}\left[1_A M_s\right] + \mathbb{E}\left[1_{A^c} M_t\right]$$

and

$$0 = \mathbb{E}\left[M_t\right] = \mathbb{E}\left[1_A M_t\right] + \mathbb{E}\left[1_{A^c} M_t\right]$$

which yields that

$$\mathbb{E}\left[1_A M_s\right] = \mathbb{E}\left[1_A M_t\right]$$

for every $A \in \mathscr{F}_s$. Therefore M is a martingale. □

3.5 Local (super-, sub-) martingales

In other areas of analysis, we often prove results by showing that a statement holds 'locally'. This is made difficult for stochastic processes by the fact that these depend on the random sample point $\omega \in \Omega$, for which we have no topology. One useful way around this challenge is to use stopping times, which helpfully 'localize' in both time and space.

We say a sequence of stopping times, T_n ($n = 1, 2, \cdots$) a localizing sequence if $T_{n+1} \geq T_n$ for every $n = 1, 2, \cdots$, and $T_n \to \infty$ as $n \to \infty$.

Definition 3.41. *We say that a process X* locally *has some property* C *if there exists a sequence of stopping times T_n (where $n = 1, 2, \ldots$) such that the stopped processes X^{T_n} have property* C *for every n, and $T_n \uparrow \infty$ almost surely. The sequence T_n is said to* localize *or* reduce *X.*

When we work with processes locally, it is then useful to reconstruct the 'global' process from its local versions.

Exercise 3.42. *Suppose given a sequence of stopping times $T_n \uparrow \infty$ and a sequence of processes $\{Y^{(n)} : n = 1, 2, \ldots\}$ such that, for all $n \leq m$,*

$$Y^{(n)} = Y^{(m)} \quad \text{on } \{t \leq T_n\}$$

*up to indistinguishability, there exists a process Y such that $Y = Y^{(n)}$ on $\{t \leq T_n\}$ for all n. [*Hint: Let $T_0 = 0$. Check that $Y := 1_{t=0} Y_0^1 + \sum_{n=0}^{\infty} \mathbf{1}_{T_n < t \leq T_{n+1}} Y_t^{n+1}$ do the job.*]*

We say $X = (X_t)_{t \geq 0}$ is a *local super-martingale*, if there is an increasing sequence of stopping times T_n such that $T_n \uparrow \infty$ and $X^{T_n} = (X_{t \wedge T_n})_{t \geq 0}$ is a super-martingale for every $n = 1, 2, \ldots$. We define the notions of local sub-martingale, local martingales in a similar way. Any martingale is a local martingale, but the converse is false.

Proposition 3.43. *Let $M = (M_t)_{t \geq 0}$ be a continuous local martingale.*

1) If M is non-negative and $M_0 \in L^1$, M is a super-martingale.

2) Suppose there exists a random variable $Z \in L^1$ with $|M_t| \leq Z$ for every $t \geq 0$, then M is a uniformly integrable martingale.

3) Let $T_n = \inf\{t \geq 0 : |M_t| \geq n\}$ (for $n = 1, 2, \cdots$). Then $T_n \uparrow \infty$ and M^{T_n} is a bounded martingale for every n.

4) If T is a stopping time, then M^T is also a local martingale.

Proof. 1) By definition, there exists a sequence T_n of stopping times such that $T_n \uparrow \infty$ and M^{T_n} is a martingale for every n. Then $M_{s \wedge T_n} = \mathbb{E}[M_{t \wedge T_n} | \mathscr{F}_s]$ for any $t > s \geq 0$ and for every n. By Fatou's lemma we have

$$\mathbb{E}[M_t | \mathscr{F}_s] \leq \liminf_{n \to \infty} \mathbb{E}[M_{t \wedge T_n} | \mathscr{F}_s] = M_s$$

for every $t > s \geq 0$. In particular, $\mathbb{E}[M_t] \leq \mathbb{E}[M_0] < \infty$, so that M_t is integrable for all $t \geq 0$. Therefore $M = (M_t)_{t \geq 0}$ is a super-martingale.

2) Use Dominated Convergence Theorem.

3) Since M_t are real random variable and $t \mapsto M_t$ is continuous so that $|M_t| < \infty$ for every t almost surely, therefore $T_n \uparrow \infty$. For each n, $|M_{T_n \wedge t}| \leq n$. On the other hand, by assumption, there is a sequence of stopping times $S_m \uparrow \infty$ ($m = 1, 2, \ldots$) such that M^{S_m} is a martingale for every m. Hence $\mathbb{E}[M_{S_m \wedge t} | \mathscr{F}_s] = M_{S_m \wedge s}$ for each m. By optional stopping to obtain

$$\mathbb{E}[M_{S_m \wedge T_n \wedge t} | \mathscr{F}_s] = M_{S_m \wedge T_n \wedge s}.$$

Letting $m \to \infty$, using Dominated Convergence Theorem, we obtain $\mathbb{E}[M_{T_n \wedge t} | \mathscr{F}_s] = M_{T_n \wedge s}$. That is M^{T_n} is a martingale for each n. □

Remark 3.44. *In the definition of a local sub-, super- or local martingale, $M_0^{T_n} = M_{T_n \wedge 0} = M_0$ should be integrable, which forces M_0 is integrable. Some authors prefer a constant random variable is a local martingale, so that a local martingale to them means $1_{\{T_n > 0\}} M_{T_n \wedge t}$ (for $t \geq 0$) is a martingale. We do not use this convention.*

3.6 Square integrable martingales

In this section we establish the major tool, namely the quadratic variation process associated with a continuous local martingale, for establishing Itô's theory of stochastic integration[23].

3.6.1 Functions with finite p-variation

There are two ways, but equivalent, to define the various variations of a function $\rho : [0, \infty) \mapsto \mathbb{R}$, which we shall give a quick review. These notions and notations later on shall be applied to sample paths of stochastic processes.

First some notations. Let $\mathscr{P}$ denote the collection of all partitions $D : 0 = t_0 < t_1 < \cdots < t_i < \cdots$ such that $t_i \to \infty$ as $i \to \infty$. An element of $\mathscr{P}$can be identified with the collection of strictly increasing sequence $\{t_i : i = 0, 1, \ldots\}$ such that $t_0 = 0$ and $t_i \to \infty$. We simply denote a typical element $D = \{t_i : i = 0, 1, \ldots\}$. The mesh of D is defined as $|D| = \sup_{i \geq 1}(t_i - t_{i-1})$ (which can be ∞).

Similarly, for any $b > a \geq 0$, we shall use $\mathscr{P}([a, b])$ to denote all finite partitions $D : \{a = t_0 < t_1 < \cdots < t_m = b\}$ (for some $m = m(D) \in \mathbb{N}$), and similarly the mesh $|D| = \max_i(t_i - t_{i-1})$.

Let $\rho : [a, b] \mapsto \mathbb{R}$ be a function, and $p \geq 1$ be a number. The p-variation of ρ (over bounded interval $[a, b]$) with respect to a finite partition $D : \{a = t_0 < t_1 <$

[23]Itô's integrals were first defined against Brownian motion in K. Itô: Stochastic integral, *Proc. Imp. Acad. Tokyo,* 20 (1944), 519-524. It was recognized that the key to extend Itô's theory to other martingales is the existence of the quadratic variation process of a martingale, in H. Kunita and S. Watanabe: On square integrable martingales, *Nagoya Math. J.,* 30 (1967), 209-245.

$\cdots < t_m = b\}$ is defined by

$$V^{\{D,p\}}(\rho|_{[a,b]}) = \sum_{i=1}^{m(D)} |\rho(t_i) - \rho(t_{i-1})|^p.$$

We say that ρ has finite p-variation over $[a,b]$ if

$$V^{\{p\}}(\rho|_{[a,b]}) = \lim_{D \in \mathscr{P}([a,b]), |D| \to 0} V^{\{D,p\}}(\rho|_{[a,b]})$$

exists, in the sense that for every $\varepsilon > 0$, there is $\delta > 0$ such that

$$\left| V^{\{D,p\}}(\rho|_{[a,b]}) - V^{\{p\}}(\rho|_{[a,b]}) \right| < \varepsilon$$

for any finite partition D of $[a,b]$ with $|D| < \delta$.

If $\rho : [0,\infty) \mapsto \mathbb{R}$ is a real valued function (i.e. a path in $\mathbb{R}$), then we define $V_t^{\{D,p\}}(\rho) = V^{\{D,p\}}(\rho|_{[0,t]})$ for $t \geq 0$. We say ρ is of finite p-variation if $V_t^{\{p\}}(\rho) = V^{\{p\}}(\rho|_{[0,t]})$ exists for every $t \geq 0$. We say a function $\rho : [0,\infty) \mapsto \mathbb{R}$ is of finite variation, if it is of finite p-variation with $p = 1$. We shall use $V_t(\rho)$ to denote $V_t^{\{1\}}(\rho)$ for simplicity. $V(\rho) : t \mapsto V_t(\rho)$ is called the total variation (function).

Since we shall deal with processes of finite p-variation (for us there are two interesting cases: $p = 1$ and $p = 2$), and their associated p-variation processes, it will be useful to use partitions independent of t, i.e. those partitions in $\mathscr{P}$. Namely, if $D = \{t_i : i = 1,2,\ldots\} \in \mathscr{P}$ so that $0 = t_0 < t_1 < \cdots < t_i < \cdots$ with $t_i \to \infty$ when $i \to \infty$, then define

$$T_t^{\{D,p\}}(\rho) = \sum_{i=0}^{\infty} |\rho(t_{i+1} \wedge t) - \rho(t_i \wedge t)|^p \quad \text{for } t \geq 0.$$

The right-hand side is indeed a finite sum since $t_i \to \infty$ when $i \to \infty$.

Lemma 3.45. *Let $p \geq 1$ and $\rho : [0,\infty) \mapsto \mathbb{R}$ be a path in $\mathbb{R}$.*

1) ρ is of finite p-variation if and only if for every $t > 0$, $\lim_{D \in \mathscr{P}, |D| \to 0} T_t^{\{D,p\}}(\rho) = T_t^{\{p\}}(\rho)$ exists, in the sense that for every $\varepsilon > 0$, there is $\delta > 0$ such that

$$\left| T_t^{\{D,p\}}(\rho) - T_t^{\{p\}}(\rho) \right| < \varepsilon$$

for every $D \in \mathscr{P}$ with $|D| < \delta$. In this case $V_t^{\{p\}}(\rho) = T_t^{\{p\}}(\rho)$ for every $t > 0$.

2) Suppose ρ has finite p-variation, and if $D^{(n)} = \{t_i^{(n)} : i = 1,2,\cdots\}$ is a sequence in $\mathscr{P}$, where $n = 1,2,\ldots$, such that $|D^{(n)}|$ converges to 0 when $n \to \infty$, then $\lim_{n\to\infty} T_t^{\{D^{(n)},p\}}(\rho) = V_t^{\{p\}}(\rho)$ for every $t > 0$.

3) If ρ is a continuous path with finite p-variation, then $t \mapsto V_t^{\{p\}}(\rho)$ is continuous, non-decreasing, with initial value zero.

Remark 3.46. *If ρ is of finite variation, then $t \mapsto V_t(\rho)$ is non-decreasing with initial value zero. If in addition ρ is continuous, then $t \mapsto V_t(\rho)$ is continuous too.*

Next we assume that ρ is of finite variation. Let $\rho^{(\pm)}(t) = \frac{1}{2}(V_t(\rho) \pm \rho(t))$ for $t \geq 0$. Then both $\rho^{(+)}$ and $\rho^{(-)}$ are non-decreasing on $[0,\infty)$, and $\rho(t) = \rho^{(+)}(t) - \rho^{(-)}(t)$ for any $t \geq 0$. Moreover if ρ is continuous, then both $\rho^{(+)}$ and $\rho^{(-)}$ are continuous too.

Exercise 3.47. *If $\rho(t) = \rho_1(t) - \rho_2(t)$ for any $t \geq 0$, and ρ_1, ρ_2 are non-decreasing, then ρ has finite variations (over finite intervals).*

Therefore in order to definite integral like $\int f(t)d\rho(t)$ against ρ which is of finite variation, we only need to do define integrals against ρ which is non-decreasing. This is however done in Paper A4 via Lebesgue's theory of integration.

For the propose of applications later on, let us consider a function $\rho : [0,\infty) \mapsto \mathbb{R}$ which is non-decreasing and right continuous on $[0,\infty)$, with $\rho(0) = 0$. According to the standard theory established in Lebesgue's integration, there is a unique measure[24], denoted by m_ρ, on $(\mathbb{R}_+, \mathscr{B}(\mathbb{R}_+))$ such that

$$m_\rho((a,b]) = \rho(b) - \rho(a)$$

for $b > a \geq 0$, and $m_\rho(\{0\}) = 0$.

As a consequence, if f is a function, such that $1_{(0,t]}f$ is integrable with respect to the measure m_ρ, then the integral

$$(f \cdot \rho)(t) = \int_0^t f(s)d\rho(s) := \int 1_{(0,t]}(s)f(s)m_\rho(ds)$$

where the first two are the notations we shall use and the last one is the Lebesgue integral defined in Paper A4. If ρ is of finite variation, then $\rho = \rho^{(+)} - \rho^{(-)}$, where $\rho^{(+)}$ and $\rho^{(-)}$ are non-decreasing, we naturally define the integral

$$\int_0^t f(s)d\rho(s) = (f \cdot \rho^{(+)})(t) - (f \cdot \rho^{(-)})(t)$$

as long as $1_{(0,t]}f$ is integrable with respect to both measures $m_{\rho^{(+)}}$ and $m_{\rho^{(-)}}$. Finally we shall use $\int_0^t f(s)|d\rho(s)|$ to denote the integral of $f1_{(0,t]}$ against the measure $m_{V(\rho)}$, the Lebesgue-Stieltjes measure associated with the total variation function $V(\rho)$.

Remark 3.48. *A function $\rho \in C^1$ (i.e. differentiable with continuous derivative) is of finite variation and $\int f(s)d\rho(s) = \int f(s)\rho'(s)ds$.*

[24] This is the Lebesgue-Stieltjes measure associated with an non-decreasing function ρ, constructed in Paper A4.

If ρ is non-decreasing and right-continuous (or with finite variation over any finite interval), and if f is left continuous, then

$$\int_0^t f(s)d\rho(s) = \lim_{D\in\mathscr{P}([0,t]),|D|\to 0} \sum_l f(t_{l-1})\left(\rho(t_l)-\rho(t_{l-1})\right)$$

where the limit is taken over finite partitions D of $[0,t]$: $0=t_0<t_1<\cdots<t_m=t$.

Exercise 3.49 (Associativity)**.** *Let ρ be right continuous with finite variation, f be Borel measurable functions, and f be integrable with respect to $m_{\rho_\pm}$. Prove that $f\cdot\rho$ is of finite variation.*

Suppose g is integrable is integrable with respect to $m_{(f\cdot\rho)_\pm}$. Prove that gf is integrable with respect to $m_{\rho_\pm}$ and

$$\int_0^t g(s)d(f\cdot\rho)(s) = \int_0^t g(s)f(s)d\rho(s).$$

Exercise 3.50 (Stopping)**.** *Let ρ be right continuous withfinite variation and fix $t\geq 0$. Set $\rho^t(s)=\rho(t\wedge s)$. Then ρ^t is of finite variation and for any measurable ρ-integrable function f*

$$\int_0^{u\wedge t} f(s)d\rho(s) = \int_0^u f(s)d\rho^t(s) = \int_0^u f(s)\mathbf{1}_{[0,t]}(s)d\rho(s)$$

for every $u\geq 0$.

Exercise 3.51 (Chain-rule)**.** *If F is a C^1 function and ρ is continuouswith finite variation, then $F(\rho(t))$ is also of finite variation and*

$$F(\rho(t)) = F(\rho(0)) + \int_0^t F'(\rho(s))d\rho(s).$$

*[*Hint. The statement is trivially true for $F(x)=x$. It is straightforward to check that if the statement is true for F, then it is also true for $xF(x)$. Hence, by induction, the statement holds for all polynomials. To complete the proof, approximate $F\in C^1$ by a sequence of polynomials.*]*

Proposition 3.52 (Change of variables)**.** *If ρ is non-decreasing and right-continuous then so is its 'right inverse'*

$$\rho^{-1}(s) := \inf\{t\geq 0 : \rho(t)>s\},$$

where $\inf\emptyset=+\infty$. *Let $\rho(0)=0$. Then, for any Borel measurable function $f\geq 0$ on $\mathbb{R}_+$, we have*

$$\int_0^\infty f(u)d\rho(u) = \int_0^{\rho(\infty)} f(\rho^{-1}(s))ds.$$

Proof. If $f(u) = 1_{[0,v]}(u)$, then the claim becomes

$$\rho(v) = \int_0^\infty 1_{\{c(s) \le v\}} ds = \inf\{s : \rho^{-1}(s) > v\},$$

and equality holds by definition of ρ^{-1}. Take differences to get indicators of sets $(u, v]$. The Monotone Class Theorem allows us to extend to functions of compact support and then take increasing limits to obtain the formula in general. □

3.6.2 Martingale spaces

Now we shall apply the theory above to the study of square integral martingales.

Theorem 3.53. *A* continuous *local martingale $M = (M_t)_{t\ge 0}$ with $M_0 = 0$ a.s., is a process of finite variation*[25] *if and only if M is indistinguishable from zero. [*The continuity assumption is critical here.*]*

Proof. Suppose M is a continuous local martingale and of finite variation. Let

$$T_n = \inf\left\{t \ge 0 : \int_0^t |dM_s| \ge n\right\} = \inf\{t \ge 0 : V_t(M) \ge n\},$$

which are stopping times, since $V_t(M) = \int_0^t |dM_s|$ is continuous and adapted. Since M has finite variation on $[0,t]$ for any $t \ge 0$, so that $T_n \uparrow \infty$. By definition

$$|M_{t\wedge T_n}| \le \left|\int_0^{t\wedge T_n} dM_s\right| \le \int_0^{t\wedge T_n} |dM_u| \le n$$

which implies that $N := M^{T_n}$ is a bounded martingale. Let $t > 0$ and $D : 0 = t_0 < t_1 < t_2 < \ldots < t_m = t$ be any finite partition of $[0,t]$. Then

$$\begin{aligned}
\mathbb{E}[N_t^2] &= \sum_{i=1}^m \mathbb{E}[N_{t_i}^2 - N_{t_{i-1}}^2] = \sum_{i=1}^m \mathbb{E}\big[(N_{t_i} - N_{t_{i-1}})^2\big] \\
&\le \mathbb{E}\Big[(\sup_i |N_{t_i} - N_{t_{i-1}}|) \cdot \underbrace{\sum |N_{t_i} - N_{t_{i-1}}|}_{\le V_t(N) = V_{t\wedge T_n}(M) \le n}\Big] \\
&\le n\mathbb{E}\big[\sup_i |N_{t_i} - N_{t_{i-1}}|\big]
\end{aligned}$$

for any partition D. Since N is continuous, so it is uniformly continuous on $[0,t]$ (for each $\omega \in \Omega$), so that $\sup_i |N_{t_i} - N_{t_{i-1}}| \to 0$ as $|D| \to 0$. Moreover, $\sup_i |N_{t_i} - N_{t_{i-1}}| \le V_t(N) \le n$ for every D, thus by Dominated Convergence Theorem, $\mathbb{E}\big[\sup_i |N_{t_i} - N_{t_{i-1}}|\big] \to 0$ as $|D| \to 0$. Hence $\mathbb{E}[N_t^2] = 0$, so that $M_{T_n \wedge t} = 0$ for every n and $t > 0$. Therefore $M = 0$. □

[25] Naturally, a (real valued) stochastic process $X = (X_t)_{t\ge 0}$ is a process of finite variation, if for almost all $\omega \in \Omega$, sample paths $t \mapsto X_t(\omega)$ are of finite variation.

We shall use $\mathscr{M}_2$ to denote the vector space of all *right continuous*, square integrable martingales [26] on $(\Omega, \mathscr{F}, \mathbb{P})$ with respect to $(\mathscr{F}_t)_{t\geq 0}$ and initial value $M_0 = 0$. $\mathscr{M}_2^c$ denotes those martingales in $\mathscr{M}_2$ which have continuous sample paths. $\mathscr{M}_2^c$ is a subspace of $\mathscr{M}_2$.

Exercise 3.54. *Let $M = (M_t)_{t\geq 0} \in \mathscr{M}_2$.*

1) Prove that $(M_t^2)_{t\geq 0}$ is a sub-martingale and

$$\mathbb{E}\left(M_t^2 - M_s^2|\mathscr{F}_s\right) = \mathbb{E}\left[(M_t - M_s)^2|\mathscr{F}_s\right]$$

for any $t > s \geq 0$. In particular $t \mapsto \mathbb{E}\left[M_t^2\right]$ is non-decreasing.

Hint: use the equality

$$\mathbb{E}\left[(M_t - M_s)^2|\mathscr{F}_s\right] = \mathbb{E}\left[(M_t^2 - M_s^2 - 2M_s(M_t - M_s)|\mathscr{F}_s\right].$$

2) More general, show the following equality holds:

$$\mathbb{E}\left[(M_t - M_s)^2|\mathscr{F}_u\right] = \begin{cases} \mathbb{E}\left(M_t^2 - M_s^2|\mathscr{F}_u\right), & \text{if } t > s \geq u \geq 0, \\ \mathbb{E}\left(M_t^2 - M_u^2|\mathscr{F}_u\right) + (M_u - M_s)^2, & \text{if } t > u > s \geq 0. \end{cases}$$

If $M \in \mathscr{M}_2$ and $T > 0$ be any fixed time, then $\|M\|_{2,T} = \sqrt{\mathbb{E}\left[|M_T|^2\right]}$ and

$$\|M\|_2 = \sqrt{\sum_{n=1}^{\infty} \frac{1}{2^n}\left(1 \wedge \|M\|_{2,T}^2\right)}. \tag{7}$$

Exercise 3.55. *Prove that $(M,N) \mapsto \|M - N\|_2$ defined in (7) is a metric on $\mathscr{M}_2$.*

Lemma 3.56. *$\mathscr{M}_2$ is a complete metric space under the metric (7), and $\mathscr{M}_2^c$ is a closed subspace of $\mathscr{M}_2$.*

Proof. Suppose $\{M^{(n)} : n = 1, 2, \ldots\}$ is a Cauchy sequence in $\mathscr{M}_2$. By Kolmogorove-Doob's inequality[27]

$$\mathbb{P}\left[\sup_{0\leq t\leq T}\left|M_t^{(n)} - M_t^{(m)}\right| \geq \lambda\right] \leq \frac{1}{\lambda^2}\mathbb{E}\left[|M_T^{(n)} - M_T^{(m)}|^2\right]$$

for any $T > 0$ and $\lambda > 0$, which tends to zero as $n, m \to \infty$. Therefore there is a process $M = (M_t)_{t\geq 0}$ such that

$$\sup_{0\leq t\leq T}\left|M_t^{(n)} - M_t\right| \to 0 \quad \text{in probability}$$

as $n \to \infty$, for every $T > 0$. Hence there is a sub-sequence $M^{(n_k)}$ such that $\sup_{0\leq t\leq T}\left|M_t^{(n_k)} - M_t\right| \to 0$ almost surely. Hence there is a null set A_T, such that $\sup_{0\leq t\leq T}\left|M_t^{(n_k)}(\omega) - M_t(\omega)\right| \to$

[26] A martingale $M = (M_t)_{t\geq 0}$ is square integrable if $\mathbb{E}\left[M_t^2\right] < \infty$ for all $t \geq 0$. This is different from a martingale which is bounded in L^2-space.

[27] This is the name for Doob's maximal inequality with $p = 2$, which was proved by Kolmogorov for random walks. Here we apply the maximal inequality to martingale $M^{(n)} - M^{(m)}$.

0 for every $\omega \in \Omega \setminus A_T$, that is $M_t^{(n_k)}(\omega) \to M_t(\omega)$ uniformly on $[0,T]$ as $k \to \infty$ for every $\omega \in \Omega \setminus A_T$, and therefore $M = (M_t)_{t \geq 0}$ is right continuous almost surely on $[0,T]$. If $M^{(n)}$ are continuous, so is M on $[0,T]$ for every $T > 0$. Since $[0,\infty) = \bigcup_{n=1}^{\infty}[0,n]$, M is right continuous almost surely on $[0,\infty)$. Since for every $t \geq 0$, $\mathbb{E}\left[\left|M_t^{(n)} - M_t\right|^2\right] \to 0$ as $n \to \infty$, $M \in \mathscr{M}_2$. $\square$

Definition 3.57. *Let $\mathscr{H}^2$ be the space of right continuous square integrable martingales bounded in L^2, that is, the vector space of all right continuous martingales $M = (M_t)_{t \geq 0}$ on $(\Omega, \mathscr{F}, \mathbb{P})$ with respect to $(\mathscr{F}_t)_{t \geq 0}$ such that $\sup_{t \geq 0} \mathbb{E}\left[M_t^2\right] < \infty$[28]. $\mathscr{H}^{2,c}$ denotes those $M \in \mathscr{H}^2$ which are continuous, and $\mathscr{H}_0^{2,c}$ the subspace of all those M with initial $M_0 = 0$.*

We note that the space $\mathscr{H}^2$ is also sometimes denoted $\mathscr{M}^2$. If $M = (M_t)_{t \geq 0}$ is an L^2-bounded right-continuous martingale, then $\{M_t : t \geq 0\}$ is uniformly integrable, $M_t \mapsto \mathbb{E}\left[M_t^2\right]$ is non-decreasing, and $M_t \to M_\infty$ almost surely and in L^2 as $t \to \infty$. Moreover

$$\sup_{t \geq 0} \mathbb{E}\left[M_t^2\right] = \lim_{t \to \infty} \mathbb{E}\left[M_t^2\right] = \mathbb{E}\left[M_\infty^2\right]$$

and $M_t = \mathbb{E}\left[M_\infty | \mathscr{F}_t\right]$ for $t \geq 0$. Therefore there is a one-to-one correspondence between $\mathscr{H}^2$ and $L^2(\Omega, \mathscr{F}_\infty, \mathbb{P})$, so that $\mathscr{H}^2$ is a Hilbert space[29] under the following norm

$$\|M\|_{\mathscr{H}^2} = \sqrt{\mathbb{E}\left[M_\infty^2\right]} = \sqrt{\sup_{t \geq 0} \mathbb{E}\left[M_t^2\right]}.$$

Doob's L^2 inequality implies that

$$\|M\|_{\mathscr{H}^2}^2 \leq \mathbb{E}\left[\sup_{t \geq 0} M_t^2\right] \leq 4\|M\|_{\mathscr{H}^2}^2 \qquad \text{for } M \in \mathscr{H}^2.$$

Exercise 3.58. *$\mathscr{H}^{2,c}$ is a closed subspace of $\mathscr{H}^2$. [*Hint: use Doob's L^2-inequality as in the proof of Lemma *3.56]*

3.6.3 Quadratic variation processes

Itô made use of two facts to define his stochastic integrals for Brownian motion, namely, 1) a Brownian motion $B = (B_t)_{t \geq 0}$ is a continuous martingale, and 2) $\{B_t^2 - t : t \geq 0\}$ is a continuous martingale too. In order to generalize Itô's theory of stochastic integration for Brownian motion to a continuous local martingale $M = (M_t)_{t \geq 0}$, we shall construct a continuous, non-decreasing and adapted process $\langle M, M \rangle$, called the *quadratic process* associated with M, such that $M_t^2 - \langle M, M \rangle_t$ (for $t \geq 0$) is a martingale.

[28] Since $\{M_t^2; t \geq 0\}$ is a sub-martingale, so that $t \mapsto \mathbb{E}\left[M_t^2\right]$ is non-decreasing, and therefore $\sup_{t \geq 0} \mathbb{E}\left[M_t^2\right] = \lim_{t \to \infty} \mathbb{E}\left[M_t^2\right]$.

[29] A complete normed space whose norm $\|M\|_{\mathscr{H}^2}$ is defined via an inner product $(M,N) = \mathbb{E}\left[M_\infty N_\infty\right]$.

To this end we shall first introduce some notation. If $D=\{t_i\}\in\mathscr{P}$ is a partition of $[0,\infty)$: $0=t_0<t_1<\cdots<t_i<\cdots$ such that $t_i\to\infty$, and if $X=(X_t)_{t\geq 0}$, then define $T^D(X)$ to be a process defined by

$$T_t^D(X)=\sum_{i=0}^{\infty}\left(X_{t_{i+1}\wedge t}-X_{t_i\wedge t}\right)^2 \quad \text{for } t\geq 0.$$

If X is adapted, so is $T^D(X)$, and if X is continuous (right continuous), so is $T^D(X)$[30], for every partition $D\in\mathscr{P}$.

In what follows, we shall work with a complete probability space $(\Omega,\mathscr{F},\mathbb{P})$ and a filtration $(\mathscr{F}_t)_{t\geq 0}$ which satisfies the *usual conditions*. In particular, $\mathscr{F}_{t+}=\mathscr{F}_t$ for every $t\geq 0$.

Let $M=(M_t)_{t\geq 0}$ be a continuous local martingale. By using reducing stopping times $T_n=\inf\{t\geq 0:|M_t|\geq n\}$, M^{T_n} is a bounded and continuous martingale for each n. Therefore we shall first construct quadratic processes for continuous bounded martingales.

Lemma 3.59. *Suppose $M=(M_t)_{t\geq 0}$ is a square integrable martingale, and $D:0=t_0<t_1<\cdots<t_i<\cdots$ is a partition of $[0,\infty)$ such that $t_i\to\infty$ when $i\to\infty$, then $\{M_t^2-T_t^D(M):t\geq 0\}$ is a martingale*[31].

Proof. Since $t_i\to\infty$, $T_t^D(M)$ is integrable for each $t>0$. For $t>s\geq 0$, there is a unique k such that $t_k<s\leq t_{k+1}$, by Exercise 3.54, so that

$$\mathbb{E}\left[(M_{t_{i+1}\wedge t}-M_{t_i\wedge t})^2|\mathscr{F}_s\right]=\mathbb{E}\left(M^2_{t_{i+1}\wedge t}-M^2_{t_i\wedge t}|\mathscr{F}_s\right)$$

for $i>k$, and

$$\begin{aligned}\mathbb{E}\left[(M_{t_{k+1}\wedge t}-M_{t_k\wedge t})^2|\mathscr{F}_s\right]&=\mathbb{E}\left[(M_{t_{k+1}\wedge t}-M_s+M_s-M_{t_k\wedge t})^2|\mathscr{F}_s\right]\\&=(M_s-M_{t_k})^2+\mathbb{E}\left[(M_{t_{k+1}\wedge t}-M_s)^2|\mathscr{F}_s\right]\\&=(M_{t_{k+1}\wedge s}-M_{t_k\wedge s})^2+\mathbb{E}\left[(M_{t_{k+1}\wedge t}-M_s)^2|\mathscr{F}_s\right].\end{aligned}$$

Hence

$$\begin{aligned}\mathbb{E}\left[T_t^D(M)|\mathscr{F}_s\right]&=\sum_{i<k}(M_{t_{i+1}}-M_{t_i})^2+\mathbb{E}\left[(M_{t_{k+1}\wedge t}-M_{t_k\wedge t})^2|\mathscr{F}_s\right]\\&\quad+\mathbb{E}\left[\sum_{i>k}(M_{t_{i+1}\wedge t}-M_{t_i\wedge t})^2|\mathscr{F}_s\right]\\&=T_s^D(M)+\mathbb{E}\left[(M_{t_{k+1}\wedge t}-M_s)^2+\sum_{i>k}(M_{t_{i+1}\wedge t}-M_{t_i\wedge t})^2|\mathscr{F}_s\right]\\&=T_s^D(M)+\mathbb{E}\left[M_t^2-M_s^2|\mathscr{F}_s\right]\end{aligned}$$

which implies that $\{M_t^2-T_t^D(M):t\geq 0\}$ is a martingale. □

[30] However, we should emphasize that $t\mapsto T_t^D(X)$ is in general not non-decreasing!

[31] As we have indicated, $T^D(X)$ for this case is adapted, continuous with initial zero, but in general fails to be an increasing process.

Lemma 3.60. *Suppose $M=(M_t)_{t\geq 0}$ is a bounded martingale, then $\mathbb{E}\left[T_t^D(M)\right] = \mathbb{E}\left[M_t^2 - M_0^2\right]$ and $\mathbb{E}\left[(T_t^D(M))^2\right]$ is bounded uniformly in $t\geq 0$ and $D\in\mathscr{P}$.*

Proof. Suppose $|M_t(\omega)|\leq K$ (for all $t\geq 0$ and $\omega\in\Omega$) for some constant K. The equality $\mathbb{E}\left[T_t^D(M)\right] = \mathbb{E}\left[M_t^2 - M_0^2\right]$ follows immediately from Lemma 3.54, so that $\mathbb{E}\left[T_t^D(M)\right]$ is bounded uniformly in t and the partition D. Let us consider $\xi_t^D := T_t^D(M) - (M_t^2 - M_0^2)$ which has mean zero, so that $T_t^D(M) = \xi_t^D + M_t^2 - M_0^2$. Since M is bounded, we therefore only need to show that $\mathbb{E}\left[\left(\xi_t^D\right)^2\right]$ is bounded uniformly in t and partition $D\in\mathscr{P}$. Since

$$M_t^2 - M_0^2 = \sum_i (M_{t_{i+1}\wedge t}^2 - M_{t_i\wedge t}^2)$$

and therefore

$$\xi_t^D = \sum_i \left((M_{t_{i+1}\wedge t} - M_{t_i\wedge t})^2 - (M_{t_{i+1}\wedge t}^2 - M_{t_i\wedge t}^2)\right).$$

For simplify our notations, let us introduce

$$I_i^D = (M_{t_{i+1}\wedge t} - M_{t_i\wedge t})^2 - (M_{t_{i+1}\wedge t}^2 - M_{t_i\wedge t}^2)$$

so that $\xi_t^D = \sum_i I_i^D(t)$. Note each $I_i^D(t)$ has mean zero as we have seen before:

$$\mathbb{E}\left[(M_t - M_s)^2 | \mathscr{F}_s\right] = \mathbb{E}\left[M_t^2 - M_s^2 | \mathscr{F}_s\right]$$

for any $t>s$ (where M is square integrable martingale).

$$\begin{aligned}\mathbb{E}\left[\left(\xi_t^D\right)^2\right] &= \mathbb{E}\left[\left(\sum_i I_i^D(t)\right)^2\right] \\ &= \sum_i \mathbb{E}\left(I_i^D(t)\right)^2 + 2\sum_{i<j}\mathbb{E}\left(I_i^D(t) I_j^D(t)\right).\end{aligned}$$

Next we use the important computation again, to show that when $i<j$, the expectation $\mathbb{E}\left(I_i^D(t)I_j^D(t)\right)$ for $i\neq j$ vanishes. In fact, if $i<j$, then $t_j\wedge t\geq t_{i+1}\wedge t$, so that $I_i^D(t)$ is $\mathscr{F}_{t_j\wedge t}$ measurable, thus

$$\mathbb{E}\left[I_i^D(t)I_j^D(t)|\mathscr{F}_{t_j\wedge t}\right] = I_i^D(t)\mathbb{E}\left[I_j^D(t)|\mathscr{F}_{t_j\wedge t}\right] = 0$$

and therefore $\mathbb{E}\left[I_i^D(t)I_j^D(t)\right] = 0$ for $i<j$. It follows that

$$\begin{aligned}\mathbb{E}\left[\left(\xi_t^D\right)^2\right] &= \sum_i \mathbb{E}\left(I_i^D(t)\right)^2 \\ &= \sum_i \mathbb{E}\left((M_{t_{i+1}\wedge t} - M_{t_i\wedge t})^2 - (M_{t_{i+1}\wedge t}^2 - M_{t_i\wedge t}^2)\right)^2 \\ &= 4\mathbb{E}\sum_{i=0}^{j} M_{t_{i+1}\wedge t}^2 (M_{t_{i+1}\wedge t} - M_{t_i\wedge t})^2 \\ &\leq 4K^2\mathbb{E}\sum_{i=0}^{j}(M_{t_{i+1}\wedge t} - M_{t_i\wedge t})^2 = 4K^2\mathbb{E}\left[T_t^D(M)\right]\end{aligned}$$

which is therefore bounded uniformly in t and $D \in \mathscr{P}$. Hence $\mathbb{E}\left[(T_t^D(M))^2\right]$ is bounded uniformly in $D \in \mathscr{P}$ and $t \geq 0$. $\square$

Suppose $D = \{t_i\} \in \mathscr{P}$ is a partition of $[0,\infty)$, then $|D| = \sup_{i\geq 1}(t_i - t_{i-1})$ denotes the mesh size. If $D_1 = \{t_i\}$ and $|D|_2 = \{t_i'\}$ be partitions with $t_0 = t_0' = 0$ and $t_i \to \infty$ and $t_i^{'} \to \infty$, then $D_1 \sqcup D_2$ denotes the partition by merging the two sequences $\{t_i : i \geq 0\} \cup \{t_i' \geq 0\}$ and list the union in an order to form a new partition of $[0,\infty)$. Then $D_1 \sqcup D_2 \in \mathscr{P}$, and $|D_1 \sqcup D_2| \leq \max\{|D_1|, |D_2|\}$.

Lemma 3.61. *Let $M = (M_t)_{t\geq 0}$ be a bounded martingale. Suppose $D_1 = \{t_i : i \geq 1\} \in \mathscr{P}$ and $D_2 = \{s_i : i \geq 1\} \in \mathscr{P}$ such that $D_1 \subset D_2$ (i.e. D_2 is finer than D_1). Let $X = T^{D_1}(M)$. Then for every $t \geq 0$*

$$T_t^{D_2}(X) \leq T_t^{D_2}(M) \sup_{i\geq 0}\left(M_{s_{i+1}\wedge t} + M_{s_i\wedge t} - 2M_{t_{l(i)}\wedge t}\right)^2$$

where $l = l(i)$ is the unique non-negative integer such that

$$t_l \leq s_i < s_{i+1} \leq t_{l+1} \textit{ for } i = 0,1,\ldots.$$

Proof. In fact, for every i we have

$$\begin{aligned} T^{D_1}_{s_{i+1}\wedge t}(M) - T^{D_1}_{s_i\wedge t}(M) &= (M_{s_{i+1}\wedge t} - M_{t_{l(i)}\wedge t})^2 - (M_{s_i\wedge t} - M_{t_{l(i)}\wedge t})^2 \\ &= (M_{s_{i+1}\wedge t} - M_{s_i\wedge t})(M_{s_{i+1}\wedge t} + M_{s_i\wedge t} - 2M_{t_{l(i)}\wedge t}) \end{aligned}$$

and therefore it follows immediately that

$$\begin{aligned} T_t^{D_2}(X) &= \sum_i \left(T^{D_1}_{s_{i+1}\wedge t}(M) - T^{D_1}_{s_i\wedge t}(M)\right)^2 \\ &\leq T_t^{D_2}(M) \sup_i (M_{s_{i+1}\wedge t} + M_{s_i\wedge t} - 2M_{t_{l(i)}\wedge t})^2 \end{aligned}$$

for any $t \geq 0$. The proof is complete. $\square$

Lemma 3.62. *Let $M = (M_t)_{t\geq 0}$ be a bounded and continuous martingale, and $D^{(n)} \in \mathscr{P}$ (for $n = 1,2,\ldots$) be a sequence partitions such that $|D^{(n)}| \to 0$ as $n \to \infty$. Then $M^2 - T^{D^{(n)}}(M)$ (for $n = 1,2,\ldots$) is Cauchy sequence in $\mathscr{M}_2^c$. Therefore $\langle M,M\rangle_t = \lim_{n\to\infty} T_t^{D^{(n)}}(M)$ (in probability) exists. $\langle M,M\rangle$ is continuous, non-decreasing with initial zero, such that $\{M_t^2 - \langle M,M\rangle_t : t \geq 0\}$ is continuous square integrable martingale.*

Proof. We shall use the following notations. Let $D^{(n)} = \{t_i^{(n)} : i \geq 1\}$ and

$$D^{(n,m)} = D^{(n)} \sqcup D^{(m)} = \{t_i^{(n,m)} : i \geq 1\}.$$

Let $N^{(n)} = M^2 - T^{D^{(n)}}(M)$ and $X^{(n)} = T^{D^{(n)}}(M)$. Let $X^{(n,m)} = X^{(n)} - X^{(m)} = N^{(m)} - N^{(n)}$. Then $X^{(n,m)}$ are continuous martingales with initial zero, and bounded on

$[0,\infty)\times\Omega$. We shall show that $\mathbb{E}\left[|X_t^{(n,m)}|^2\right]\to 0$ as $n,m\to\infty$ which implies that $N^{(n)}:=M^2-T^{D^{(n)}}(M)$ (for $n=1,2,\ldots$) is a Cauchy sequence in $\mathscr{M}_2^c$. All other claims follow then immediately.

Since $X^{(n,m)}$ is a continuous martingale bounded on $[0,\infty)\times\Omega$, so that $\left(X^{(n,m)}\right)^2 - T^{D^{(n,m)}}(X^{(n,m)})$ is a continuous martingale for every n,m, and

$$\mathbb{E}\left[\left(X_t^{(n,m)}\right)^2\right]=\mathbb{E}\left[T_t^{D^{(n,m)}}(X^{(n,m)})\right]$$

for every $t\geq 0$. Since $(a+b)^2\leq 2a^2+2b^2$, we therefore deduce that

$$\mathbb{E}\left[T_t^{D^{(n,m)}}(X^{(n,m)})\right]\leq 2\mathbb{E}\left[T_t^{D^{(n,m)}}(X^{(n)})\right]+2\mathbb{E}\left[T_t^{D^{(n,m)}}(X^{(m)})\right]$$

for every $t\geq 0$. Let us show that $\mathbb{E}\left[T_t^{D^{(n,m)}}(X^{(n)})\right]\to 0$ and $\mathbb{E}\left[T_t^{D^{(n,m)}}(X^{(m)})\right]\to 0$ as $n,m\to\infty$. By the previous lemma we have

$$T_t^{D^{(n,m)}}(X^{(n)})\leq T_t^{D^{(n,m)}}(M)\sup_i\left(M_{t_{i+1}^{(n,m)}\wedge t}+M_{t_i^{(n,m)}\wedge t}-2M_{t_{l(i)}^{(n)}\wedge t}\right)^2$$

where

$$t_{l(i)}^{(n)}\leq t_i^{(n,m)}<t_i^{(n,m)}\leq t_{l(i)+1}^{(n)}.$$

Since as $n\to\infty$, $t_{i+1}^{(n,m)}-t_{l(i)}^{(n)}\leq |D^{(n)}|\to 0$ and $t_{i+1}^{(n,m)}-t_{l(i)}^{(n)}\leq |D^{(n)}|\to 0$, therefore

$$\sup_i\left(M_{t_{i+1}^{(n,m)}\wedge t}+M_{t_i^{(n,m)}\wedge t}-2M_{t_{l(i)}^{(n)}\wedge t}\right)\to 0$$

for every $t\geq 0$, as M is continuous so it is uniformly continuous on $[0,t]$. Hence

$$\mathbb{E}\left[T_t^{D^{(n,m)}}(X^{(n)})\right]\leq\sqrt{\mathbb{E}\left[(T_t^{D^{(n,m)}}(M))^2\right]}\sqrt{\mathbb{E}\sup_i\left(M_{t_{i+1}^{(n,m)}\wedge t}+M_{t_i^{(n,m)}\wedge t}-2M_{t_{l(i)}^{(n)}\wedge t}\right)^4}$$

for every $t\geq 0$. By Lemma 3.60, $\sqrt{\mathbb{E}\left[(T_t^{D^{(n,m)}}(M))^2\right]}$ is bounded in n,m and t, and

$$\sup_i\left(M_{t_{i+1}^{(n,m)}\wedge t}+M_{t_i^{(n,m)}\wedge t}-2M_{t_{l(i)}^{(n)}\wedge t}\right)^4$$

is bounded too. Thus by Lebesgue's bounded convergence theorem (Paper A4),

$$\mathbb{E}\left[\sup_i\left(M_{t_{i+1}^{(n,m)}\wedge t}+M_{t_i^{(n,m)}\wedge t}-2M_{t_{l(i)}^{(n)}\wedge t}\right)^4\right]\to 0$$

and therefore $\mathbb{E}\left[T_t^{D^{(n,m)}}(X^{(n)})\right]\to 0$ as $n\to\infty$. Similarly $\mathbb{E}\left[T_t^{D^{(n,m)}}(X^{(m)})\right]\to 0$ as $m\to\infty$ which completes the proof. □

Exercise 3.63. *Let $M = (M_t)_{t\geq 0}$ be a continuous and bounded martingale. Prove that $\langle M^T, M^T \rangle = \langle M, M \rangle^T$ for any stopping time. [*Hint: use the fact that, by the optional stopping theorem, $(M^T)^2 - \langle M, M \rangle^T$ is a martingale for every stopping time.*]*

Theorem 3.64. [32] *Let $M = (M_t)_{t\geq 0}$ be a* continuous *local martingale. There exists a unique (up to indistinguishability) non-decreasing, continuous and adapted process $\{\langle M, M \rangle_t : t \geq 0\}$ with $\langle M, M \rangle_0 = 0$, such that $\{M_t^2 - \langle M, M \rangle_t : t \geq 0\}$ is a continuous local martingale. The process $\langle M, M \rangle$ is called the* quadratic variation *of M, or simply* the increasing process of *M, and is often denoted $\langle M, M \rangle_t = \langle M \rangle_t$.*

Proof. Let $T^{(n)} = \inf\{t \geq 0 : |M_t| \geq n\}$ (where $n = 1, 2, \ldots$). Then $T^{(n)} \uparrow \infty$ and each $M^{T^{(n)}}$ is a continuous and bounded martingale, so its quadratic process $\langle M^{T_n}, M^{T_n} \rangle$ exists. If $n \geq m$, then

$$\langle M^{T_n}, M^{T_n} \rangle^{T_m} = \langle M^{T_m}, M^{T_m} \rangle$$

which implies that

$$\langle M^{T_n}, M^{T_n} \rangle = \langle M^{T_m}, M^{T_m} \rangle \quad \text{on } \{t \leq T_m\}$$

for any $n \geq m$, so there is a unique process $\langle M, M \rangle$ such that $\langle M, M \rangle = \langle M^{T_n}, M^{T_n} \rangle$ on $\{t \leq T_n\}$ for every $n = 1, 2, \ldots$. Clearly $\langle M, M \rangle$ is continuous, non-decreasing and adapted with initial zero, and $\langle M, M \rangle^{T_n} = \langle M^{T_n}, M^{T_n} \rangle$, which implies that $\left(M^2 - \langle M, M \rangle\right)^{T_n}$ is a continuous martingale. The proof is complete. □

We next shall give further information about the quadratic variation process of a continuous local martingale. To this end we introduce a few notations. Recall that for every $t \geq 0$, $\mathscr{P}[0,t]$ denotes the collection of all finite partitions $D = \{t_i : i = 0, \ldots, m\}$ where $0 = t_0 < t_1 < \cdots < t_m = t\}$, $m = m(D)$ is an non-negative integer, and $|D| = \max_i (t_i - t_{i-1})$.

Proposition 3.65. *Let $M = (M_t)_{t\geq 0}$ be a continuous local martingale. Then for every $t \geq 0$, and for any sequence of finite partitions $D^{(n)}(t) = \{t_i^{(n)}\}$ of $[0,t]$ such that $|D^{(n)}(t)| \to 0$ as $n \to \infty$, we have*[33]

$$\langle M, M \rangle_t = \lim_{n\to\infty} \sum_{i=1}^{m(D^{(n)}(t))} (M_{t_i^{(n)}} - M_{t_{i-1}^{(n)}})^2 \quad \textit{in probability.}$$

Proof. Let $T^{(n)} = \inf\{t \geq 0 : |M_t| \geq n\}$. Then for any sequence of partitions $D^{(m)} = \{s_i^{(m)} : i \geq 0\}$ of $[0, \infty)$ such that $s_i^{(m)} \to \infty$ as $i \to \infty$, and $|D^{(m)}| \to 0$,

$$T_t^{D^{(m)}}(M^{T^{(n)}}) \to \left\langle M^{T^{(n)}} \right\rangle_t \quad \text{in probability}$$

[32]This theorem is a consequence of Doob-Meyer's decomposition for sub-martingales, which says that every "class D" right continuous sub-martingale is a sum of a right continuous local martingale and a "predictable" non-decreasing process.

[33]Therefore $t \mapsto \langle M, M \rangle_t$ is non-decreasing !

as $m \to \infty$, for every $t \geq 0$, where $n = 1, 2, \ldots$. Hence for every $\delta > 0$ and for every $n = 1, 2, \ldots$

$$\begin{aligned}\mathbb{P}\left[\left|T_t^{D^{(m)}}(M) - \langle M \rangle_t\right| > \delta\right] &\leq \mathbb{P}\left[\left|T_t^{D^{(m)}}(M) - \langle M \rangle_t\right| > \delta, t < T^{(n)}\right] \\ &+ \mathbb{P}\left[t \geq T^{(n)}\right] \\ &\leq \mathbb{P}\left[\left|T_t^{D^{(m)}}(M^{T^{(n)}}) - \left\langle M^{T^{(n)}} \right\rangle\right| > \delta\right] \\ &+ \mathbb{P}\left[t \geq T^{(n)}\right]\end{aligned}$$

By letting $m \to \infty$

$$\limsup_{m\to\infty} \mathbb{P}\left[\left|T_t^{D^{(m)}}(M) - \langle M \rangle_t\right| > \delta\right] \leq \mathbb{P}\left[t \geq T^{(n)}\right]$$

for every n. Since $T^{(n)} \uparrow \infty$, so that $\mathbb{P}\left[t \geq T^{(n)}\right] \to 0$ as $n \to \infty$ for every $t \geq 0$. Therefore

$$\lim_{m\to\infty} \mathbb{P}\left[\left|T_t^{D^{(m)}}(M) - \langle M \rangle_t\right| > \delta\right] = 0$$

for every $\delta > 0$ and $t \geq 0$, which completes the proof. □

Proposition 3.66. *Let $M = (M_t)_{t\geq 0}$ be a continuous, square integrable martingale, i.e. $\mathbb{E}\left[M_t^2\right] < \infty$ for every $t \geq 0$. Then $M^2 - \langle M, M \rangle$ is a continuous martingale.*

Proof. According to Doob's L^2-inequality $\mathbb{E}\left[\sup_{s\leq t} M_s^2\right] \leq 4\mathbb{E}\left[M_t^2\right] < \infty$ for every $t \geq 0$, that is, $\sup_{s\leq t} M_s^2$ is integrable. Let $T^{(n)} = \inf\{t \geq 0 : |M_t| \geq n\}$. Then $(M_{t\wedge T^{(n)}})^2 - \langle M \rangle_{t\wedge T^{(n)}}$ is a square integrable martingale (cf. Lemma 3.62), so that

$$\mathbb{E}\left[\langle M \rangle_{t\wedge T^{(n)}}\right] = \mathbb{E}\left[(M_{t\wedge T^{(n)}})^2\right] \leq \mathbb{E}\left[\sup_{s\leq t} M_s^2\right] < \infty$$

for every n. Letting $n \to \infty$, by Fatou's lemma

$$\mathbb{E}\left[\langle M \rangle_t\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[\langle M \rangle_{t\wedge T^{(n)}}\right] \leq \mathbb{E}\left[\sup_{s\leq t} M_s^2\right] < \infty$$

and therefore $M_t^2 - \langle M \rangle_t$ is integrable for every $t \geq 0$. Moreover for $t > s \geq 0$,

$$\mathbb{E}\left[M^2_{t\wedge T^{(n)}} - \langle M \rangle_{t\wedge T^{(n)}} | \mathscr{F}_s\right] = M^2_{s\wedge T^{(n)}} - \langle M \rangle_{s\wedge T^{(n)}}$$

for every n. Since

$$\left|M^2_{t\wedge T^{(n)}} - \langle M \rangle_{t\wedge T^{(n)}}\right| \leq \sup_{s\leq t} M_s^2 + \langle M \rangle_t \qquad \text{for any } n$$

and $\sup_{s\leq t} M_s^2 + \langle M \rangle_t$ is integrable, so by Dominated Convergence Theorem (Paper A4) and the fact that we obtain

$$\mathbb{E}\left[M_t^2 - \langle M \rangle_t | \mathscr{F}_s\right] = M_s^2 - \langle M \rangle_s .$$

Therefore $M^2 - \langle M \rangle$ is a martingale. □

Exercise 3.67. *1) Let $M = (M_t)_{t\geq 0}$ be a continuous local martingale. Prove that M is a square integrable martingale if and only if $\langle M \rangle_t$ is integrable for every $t \geq 0$.*

2) Let $M = (M_t)_{t\geq 0}$ be a continuous martingale bounded in L^2, i.e. $\sup_{t\geq} \mathbb{E}\left[M_t^2\right] < \infty$. Then $M^2 - \langle M, M \rangle$ is a continuous uniformly integrability martingale.

Here is the proof. According to Doob's L^2-inequality $\mathbb{E}\left[\sup_{s\leq t} M_s^2\right] \leq 4\sup_{t\geq 0} \mathbb{E}\left[M_t^2\right]$ for every $t \geq 0$. Letting $t \to \infty$

$$\mathbb{E}\left[\sup_{s\geq 0} M_s^2\right] \leq 4 \sup_{t\geq 0} \mathbb{E}\left[M_t^2\right] < \infty.$$

Therefore $\sup_s M_s^2$ is integrable. Since

$$\mathbb{E}\left[\langle M, M \rangle_t\right] = \mathbb{E}\left[M_t^2 - M_0^2\right] \leq \sup_{t\geq 0} \mathbb{E}\left[M_t^2\right]$$

so by MCT

$$\mathbb{E}\left[\langle M, M \rangle_\infty\right] \leq \sup_{t\geq 0} \mathbb{E}\left[M_t^2\right] < \infty.$$

Since

$$\left|M_t^2 - \langle M, M \rangle_t\right| \leq \sup_s M_s^2 + \langle M, M \rangle_\infty$$

for every $t \geq 0$. Hence $\{M_t^2 - \langle M, M \rangle_t ; t \geq 0\}$ is uniformly integrable.

We can see that the quadratic variation of a martingale is telling us something about how its variance increases with time. We also need an analogous quantity for the 'covariance' between two martingales. This is most easily defined through *polarisation*. The *quadratic co-variation* between two continuous local martingales M, N is defined by

$$\langle M, N \rangle := \frac{1}{2} \left(\langle M+N, M+N \rangle - \langle M, M \rangle - \langle N, N \rangle\right). \tag{8}$$

It is often called the (angle) bracket process of M and N. The following properties about the mutual quadratic variation process can be derived immediately.

Proposition 3.68. *Let M, N be two continuous local martingales.*

1) $\langle M, N \rangle$ is the unique continuous, adapted process with finite variation and initial zero, such that $MN - \langle M, N \rangle$ is a continuous local martingale.

2) Let $D^{(n)} = \{t_i^{(n)}\} \in \mathscr{P}$ be a sequence of partitions of $[0, \infty)$ with $|D^{(n)}| \to 0$ as $n \to \infty$. Then

$$\langle M, N \rangle_t = \lim_{n\to\infty} \sum_i (M_{t_{i+1}^{(n)} \wedge t} - M_{t_i^{(n)} \wedge t})(N_{t_{i+1}^{(n)} \wedge t} - N_{t_i^{(n)} \wedge t}) \quad \text{in probability,} \tag{9}$$

for every $t \geq 0$.

3) $|\langle M, N \rangle_t|^2 \leq \langle M \rangle_t \langle N \rangle_t$ for any $t \geq 0$.

Exercise 3.69. *Let M,N be two continuous local martingales, and T be a stopping time. Prove that*

$$\langle M^T,N^T\rangle_t = \langle M^T,N\rangle_t = \langle M,N^T\rangle_t = \langle M,N\rangle_{T\wedge t} \tag{10}$$

for every $t \geq 0$. Hint*: By applying the sequence of stopping times $T^{(n)} = \inf\{t \geq 0 : |M_t| + |N_t| \geq n\}$, we only need to verify these equations for bounded and continuous martingales M and N. Then, by optional stopping, $M^T N^T - \langle M,N\rangle^T$ is a martingale, so that $\langle M^T,N^T\rangle = \langle M,N\rangle^T$. Next show, again by optional stopping, $M^T N^T - \langle M,N^T\rangle$ is a martingale, which shall gives that $\langle M^T,N^T\rangle = \langle M,N^T\rangle$ which however follows from (9) immediately.*

Theorem 3.70 (Kunita–Watanabe inequality)**.** *Let M,N be continuous local martingales and K,H two measurable processes. Then for all $0 \leq t \leq \infty$,*

$$\int_0^t |H_s||K_s||d\langle M,N\rangle_s| \leq \left(\int_0^t H_s^2 d\langle M\rangle_s\right)^{1/2} \left(\int_0^t K_s^2 d\langle N\rangle_s\right)^{1/2} \quad a.s.. \tag{11}$$

We shall introduce a large class of stochastic processes with which Itô's stochastic integrals may be defined, which is on the other hand rich enough for defining stochastic models in applications in various scientific areas such as quantitative finance, turbulence modeling and etc.

A continuous, adapted process $X = (X_t)_{t\geq 0}$ is called a *continuous semimartingale* if

$$X_t - X_0 = M_t + A_t, \quad t \geq 0 \tag{12}$$

where $M = (M_t)_{t\geq 0}$ is a continuous local martingale with initial $M_0 = 0$, $A = (A_t)_{t\geq 0}$ is a continuous, adapted process of finite variation with initial$A_0 = 0$. M is called the martingale part, and A is the variational part of X.

Exercise 3.71. *The decomposition is unique (up to indistinguishability).*

Proposition 3.72. *Let X be a continuous semimartingale with its martingale part M and variational part A. Let $D^{(n)} = \{t_i^{(n)}\} \in \mathscr{P}$ be a sequence of partitions with $|D^{(n)}| \to 0$ when $n \to \infty$. Then*

$$\lim_{n\to\infty} \sum_{i=1}^{m(D^{(n)})} (X_{t_i^{(n)}\wedge t} - X_{t_{i-1}^{(n)}\wedge t})^2 = \langle M,M\rangle_t \quad \text{in probability,}$$

for every $t \geq 0$. Therefore a continuous semimartingale is of finite quadratic variation[34] *and $\langle X,X\rangle = \langle M,M\rangle$.*

Proof. For every $t \geq 0$

$$\sum_i (X_{t_i} - X_{t_{i-1}})^2 = \underbrace{\sum_i (M_{t_i} - M_{t_{i-1}})^2}_{(i)} + \underbrace{\sum_i (A_{t_i} - A_{t_{i-1}})^2}_{(ii)} + 2\underbrace{\sum_i (M_{t_i} - M_{t_{i-1}})(A_{t_i} - A_{t_{i-1}})}_{(iii)}.$$

[34] Here we use the convention that $\langle X,X\rangle_t = \lim_{|D|\to 0} \sum_i (X_{t_i\wedge t} - X_{t_{i-1}\wedge t})^2$ as long as the quadratic variation exists.

It follows from the properties of M and A that, as the partition mesh $|D| \to 0$,

$$(i) \to \langle M, M \rangle_t,$$
$$(ii) \leq \sup_{1 \leq i \leq n_m} |A_{t_i} - A_{t_{i-1}}| \cdot V_t(A) \to 0 \quad \text{a.s.} ,$$
$$(iii) \leq \sup_{1 \leq i \leq n_m} |M_{t_i} - M_{t_{i-1}}| \cdot V_t(A) \to 0 \quad \text{a.s.} .$$

□

If X, Y are two continuous semimartingales, we can define their co-variation $\langle X, Y \rangle$ via the polarisation formula that we used for martingales. If $X_t = X_0 + M_t + A_t$ and $Y_t = Y_0 + N_t + V_t$, then $\langle X, Y \rangle_t = \langle M, N \rangle_t$.

4 Brownian Motion

In this section we give the definition and the construction of Brownian motion, and discuss several important aspects of Brownian motion.

4.1 Definition, and properties

Our fundamental building block will be Brownian motion. It is often described as an 'infinitesimal random walk', so to motivate the definition, we take a quick look at simple (discrete time) random walk.

Exercise 4.1. *The symmetric random walk $\{S_n : n \geq 0\}_{n \geq 0}$,where $S_0 = 0$, $S_n = \sum_{i=1}^n \xi_i$ for $n \geq 1$, ξ_i i.i.d. with distribution given by $\mathbb{P}[\xi_i = -1] = 1/2 = \mathbb{P}[\xi_i = 1]$. Then $cov(S_n, S_m) = n \wedge m$. Since $\mathbb{E}\xi_i = 0$ and $var(\xi_i) = 1$, by Central Limit Theorem*

$$\mathbb{P}\left[\frac{S_n}{\sqrt{n}} \leq x\right] \to \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \text{ as } n \to \infty.$$

and

$$\mathbb{P}\left[\frac{S_{[nt]}}{\sqrt{n}} \leq x\right] \to \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi t}} e^{-y^2/2t} dy \text{ as } n \to \infty,$$

where $[nt]$ denotes the integer part of nt for every $t > 0$.

Heuristically at least, passage to the limit from simple random walk suggests the following definition of Brownian motion.

Definition 4.2 (Brownian motion)**.** *An $\mathbb{R}^d$-valued continuous process $B = (B_t)_{t \geq 0}$ on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is a Brownian motion, if it has independent increments: for any $0 \leq t_0 < t_1 < t_2 < \cdots < t_n$, the random variables B_{t_0}, $B_{t_1} - B_{t_0}$, ..., $B_{t_n} - B_{t_{n-1}}$ are independent, and the increment $B_t - B_s$ for any $t > s \geq 0$, has a (d-dimensional) normal distribution $N(0, (t-s)I)$. Suppose the distribution of B_0 is μ, a probability measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$,then B is called Brownian motion with initial distribution μ. If $B_0 = \xi$ almost surely for some $\xi \in \mathbb{R}^d$,then B is called Brownian motion started from ξ. A standard Brownian motion is a Brownian motion started from 0.*

We shall leave the existence (or the construction) of Brownian motion as a topic to be discussed late on, and we shall look at the important properties first.

If $B = (B_t)_{t\geq 0}$ is a Brownian motion (in short, BM) of d-dimensions, $B_0 = 0$, and $\mathscr{F}_t^0 = \sigma\{B_s : s \leq t\}$ (for $t \geq 0$) the natural filtration generated by the process B. Let $s \geq 0$ be any but fixed. For any partition $D : 0 \leq t_0 < \cdots < t_n \leq s$, and any $t > s$, by definition, B_{t_0}, $B_{t_1} - B_{t_0}$ $\cdots$, $B_{t_n} - B_{t_{n-1}}$ and $B_t - B_s$ are independent, and therefore the family $\{B_{t_1}, \cdots, B_{t_n}\}$ and $B_t - B_s$ are independent. Hence $\sigma\{B_{t_0}, \cdots, B_{t_n}\}$ and $B_t - B_s$ are independent for any $t_i \leq s$, and $t > s$. Therefore $B_t - B_s$ and $\mathscr{F}_s^0$ are independent for any $t > s \geq 0$. By continuity of the sample path $t \mapsto B_t$, we may therefore conclude that $B_t - B_s$ and $\mathscr{F}_{s+}^0$ are independent for any $t > s \geq 0$.

4.1.1 Martingale property

Let us begin with the following simple but very important property of Brownian motion.

Proposition 4.3. *Let $B = (B^1, \cdots, B^d)$ be a d-dimensional standard Brownian motion, $B_0 = 0$. Then*

1) each B^i is a continuous square integrable martingale, and

2) the covariation process[35] $\langle B^i, B^j \rangle_t = \delta_{ij} t$, that is, $\{B_t^i B_t^j - \delta_{ij} t; t \geq 0\}$ is a martingale for every pair i, j.

Proof. $\mathbb{E}\left[|B_t^i|^2\right] = t$ for every i, so B is square integrable. Let $\mathscr{F}_t^0 = \sigma\{B_s : s \leq t\}$ be the natural filtration of B. Then $B_t - B_s$ and $\mathscr{F}_s^0$ are independent for every $t > s \geq 0$ by the independence of increments of B over disjoint intervals. Therefore, for $t > s \geq 0$, $\mathbb{E}\left[B_t^i - B_s^i | \mathscr{F}_s^0\right] = \mathbb{E}\left[B_t^i - B_s^i\right] = 0$. On the other hand

$$\begin{aligned}
\mathbb{E}\left[(B_t^i)^2 - (B_s^i)^2 | \mathscr{F}_s^0\right] &= \mathbb{E}\left[|B_t^i - B_s^i|^2 + 2B_s^i(B_t^i - B_s^i) | \mathscr{F}_s^0\right] \\
&= \mathbb{E}\left[|B_t^i - B_s^i|^2 | \mathscr{F}_s^0\right] = \mathbb{E}|B_t^i - B_s^i|^2 \\
&= t - s
\end{aligned}$$

so that $(B_t^i)^2 - t$ (for $t \geq 0$) is a martingale, which implies that $\langle B^i, B^i \rangle_t = t$. Similarly if $i \neq j$ and $t > s \geq 0$

$$\begin{aligned}
\mathbb{E}\left[B_t^i B_t^j - B_s^i B_s^j | \mathscr{F}_s^0\right] &= \mathbb{E}\left[(B_t^i - B_s^i)(B_t^j - B_s^j) + B_s^i(B_t^j - B_s^j) + B_s^j(B_t^i - B_s^i) | \mathscr{F}_s^0\right] \\
&= \mathbb{E}\left[(B_t^i - B_s^i)(B_t^j - B_s^j) | \mathscr{F}_s^0\right] \\
&= \mathbb{E}\left[(B_t^i - B_s^i)(B_t^j - B_s^j)\right] \\
&= \mathbb{E}(B_t^i - B_s^i)\mathbb{E}(B_t^j - B_s^j) = 0
\end{aligned}$$

which implies that $B^i B^j$ (for $i \neq j$) is a continuous martingale, hence $\langle B^i, B^j \rangle = 0$ for $i \neq j$. □

[35] $\delta_{ij} = 1$ or 0 according to $i = j$ or not.

It follows the following sample property of Brownian motion.

Corollary 4.4. *Let $B = (B_t)_{t\geq 0}$ be a one dimensional standard Brownian motion.*

1) For every $t > 0$ the quadratic variation

$$\lim_{D:0=t_0<\cdots<t_n=1,|D|\to 0} \sum_{i=1}^{n} |B_{t_i} - B_{t_{i-1}}|^2 = t \qquad \text{in probability.}$$

2) Brownian sample paths are of infinite variation on any non-trivial interval almost surely.

3) Brownian sample paths are almost surely nowhere locally Hölder continuous of order $\gamma > \frac{1}{2}$.

In fact, a very precise statement is possible.

Theorem 4.5. *For B a Brownian motion,*

$$\limsup_{\varepsilon\downarrow 0} \sup_{0\leq s<t\leq 1, t-s\leq\varepsilon} \frac{|B_t - B_s|}{\sqrt{2\varepsilon\log(1/\varepsilon)}} = 1 \qquad a.s.$$

Consequently, Brownian sample paths are almost surely nowhere locally Hölder continuous of order $\gamma = 1/2$, and the 2-variation is almost surely infinite.

Proof. [Lévy's modulus of continuity (*Not Examinable*)] Omitted (proof is a careful calculation with estimates of normal random variables, see, for example, H. P. McKean: *Stochastic Integrals*. Academic Press (1969), Section 1.6, page 14.) □

Proposition 4.6. *Let $B = (B^1, \cdots, B^d)$ be a d-dimensional standard Brownian motion, and $\mathscr{F}_t^0 = \sigma\{B_s : s \leq t\}$.*

1) For every $\xi \in \mathbb{R}^d$ and $t > s \geq 0$[36]

$$\mathbb{E}\left[e^{i\xi\cdot(B_t - B_s)} | \mathscr{F}_s^0\right] = e^{-\frac{1}{2}(t-s)|\xi|^2}.$$

2) Let $\xi \in \mathbb{R}^d$. Define $M_t = e^{\xi\cdot B_t - \frac{1}{2}|\xi|^2 t}$ for $t \geq 0$. Then $M = (M_t)_{t\geq 0}$ is a continuous square integrable martingale.

Proof. We know that $B_t - B_s$ is independent of $\mathscr{F}_s^0$ and has a normal distribution $N(0, (t-s)I)$ for any $t > s \geq 0$, so that

$$\mathbb{E}\left[e^{\mathrm{i}\xi\cdot(B_t - B_s)} | \mathscr{F}_s^0\right] = \mathbb{E}\left[e^{\mathrm{i}\xi\cdot(B_t - B_s)}\right] = e^{-\frac{1}{2}(t-s)|\xi|^2}$$

[36] This means the (complex valued) process $\{e^{i\xi\cdot B_t + \frac{1}{2}|\xi|^2 t}; t \geq 0\}$ is a martingale. That is, $\left\{e^{\frac{1}{2}|\xi|^2 t}\cos(\xi\cdot B_t); t \geq 0\right\}$ and $\left\{e^{\frac{1}{2}|\xi|^2 t}\sin(\xi\cdot B_t); t \geq 0\right\}$ are martingales. We shall see later on (after establishing the Itô formula), that the reason why these are martingales, because $u(x,t) := e^{\frac{1}{2}|\xi|^2 t}\cos(\xi\cdot x)$ (similarly for the sin one) satisfies the (backward) heat equation $\frac{\partial}{\partial t}u + \frac{1}{2}\Delta u = 0$.

which[37] proves 1). To prove the second item, we notice that 1) still holds for complex vector ξ, as both sides are analytic in ξ. Therefore by replacing ξ by $-\mathrm{i}\xi$ we then obtain

$$\mathbb{E}\left[e^{\xi\cdot(B_t-B_s)}|\mathscr{F}_s^0\right]=e^{\frac{1}{2}(t-s)|\xi|^2}\quad \text{for any } t>s\geq 0,$$

(which can be indeed checked directly), rearranging the terms to deduce that

$$\mathbb{E}\left[M_t|\mathscr{F}_s^0\right]=M_s\quad \text{for } t>s\geq 0$$

which completes the proof. □

The martingales constructed in Proposition 4.6 are very useful tools in applications. We shall now give a trivial improvement over the previous proposition, which is however very important.

Proposition 4.7. *Let $B=(B^1,\cdots,B^d)$ be a d-dimensional standard Brownian motion, $B_0=0$, and $\mathscr{F}_t^0=\sigma\{B_s:s\leq t\}$ and $\mathscr{F}_{t+}^0=\bigcap_{u>t}\mathscr{F}_u^0$ for every $t\geq 0$, so that $(\mathscr{F}_{t+}^0)_{t\geq 0}$ is a right continuous filtration.*

1) B^i, B^iB^j (for $i\neq j$) and $(B_t^i)^2-t$ (for $t\geq 0$) are all $(\mathscr{F}_{t+}^0)$-martingales.

2) For every $\varepsilon>0$, $\{B_{t+\varepsilon}-B_\varepsilon:t\geq 0\}$ is a standard Brownian motion, independent of $\mathscr{F}_{\varepsilon+}^0$.

3) For every $\xi\in\mathbb{R}^d$, $N_t=e^{i\xi\cdot B_t+\frac{1}{2}|\xi|^2t}$ (for $t\geq 0$) and $M_t=e^{\xi\cdot B_t-\frac{1}{2}|\xi|^2t}$ (for $t\geq 0$) are $(\mathscr{F}_{t+}^0)$-martingales. That is,

$$\mathbb{E}\left[e^{i\xi\cdot(B_t-B_s)}|\mathscr{F}_{s+}^0\right]=e^{-\frac{1}{2}(t-s)|\xi|^2}\quad \textit{for } t>s\geq 0.$$

Proof. Let us prove 3) only, proofs of other two as exercises. Let us prove that N is an $(\mathscr{F}_{t+}^0)$-martingale, the proof of others is similar. Let $t>s\geq 0$, and for every $\varepsilon>0$ such that $t>s+\varepsilon$. Then from the previous proposition, for every $A\in\mathscr{F}_{s+}^0\subset\mathscr{F}_{s+\varepsilon}^0$, we have

$$\mathbb{E}\left[e^{\mathrm{i}\xi\cdot(B_t-B_{s+\varepsilon})}1_A\right]=e^{-\frac{1}{2}(t-s-\varepsilon)|\xi|^2}\mathbb{P}(A).$$

Since $B_{t+\varepsilon}\to B_t$ when $\varepsilon\downarrow 0$, by using the Dominated Convergence Theorem, we obtain

$$\mathbb{E}\left[e^{\mathrm{i}\xi\cdot(B_t-B_s)}1_A\right]=e^{-\frac{1}{2}(t-s)|\xi|^2}\mathbb{P}(A)\quad \text{for any } A\in\mathscr{F}_{s+}^0$$

for any $t>s\geq 0$, which is equivalent to that N is an $(\mathscr{F}_{t+}^0)$-martingale. □

As an application of the exponential martingales associated with Brownian motion, we have the following interesting consequence.

[37] We recall that a random variable Z has a normal distribution $N(a,\sigma^2)$ if and only if its characteristic function $\mathbb{E}\left[e^{i\xi Z}\right]=\exp\left(ia\xi-\frac{1}{2}\sigma^2|\xi|^2\right)$.

Proposition 4.8. *Let $B = (B_t)_{t\geq 0}$ be a one dimensional standard Brownian motion, $B_0 = 0$. Let $a \in \mathbb{R}$ and define $\tau_a = \inf\{t > 0 : B_t = a\}$. If $a \neq 0$, then $\mathbb{P}[\tau_a < \infty] = 1$, and $\mathbb{E}[e^{-s\tau_a}] = e^{-\sqrt{2s}|a|}$ for any $s > 0$*[38].

Proof. The first hitting time τ_a is a stopping time with respect to $(\mathscr{F}^0_{t+})_{t\geq 0}$. Since $-B$ is also a standard Brownian motion, so $\mathbb{P}[\tau_a < \infty] = \mathbb{P}[\tau_{-a} < \infty]$, hence we may assume that $a > 0$. For every real ξ, $M_t = e^{\xi B_t - \frac{\xi^2}{2}t}$ (where $t \geq 0$) is a martingale, by optional stopping theorem $\mathbb{E}[M_{t\wedge\tau_a}] = \mathbb{E}[M_0] = 1$. While for $\xi > 0$, $0 < M_{t\wedge\tau_a} \leq e^{\xi a}$ and

$$M_{t\wedge\tau_a} = e^{\xi B_{t\wedge\tau_a} - \frac{\xi^2}{2}t\wedge\tau_a} \to 1_{\{\tau_a<\infty\}} e^{\xi a - \frac{\xi^2}{2}\tau_a}$$

as $t \to \infty$. Apply Bounded Convergence Theorem, we obtain from $\mathbb{E}[M_{t\wedge\tau_a}] = 1$ (letting $t \to \infty$) that

$$\mathbb{E}\left[1_{\{\tau_a<\infty\}} e^{\xi a - \frac{\xi^2}{2}\tau_a}\right] = 1.$$

Rearranging the terms to obtain that

$$\mathbb{E}\left[1_{\{\tau_a<\infty\}} e^{-\frac{\xi^2}{2}\tau_a}\right] = e^{-\xi a} \quad \text{for any } \xi > 0.$$

Next using Bounded Convergence Theorem again but let $\xi \downarrow 0$ to obtain that $\mathbb{P}[\tau_a < \infty] = 1$. Thus $\mathbb{E}\left[e^{-\frac{\xi^2}{2}\tau_a}\right] = e^{-\xi a}$ and the Laplace transform of τ_a follows immediately. □

In fact the converse of 1) in Proposition 4.7 is also true, which motivates the following definition.

Remark. Let us consider $\tau_0 = \inf\{t > 0 : B_t = 0\}$ where B is a standard real Brownian motion, and prove that $\mathbb{P}[\tau_0 = 0] = 1$, so τ_0 has distribution δ_0. To this end, for every $\varepsilon > 0$, consider $T_\varepsilon = \inf\{t > \varepsilon : B_t = 0\}$. Then $T_\varepsilon \downarrow \tau_0$ as $\varepsilon \downarrow 0$. On the other hand

$$\begin{aligned} T_\varepsilon &= \varepsilon + \inf\{t > 0 : B_{t+\varepsilon} =\} = \varepsilon + \inf\{t > 0 : B_{t+\varepsilon} - B_\varepsilon = -B_\varepsilon\} \\ &= \varepsilon + \inf\{t > 0 : X_t = -B_\varepsilon\} \end{aligned}$$

where $X_t = B_{t+\varepsilon} - B_\varepsilon$ (for $t \geq 0$) is a standard Brownian motion independent of $\mathscr{F}^0_\varepsilon$. Since B_ε has a normal distribution $N(0, \varepsilon)$, so that $|B_\varepsilon| \neq 0$ almost surely.

[38] As a consequence, by differentiating in s under integration (which is justified, Paper A4), we have $\mathbb{E}\left[\tau_a e^{-s\tau_a}\right] = \frac{|a|}{\sqrt{2s}} e^{-\sqrt{2s}|a|}$ for any $a \neq 0$. Letting $s \downarrow 0$, we prove that $\mathbb{E}[\tau_a] = \infty$ for any $a \neq 0$. That is, τ_a is finite almost surely, but not integrable for $a \neq 0$.

Therefore for any $s > 0$

$$\begin{aligned}
\mathbb{E}\left[e^{-sT_\varepsilon}\right] &= e^{-s\varepsilon}\mathbb{E}\left(\mathbb{E}\left[e^{-s\tau_{B_\varepsilon}}|\mathscr{F}^0_\varepsilon\right]\right) \\
&= e^{-s\varepsilon}\mathbb{E}\left(\mathbb{E}\left[e^{-\sqrt{2s}|B_\varepsilon|}|\mathscr{F}^0_\varepsilon\right]\right) \\
&= e^{-s\varepsilon}\mathbb{E}\left[e^{-\sqrt{2s}|B_\varepsilon|}\right] \\
&= e^{-s\varepsilon}\int_{-\infty}^{\infty} e^{-\sqrt{2s}|x|}\frac{1}{\sqrt{2\pi\varepsilon}}e^{-\frac{x^2}{2\varepsilon}}\mathrm{d}x \\
&= e^{-s\varepsilon}\int_{-\infty}^{\infty} e^{-\sqrt{2s}|\sqrt{\varepsilon}x|}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\mathrm{d}x.
\end{aligned}$$

Letting $\varepsilon \downarrow 0$ we obtain that

$$\mathbb{E}\left[e^{-s\tau_0}\right] = 1 \quad \text{for every } s > 0$$

we therefore must have $\mathbb{P}[\tau_0 = 0] = 1$, which is what we have expected.

Definition 4.9. *Let $(\mathscr{F}_t)_{t\geq 0}$ be a filtration on $(\Omega, \mathscr{F}, \mathbb{P})$. An $\mathbb{R}^d$-valued continuous process $X = (X_t)_{t\geq 0}$ is called a d-dimensional $(\mathscr{F}_t)$-Brownian motion*[39]*, if X is adapted to $(\mathscr{F}_t)_{t\geq 0}$ and*

$$\mathbb{E}\left[e^{i\xi\cdot(X_t - X_s)}|\mathscr{F}_s\right] = e^{-\frac{1}{2}(t-s)|\xi|^2} \tag{13}$$

for every $\xi \in \mathbb{R}^d$ and $t > s \geq 0$.

The equation (13) is equivalent to say that $e^{\mathrm{i}\xi\cdot(X_t - X_0)+\frac{1}{2}|\xi|^2 t}$ (for $t \geq 0$) is a continuous martingale with respect to $(\mathscr{F}_t)_{t\geq 0}$.

Proposition 4.10. *An $(\mathscr{F}_t)$-Brownian motion $X = (X_t)_{t\geq 0}$ must be a Brownian motion, and $X_t - X_s$ and $\mathscr{F}_s$ are independent for any $t > s \geq 0$.*

Proof. Let $t > s \geq 0$. Then $\mathbb{E}\left[e^{\mathrm{i}\xi\cdot(X_t - X_s)}\right] = e^{-\frac{1}{2}(t-s)|\xi|^2}$ for every ξ, so $X_t - X_s$ has a normal distribution $N(0, (t-s)I)$. Moreover, for every $A \in \mathscr{F}_s$ and $\xi \in \mathbb{R}^d$

$$\mathbb{E}\left[e^{\mathrm{i}\xi\cdot(X_t - X_s)}1_A\right] = e^{-\frac{1}{2}(t-s)|\xi|^2}\mathbb{P}(A) = \mathbb{E}\left[e^{\mathrm{i}\xi\cdot(X_t - X_s)}\right]\mathbb{P}(A)$$

which implies that $X_t - X_s$ and $\mathscr{F}_s$ are independent[40]. Let $\mathscr{F}^0_t = \sigma\{X_s : s \leq t\}$ be the filtration generated by X. Then $\mathscr{F}^0_s \subset \mathscr{F}_s$ for any s. Therefore $X_t - X_s$ and $\mathscr{F}^0_s$ are independent for any $t > s \geq 0$, which implies that X has independent increments. Hence X is a Brownian motion. □

[39] We follow the definition given in Ikeda and Watanabe, which is different from that given in the lecture notes written by the previous lecturers, but they are equivalent.

[40] Here we may use Fourier transform, or Fourier series to conclude that $f(B_t - B_s)$ is independent of $\mathscr{F}_s$, for example for smooth function f which decreasing fast enough at infinity, which allows us to conclude that $B_t - B_s$ is independent of $\mathscr{F}_s$.

Proposition 4.7 implies that a Brownian motion $B_t = (B_t)_{t\geq 0}$ is a Brownian motion with respect to the right continuous filtration $(\mathscr{F}^0_{t+})_{t\geq 0}$, where $\mathscr{F}^0_t = \sigma\{B_s : s \leq t\}$. Therefore, in the discussions of (super- / sub-) martingales related to Brownian motion, the underlying filtration may be assumed to be *right-continuous.*

Theorem 4.11 (Strong Markov Property)**.** *Let $(\mathscr{F}_t)_{t\geq 0}$ be right-continuous, and $X = (X_t)_{t\geq 0}$ be a d-dimensional $(\mathscr{F}_t)_{t\geq 0}$-Brownian motion. Let T be a stopping time with respect to $(\mathscr{F}_t)_{t\geq 0}$ with $\mathbb{P}[T < \infty] > 0$, $Y_t = X_{T+t} - X_T$ and $\mathscr{G}_t = \mathscr{F}_{T+t}$ for $t \geq 0$. Let $\mathbb{P}_T$ denote the conditional probability on $\{T < \infty\}$, that is $\mathbb{P}_T(G) = \mathbb{P}[G|T < \infty]$ for any $G \in \mathscr{F}$. Then $Y = (Y_t)_{t\geq 0}$ is $(\mathscr{G}_t)_{t\geq 0}$-Brownian motion started from zero*[41] *and is independent of $\mathscr{F}_T$ under $\mathbb{P}_T$.*

Proof. We may assume that $X_0 = 0$ and $\mathbb{P}[T < \infty] = 1$. Let $\xi \in \mathbb{R}^d$ and set $M_t = e^{\mathrm{i}\xi\cdot X_t + \frac{1}{2}|\xi|^2 t}$ for $t \geq 0$. Then $M = (M_t)_{t\geq 0}$ is a continuous martingale and $|M_t| \leq e^{\frac{1}{2}|\xi|^2 t}$, but it is not uniformly integrable unless $\xi = 0$. Let $t > s \geq 0$. Apply optional stopping time theorem with $T\wedge n + t \geq T \wedge n + s$ we obtain that

$$\mathbb{E}\left[M_{T\wedge n+t}|\mathscr{F}_{T\wedge n+s}\right] = M_{T\wedge n+s}$$

rearranging to obtain

$$\mathbb{E}\left[e^{\mathrm{i}\xi\cdot(X_{T\wedge n+t} - X_{T\wedge n+s})}|\mathscr{F}_{T\wedge n+s}\right] = e^{-\frac{1}{2}(t-s)|\xi|^2}$$

Therefore, if $A \in \mathscr{F}_{T\wedge k+s}$, then

$$\mathbb{E}\left[e^{\mathrm{i}\xi\cdot(X_{T\wedge n+t} - X_{T\wedge n+s})}1_A\right] = e^{-\frac{1}{2}(t-s)|\xi|^2}\mathbb{P}(A) \quad \text{for any } n \geq k.$$

Letting $n \to \infty$, and using Dominated Convergence Theorem we deduce that

$$\mathbb{E}\left[e^{\mathrm{i}\xi\cdot(X_{T+t} - X_{T+s})}1_A\right] = e^{-\frac{1}{2}(t-s)|\xi|^2}\mathbb{P}(A).$$

Since $\mathscr{F}_{T+s} = \sigma\{\mathscr{F}_{T\wedge k+s} : k \geq 1\}$, we thus conclude that

$$\mathbb{E}\left[e^{\mathrm{i}\xi\cdot(X_{T+t} - X_{T+s})}|\mathscr{F}_{T+s}\right] = e^{-\frac{1}{2}(t-s)|\xi|^2} \quad \text{for any } t > s \geq 0.$$

Therefore $Y_t = X_{T+t} - X_T$ (for $t \geq 0$) is an $(\mathscr{F}_{T+t})$-Brownian motion started at zero. □

[41]This property is one version of the so-called strong Markov property of Brownian motion. That is, a Brownian motion begins afresh at stopping times, first formulated by G. Hunt: Some theorems concerning Brownian motion, *Trans. Amer. Math. Soc.* 81, 294-319 (1956). Also E. B. Dynkin: *Markov Processes*. Springer, Berlin (1965). For a modern treatment in a rather general setting, see K. L. Chung: *Lectures from Markov Processes to Brownian Motion*. Springer-Verlag (1982), Section 2.3.

In particular, if $X = (X_t)_{t\geq 0}$ is a d-dimensional Brownian motion and T is a stopping time with $\mathbb{P}[T < \infty] > 0$, then $X_{T+t} - X_T$ has a normal distribution $N(0,t)$ under the conditional probability P_T. That is, the conditional probability[42]

$$\mathbb{P}[X_{T+t} - X_T \in \mathrm{d}x | T < \infty] = \frac{1}{(2\pi t)^{d/2}} \exp\left(-\frac{|x|^2}{2t}\right) \mathrm{d}x$$

for every $t > 0$.

The following result, the reflection principle, was known at the end of the 19th Century for random walk and appears in the famous 1900 thesis of Bachelier, which introduced the idea of modelling stock prices using Brownian motion (although since he had no formulation of the strong Markov property, his proof is not rigorous).

Theorem 4.12 (The reflection principle)**.** *Let B be a Brownian motion and τ a stopping time with $\mathbb{P}[\tau < \infty] = 1$. Then the process $\tilde{B}$ defined by*

$$\tilde{B}_t = \begin{cases} B_t & t < \tau, \\ 2B_\tau - B_t & t \geq \tau. \end{cases}$$

is a standard Brownian motion.

Proof. By definition $\tilde{B}$ is a Brownian motion up to the stopping time τ. For $t > \tau$ we write $t = \tau + t'$, and let $\bar{B}_{t'} = B_{\tau+t'} - B_\tau$, which is a Brownian motion independent of (τ, B_τ) by the strong Markov property. Using this and the symmetry of Brownian motion, so that $\bar{B} = -\bar{B}$ in distribution, for $t > \tau$ we have

$$\begin{aligned} B_t &= B_{\tau+t'} - B_\tau + B_\tau \\ &= \bar{B}_{t'} + B_\tau \\ &= -\bar{B}_{t'} + B_\tau \text{ (in distribution by symmetry)} \\ &= 2B_\tau - B_t = \tilde{B}_t. \end{aligned}$$

Thus $\tilde{B}$ has the law of Brownian motion as required. □

Theorem 4.13. *Let $B = (B_t)_{t\geq 0}$ be a real standard Brownian motion, $B_0 = 0$, and $S_t = \sup_{s\leq t} B_s$ (for $t \geq 0$) be running maximum process. Then*[43]

$$\mathbb{P}[S_t \geq b, B_t \leq a] = \frac{1}{\sqrt{2\pi t}} \int_{2b-a}^{\infty} e^{-\frac{x^2}{2t}} dx \tag{14}$$

for every $t > 0$, $b > 0$ and $b \geq a$.

[42] In particular, if T is a stopping time which is finite almost surely, then $X_{T+t} - X_{T+s}$ (where $t > s \geq 0$) has a normal distribution $N(0,(t-s)I)$. This fact is still very surprising indeed, if you think about the definition of the random variable $X_{T+t} - X_{T+s}$.

[43] This gives, for any but fixed $t > 0$, the joint distribution of (B_t, S_t). In fact by taking derivatives in a and in b, we may conclude that the PDF of the joint law of (B_t, S_t) is given by

$$f_{(B_t,S_t)}(a,b) = \frac{2(2b-a)}{\sqrt{2\pi} t^{3/2}} \exp\left(-\frac{(2b-a)^2}{2t}\right)$$

on the region $\{(a,b) : b > 0, b \geq a\}$ of $\mathbb{R}^2$.

Proof. Let $\tau_b = \inf\{t > 0 : B_t = b\}$. For $b > 0$ and $t > 0$, $\{S_t \geq b\} = \{\tau_b \leq t\}$, so that

$$\begin{aligned}\mathbb{P}[S_t \geq b, B_t \leq a] &= \mathbb{P}[\tau_b \leq t, B_t \leq a] \\ &= \mathbb{P}[\tau_b \leq t, 2B_{\tau_b} - B_t \leq a] \\ &= \mathbb{P}[\tau_b \leq t, 2b - B_t \leq a] \\ &= \mathbb{P}[\tau_b \leq t, B_t \geq 2b - a] \\ &= \mathbb{P}[B_t \geq 2b - a]\end{aligned}$$

which leads to (14), here the last equality follows from the fact that, so $\{\tau_b \leq t\} \subseteq \{B_t \geq 2b - a\}$ as $2b - a > b$. □

Remark 4.14. *We have proved that $\mathbb{P}[S_t \geq b, B_t \leq a] = \mathbb{P}[B_t \geq 2b - a]$ for $b > 0$ and $b \geq a$. For the last assertion of the theorem, taking $b = a$, observe that*

$$\begin{aligned}\mathbb{P}[S_t \geq a] &= \mathbb{P}[S_t \geq a, B_t \geq a] + \mathbb{P}[S_t \geq a, B_t \leq a] \\ &= 2\mathbb{P}[B_t \geq a] = \mathbb{P}[B_t \geq a] + \mathbb{P}[B_t \leq -a] \quad \textit{(symmetry)} \\ &= \mathbb{P}[|B_t| \geq a]\end{aligned}$$

for every $a > 0$, which implies, for every fixed time t, S_t and $|B_t|$ have the same distribution[44] .

Together with Doob's maximal inequality, we derive a useful bound for Brownian motion.

Proposition 4.15. *Let $(B_t)_{t \geq 0}$ be (one dimensional) Brownian motion with $B_0 = 0$. Then*[45]

$$\mathbb{P}\left[\sup_{s \in [0,t]} B_s \geq \lambda\right] = 2\int_{\lambda/\sqrt{t}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \leq e^{-\lambda^2/(2t)}$$

for every[46] *$t > 0$ and $\lambda > 0$.*

---

[44] Note here, for every fixed t, S_t and $|B_t|$ has the same distribution. While on the other hand $t \mapsto S_t$ is a non-decreasing process, but $t \mapsto |B_t|$ is not, so as processes $\{S_t : t \geq 0\}$ and $\{|B_t| : t \geq 0\}$ have different distributions, but they have exactly the same *one dimensional marginal* distribution. In Paper C8.1 we shall show that $\{|B_t| : t \geq 0\}$ is a continuous semimartingale.

[45] This gives the distribution of $T_b = \tau_b$ for any $b > 0$, namely

$$\mathbb{P}[T_b < t] = \mathbb{P}[S_t > b] = \mathbb{P}[|B_t| \geq b] = 2\int_{b/\sqrt{t}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

Which in particular implies that $\mathbb{P}[T_b < \infty] = 2\int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$. By differentiating in t, it follows that the distribution of T_b (when $|b| \neq 0$, so $T_b = \tau_b$) has a PDF given by

$$f_{T_b}(t) = \frac{|b|}{\sqrt{2\pi} t^{3/2}} \exp\left(-\frac{b^2}{2t}\right) \quad \text{for } t > 0.$$

Using this formula one can check T_b is not integrable if $b \neq 0$.

[46] It is worthy of pointing out that the tail bound $e^{-\lambda^2/(2t)}$ may be written as $\exp\left(-\frac{\lambda^2}{2\sup_s \mathrm{var}(B_s)}\right)$. This version of tail estimates (or concentration) has a Gaussian extension, under the name of Borell's inequality.

Proof. In fact

$$\mathbb{P}[S_t \geq \lambda] = \mathbb{P}[|B_t| \geq \lambda] = \frac{2}{\sqrt{2\pi t}} \int_{\lambda}^{\infty} e^{-\frac{x^2}{2t}} dx$$
$$= \frac{2}{\sqrt{2\pi}} \int_{\lambda/\sqrt{t}}^{\infty} e^{-\frac{x^2}{2}} dx$$

This estimate can be established by using the fact that $M_t = e^{\alpha B_t - \alpha^2 t/2}$, $t \geq 0$, is a non-negative martingale. It follows that, for $\alpha \geq 0$, using Doob's maximal inequality,

$$\begin{aligned} \mathbb{P}[S_t \geq \lambda t] &= \mathbb{P}\Big[\sup_{u \leq t} \big(e^{\alpha B_u - \alpha^2 t/2}\big) \geq e^{\alpha\lambda t - \alpha^2 t/2}\Big] \\ &\leq \mathbb{P}\Big[\sup_{u \leq t} \big(e^{\alpha B_u - \alpha^2 u/2}\big) \geq e^{\alpha\lambda t - \alpha^2 t/2}\Big] \\ &\leq e^{-\alpha\lambda t + \alpha^2 t/2} \underbrace{\mathbb{E}\left[e^{\alpha B_t - \alpha^2 t/2}\right]}_{=1}. \end{aligned}$$

The bound now follows since $\min_{\alpha \geq 0} e^{-\alpha\lambda t + \alpha^2 t/2} = e^{-\lambda^2/(2t)}$ (with the minimum achieved when $\alpha = \lambda$). □

Corollary 4.16. *Let B be a standard real-valued Brownian motion, $B_0 = 0$ and $S_t = \sup_{s \leq t} B_s$ for $t \geq 0$.*

1) Then $S_t > 0$ for $t > 0$ (of course, almost surely).

*2) For every $a \in \mathbb{R}$, let $T_a := \inf\{t \geq 0 : B_t = a\}$ (*the passage time*), with the convention that* $\inf \emptyset = \infty$*. Then for each $a \in \mathbb{R}$, $T_a < \infty$. Consequently, we have a.s.*

$$\limsup_{t\to\infty} B_t = +\infty, \quad \liminf_{t\to\infty} B_t = -\infty.$$

Proof. 1) $\mathbb{P}[S_t > 0] = \mathbb{P}[|B_t| > 0] = 1$ for every $t > 0$.

2) If $a = 0$, then $T_0 = 0$. If $a \neq 0$, then $T_a = \tau_a$, so $T_a < \infty$ almost surely. Without using $\tau_a = \inf\{t > 0 : B_t = a\}$ one may argue as the following. Write

$$1 = \mathbb{P}\Big[\sup_{0 \leq s \leq 1} B_s > 0\Big] = \lim_{\delta \downarrow 0} \mathbb{P}\Big[\sup_{0 \leq s \leq 1} B_s > \delta\Big].$$

Now, writing $c = 1/\delta$ in the Brownian scaling of Proposition 4.19 ii, we have that $B_t^{\delta} = \delta^{-1} B_{t\delta^2}$ is a Brownian motion. Thus for any $\delta > 0$,

$$\mathbb{P}\Big[\sup_{0 \leq s \leq 1} B_s > \delta\Big] = \mathbb{P}\Big[\sup_{0 \leq s \leq 1/\delta^2} B_s^{\delta} > 1\Big] = \mathbb{P}\Big[\sup_{0 \leq s \leq 1/\delta^2} B_s > 1\Big]. \tag{15}$$

If we let $\delta \downarrow 0$, we find

$$\mathbb{P}[\sup_{s \geq 0} B_s > 1] = \lim_{\delta \downarrow 0} \mathbb{P}\Big[\sup_{0 \leq s \leq 1/\delta^2} B_s > 1\Big] = \lim_{\delta \downarrow 0} \mathbb{P}[\sup_{0 \leq s \leq 1} B_s > \delta] = 1.$$

Another scaling argument shows that, for every $M > 0$,

$$\mathbb{P}[\sup_{s\geq 0} B_s > M] = 1$$

and replacing B with $-B$,

$$\mathbb{P}[\inf_{s\geq 0} B_s < -M] = 1.$$

Continuity of sample paths completes the proof of 2). □

Corollary 4.17. *The map $t \mapsto B_t$ is a.s. not monotone on any non-trivial interval.*

4.1.2 Finite dimensional distributions

It is important to be able to do "computations" with Brownian motion, at least, for computing quantities associated with Brownian motion at finite many times. To this end, we should introduce the *transition probability function*[47] $p(t,x,y)$, which is nothing but the Gaussian density with mean x and variance t. More precisely[48]

$$p(t,x,y) = p(t,y-x) := \frac{1}{(2\pi t)^{d/2}} \exp\left(-\frac{|y-x|^2}{2t}\right) \quad \text{for } t > 0, x,y \in \mathbb{R}^d.$$

1) For any fixed (t,x), $p(t,x,y)$ is a probability density function on $\mathbb{R}^d$.

2) The Chapman-Kolmorgorov equation[49] holds for $p(t,x,y)$

$$p(t+s,x,y) = \int_{\mathbb{R}^d} p(t,x,z)p(s,z,y)\mathrm{d}z$$

for any $s > 0$, $t > 0$ and $x,y \in \mathbb{R}^d$.

Let $B = (B^1, \cdots, B^d)$ be a Brownian motion of d-dimensions, with initial distribution μ, that is $\mu(A) = \mathbb{P}[B_0 \in A]$ for any $A \in \mathscr{B}(\mathbb{R}^d)$. Let $D : 0 = t_0 < t_1 < t_2 < \ldots < t_n$ be a finite partition of $[0,\infty)$. Then $X_0 := B_{t_0}, X_i := B_{t_i - t_{i-1}}$ (where $i = 1, \cdots, n$) are independent, X_0 has has distribution μ, and X_i has a normal distribution $N(0,(t_i - t_{i-1})I)$, and therefore the joint distribution of $(X_0, \cdots, X_n)$ is a

---

[47]The name is reserved for processes with Markov property, like Markov chains, here $p(t,x,y)$ really plays a role as the one step transition probability matrix. Here we are dealing with processes in continuous-time with "continuous" state space.

[48]For any fixed $t > 0$, $x \in \mathbb{R}^d$, $y \mapsto p(t,x,y)$ is a probability density function (PDF), and, for any fixed y, as a function of the pair $x \in \mathbb{R}^d, t > 0$, $p(t,x,y)$ is a solution to the heat equation $\frac{\partial}{\partial t} p = \frac{1}{2}\Delta p$, where $\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$ is the Laplace operator on $\mathbb{R}^d$. This is the reason why Brownian motion is called a diffusion process with infinitesimal generator $\frac{1}{2}\Delta$. This builds an important connection between diffusion processes, parabolic differential equations, harmonic analysis and etc. This connection will be explored further in Paper C8.2.

[49]This can be verified as an easy exercise.

probability measure on $\mathbb{R}^d \times \cdots \times \mathbb{R}^d$ (the product space of $d+1$ copies of $\mathbb{R}^d$) with its Borel σ-algebra, given by

$$\mathbb{P}\left[(X_0, \cdots, X_n) \in G\right] = \int \cdots \int_G \mu(dx_0) \prod_{i=1}^{n} p(t_i - t_{i-1}, 0, x_i) \mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

By change variable formula we thus obtain

$$\mathbb{P}\left[(B_{t_0}, \cdots, B_{t_n}) \in G\right] = \int \cdots \int_G \mu(dx_0) \prod_{i=1}^{n} p(t_i - t_{i-1}, x_{i-1}, x_i) \mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

That is, the marginal distribution of $(B_{t_0}, \cdots, B_{t_n})$, a probability measure P_D on $\mathbb{R}^d \times \cdots \times \mathbb{R}^d$, is given by

$$P_D(dx_0, \cdots, dx_n) = \mu(\mathrm{d}x_0) \prod_{i=1}^{n} p(t_i - t_{i-1}, x_{i-1}, x_i) \mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

The collection $\{P_D : D = \{t_i\}\}$ defined as above is called the family of finite dimensional marginal distributions of Brownian motion with initial distribution μ.

The following properties are easy to check via the previous formula.

Lemma 4.18. *The family $\{P_D : D = \{t_i\}\}$ is consistent. That is, if $D : 0 = t_0 < \cdots < t_n$ and*

$$D' : 0 = t_0 < \cdots t_k < s < t_{k+1} < \cdots < t_n.$$

Then

$$P_D(dx_0, \cdots, dx_n) = \int_{\mathbb{R}^d} P_{D'}(dx_0, \cdots, dx_k, dx, dx_{k+1}, \cdots, dx_n)$$

where the integral is carried out against dx.

In particular, if B is a Brownian motion started from $\xi \in \mathbb{R}^d$, i.e. $\mu(\mathrm{d}x) = \delta_\xi(\mathrm{d}x)$, then the joint distribution of $(B_{t_0}, \cdots, B_{t_n})$

$$\mathbb{P}\left[(B_{t_0}, \cdots, B_{t_n}) \in G\right] = \int \cdots \int_G \delta_\xi(\mathrm{d}x_0) \prod_{i=1}^{n} p(t_i - t_{i-1}, x_{i-1}, x_i) \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

for any $0 = t_0 < t_1 < \cdots < t_n$ and for any Borel measurable G, and therefore $(B_{t_0}, \cdots, B_{t_n})$ has a normal distribution[50].

We say a stochastic process $X = (X_t)_{t \geq 0}$ is a Gaussian process, if for any $0 \leq t_1 < \cdots < t_n$, $(X_{t_1}, \cdots, X_{t_n})$ has a normal distribution. Since a normal distribution is determined uniquely by its mean and its co-variance, therefore, the distribution of a Gaussian process is determined uniquely by its mean function $m(t) = \mathbb{E}(X_t)$ and its co-variance function $C(s,t) = \mathbb{E}\left[(X_t - m(t))(X_s - m(s))\right]$ for $s, t \in [0, \infty)$. As a consequence, an $\mathbb{R}^d$-valued stochastic process $X = (X_t)_{t \geq 0}$ is a

[50] we shall consider $\delta_a(\mathrm{d}x)$ as a normal distribution $N(a, 0)$, the degenerate Gaussian distribution.

Gaussian process if and only if for any finite many $t_i \in [0,\infty)$, and for any numbers c_i, a linear combination $\sum_i c_i X_{t_i}$ has a normal distribution[51][52].

If $B = (B_t)_{t\geq 0}$ is a real Brownian motion started from ξ, then $m(t) = \xi$ and

$$\begin{aligned}\mathbb{E}\left[(B_s - B_0)(B_t - B_0)\right] &= \mathbb{E}\left[(B_s - B_0)(B_t - B_s + B_s - B_0)\right] \\ &= \mathbb{E}\left[(B_s - B_0)^2\right] = s\end{aligned}$$

for $t \geq s \geq 0$. Hence the co-variance function of a real Brownian motion is $C(s,t) = s \wedge t$. Therefore we have the following characterization of Brownian motion as a continuous Gaussian process: a *continuous* (real) stochastic process is a standard Brownian motion, if and only if $B_0 = 0$, B is a Gaussian process with mean zero and the co-variance function $\mathbb{E}[B_s B_t] = s \wedge t$. As a consequence, we have the following corollary.

Proposition 4.19. *Let B be a real-valued Brownian motion started from 0. Then*

i. $-B_t$ *is also a Brownian motion,* *(symmetry)*

ii. $\forall c \geq 0$, cB_{t/c^2} *is a Brownian motion,* *(scaling)*

iii. $X_0 = 0$, $X_t := tB_{\frac{1}{t}}$ *is a Brownian motion,* *(time inversion)*

iv. for $t \in [0,1]$, $X_t := B_1 - B_{1-t}$ *is a Brownian motion,* *(time reversal)*

v. $\forall s \geq 0$, $\tilde{B}_t = B_{t+s} - B_s$ *is a Brownian motion independent of* $\sigma(B_u : u \leq s)$, *(simple Markov property).*

The proof is an exercise.

By using the finite dimensional marginal distributions, it is easy to formulate the simple Markov property. We shall state this important feature of Brownian motion in two forms.

Proposition 4.20. *Let $B = (B_t)_{t\geq 0}$ be an $(\mathscr{F}_t)_{t\geq 0}$ Brownian motion (where the filtration $(\mathscr{F}_t)$ is right continuous). Then B possesses the (strong) Markov property*[53]

$$\mathbb{E}\left[f(B_{t+S})|\mathscr{F}_S\right]1_{\{S<\infty\}} = \mathbb{E}\left[f(B_{t+S})|B_S\right]1_{\{S<\infty\}} = \int_{\mathbb{R}^d} f(y)p(t,B_S,y)1_{\{S<\infty\}}dy$$

[51] A random vector $Y := (X_{t_1}, \cdots, X_{t_n})$ has a normal distribution, if and only if by definition its characteristic function $\mathbb{E}\left[e^{i\xi\cdot Y}\right] = \exp\left(ia\cdot\xi - \frac{1}{2}\xi\cdot\Sigma\xi\right)$, where a is the mean vector of Y and Σ is the co-variance matrix of Y. Conversely, suppose for any $\xi = (\xi_i)$ the linear combination $Z = \sum_i \xi_i X_{t_i}$ has a normal distribution. Then since $\mathbb{E}[Z] = \xi\cdot a$ and the variance of Z is $\sum_{i,j}\xi_i\xi_j\sigma_{ij} = \xi\cdot\Sigma\xi$ (where σ_{ij} is the co-variance of X_{t_i} and X_{t_j}), so that $\mathbb{E}\left[e^{i\sum_j \xi_j X_{t_j}}\right] = \exp\left(ia\cdot\xi - \frac{1}{2}\xi\cdot\Sigma\xi\right)$. Since $\xi_1, \cdots, \xi_n$ are arbitrary, so that the joint distribution of $(X_{t_1}, \cdots, X_{t_n})$ is a normal distribution.

[52] This can be generalized to introduce Gaussian fields, namely, a family of random variables $\{X_t; t\in\Lambda\}$ (where Λ is an index set, a subset of $\mathbb{R}^n$ for example) is called a Gaussian field, if for any finite subset $\{t_i\}$ of Λ, the joint distribution of (X_{t_i}) is a normal distribution (so determined by its mean vector $\mathbb{E}[X_{t_i}]$ and its co-variance matrix $\mathrm{var}(X_{t_i}, X_{t_j})$. Hence, $\{X_t; t\in\Lambda\}$ is a Gaussian field if and only if any finite linear combination $\sum_i c_i X_{t_i}$ has a normal distribution.

[53] Therefore, it is convenient to introduce a family $\{P_t; t>0\}$, where P_t is a linear operator defined by $P_t f(x) = \int_{\mathbb{R}^d} f(y)p(t,x,y)dy$, which sends a function f on $\mathbb{R}^d$ (as long as the integral exists for

for any $t > 0$, any bounded Borel measurable function f and any stopping time S.

Proof. We may assume that $\mathbb{P}[S < \infty] = 1$, as $\{S < \infty\}$ is $\mathscr{F}_S$-measurable. By Theorem 4.11, $\{B_{t+S} - B_S; t \geq 0\}$ is an $(\mathscr{F}_{S+t})_{t\geq 0}$ standard Brownian motion independent of $\mathscr{F}_S$. Hence, for $t > 0$, $B_{t+S} - B_S$ and $\mathscr{F}_S$ are independent, and $B_{t+S} - B_S$ has a normal distribution $N(0, tI)$, so that

$$\begin{aligned}\mathbb{E}[f(B_{t+S} - B_S + B_S)|\mathscr{F}_S] &= \mathbb{E}\left[\int_{\mathbb{R}^d} p(t,0,y)f(y+B_S)\mathrm{d}y|\mathscr{F}_S\right] \\ &= \int_{\mathbb{R}^d} p(t,0,y)f(y+B_S)\mathrm{d}y \\ &= \int_{\mathbb{R}^d} p(t,B_S,y)f(y)\mathrm{d}y\end{aligned}$$

In particular this implies that $\mathbb{E}[f(B_{t+S})|\mathscr{F}_S] = \mathbb{E}[f(B_{t+S})|B_S]$, which is also called the strong Markov property. □

The another form of the Markov property, which can be verified by using the Chapman-Kolmogorove equation for $p(t,x,y)$.

Proposition 4.21. *Let $B = (B_t)_{t\geq 0}$ be an $(\mathscr{F}_t)_{t\geq 0}$ Brownian motion. Let $\mathscr{F}^0_{\geq t} = \sigma(B_s : s \geq t)$. Then for any bounded $\mathscr{F}^0_{\geq s}$ measurable random variable H, $\mathbb{E}[H|\mathscr{F}_s] = \mathbb{E}[H|B_s]$.*

Proof. For any partition $s \leq t_1 < \cdots < t_m$,

$$\mathbb{E}[f_1(B_{t_1})\cdots f_m(B_{t_m})|\mathscr{F}_s] = \mathbb{E}[f_1(B_{t_1})\cdots f_m(B_{t_m})|B_s]$$

for any bounded Borel measurable functions f_i. This can be checked by induction on m. □

Although the sample paths of Brownian motion are continuous, it does not mean that they are nice in any other sense. In fact the behaviour of Brownian motion is unlike the usual functions one encounters. Here are just a few of its strange behavioural traits.

i. Although $(B_t)_{t\geq 0}$ is continuous everywhere, it is (with probability one) differentiable nowhere.

ii. Brownian motion will eventually hit any and every real value no matter how large, or how negative. No matter how far above the axis, it will (with probability one) be back down to zero at some later time.

---

defining $P_t f(x)$) to a function $P_t f$. The Markov property may be stated as $\mathbb{E}[f(B_t)|\mathscr{F}_s] = P_{t-s}f(B_s)$ for any $t > s \geq 0$. Due to the Chapman-Kolmogorov's equation, $\{P_t; t > 0\}$ forms a semi-group in the sense that $P_t \circ P_s = P_{t+s}$ for any $t, s \geq 0$ (here we agree that $P_0 = I$ the identity operator). $\{P_t; t \geq 0\}$ is called the heat sim-group on $\mathbb{R}^d$. For any f (bounded, Borel measurable for example), $(x,t) \mapsto P_t f(x)$ solves the heat equation, namely $\frac{\partial}{\partial t}P_t f = \frac{1}{2}\Delta P_t f$ on $\mathbb{R}^d \times (0,\infty)$.

iii. Once Brownian motion hits a value, it immediately hits it again (uncountably!) *infinitely* often, and then again from time to time in the future.

iv. It doesn't matter what scale you examine Brownian motion on, it looks just the same. The paths of Brownian motion are fractals almost surely.

Theorem 4.22 (Blumenthal's 0-1 law)**.** *Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion on $(\Omega, \mathscr{F}, \mathbb{P})$, and $\mathscr{F}_t^0 = \sigma\{B_s : s \leq t\}$ for $t \geq 0$. Then for every $A \in \mathscr{F}_{0+}^0$, $\mathbb{P}[A] = 0$ or 1.*

Proof. Let $0 < t_1 < t_2 \cdots < t_k$ and let $g : \mathbb{R}^k \mapsto \mathbb{R}$ be a bounded continuous function. Also, fix $A \in \mathscr{F}_{0+}^0$. Then by continuity and dominated convergence

$$\mathbb{E}[1_A g(B_{t_1}, \dots, B_{t_k})] = \lim_{\varepsilon \downarrow 0} \mathbb{E}[1_A g(B_{t_1} - B_\varepsilon, \dots, B_{t_k} - B_\varepsilon)].$$

If $0 < \varepsilon < t_1$, the variables $B_{t_1} - B_\varepsilon, \dots, B_{t_k} - B_\varepsilon$ are independent of $\mathscr{F}_\varepsilon^0$ and thus also of $\mathscr{F}_{0+}$. It follows that

$$\begin{aligned}
\mathbb{E}[1_A g(B_{t_1}, \dots, B_{t_k})] &= \lim_{\varepsilon \downarrow 0} \mathbb{E}[1_A g(B_{t_1} - B_\varepsilon, \dots, B_{t_k} - B_\varepsilon)] \\
&= \mathbb{P}[A]\mathbb{E}[g(B_{t_1}, \dots, B_{t_k})].
\end{aligned}$$

We have thus obtained that $\mathscr{F}_{0+}^0$ is independent of $\sigma(B_{t_1}, \dots, B_{t_k})$. Since this holds for any finite collection $\{t_1, \dots, t_k\}$ of (strictly) positive reals, $\mathscr{F}_{0+}^0$ is independent of $\sigma(B_t, t > 0)$. However, $\sigma(B_t, t > 0) = \sigma(B_t, t \geq 0)$, since B_0 is the point-wise limit of B_t when $t \to 0$. Since $\mathscr{F}_{0+} \subset \sigma(B_t, t \geq 0)$, we conclude that $\mathscr{F}_{0+}$ is independent of itself and so must be trivial. □

4.2 Construction of Brownian motion

4.2.1 Constructing distributions on path spaces

For definiteness, we take real-valued processes, so $E = \mathbb{R}^d$ in this section.

Let $X = (X_t)_{t\geq 0}$ be a process taking values in E. Indistinguishability takes the *sample path* as the basic object of study, so that we could think of $(X_t(\omega), t \geq 0)$ as a path in E as a random variable taking values in the space $E^{[0,\infty)}$ of all possible paths, i.e. the space of all mappings from $[0, \infty)$ into E. This sample space $E^{[0,\infty)}$, the path space, then has to be endowed with a σ-algebra of measurable sets.

Definition 4.23. *An n-dimensional* cylinder set *in $E^{[0,\infty)}$ is a set of the form*

$$C = \left\{\omega \in E^{[0,\infty)} : \left(\omega(t_1), \dots, \omega(t_n)\right) \in A\right\}$$

for some $0 \leq t_1 < \cdots < t_n$ and $A \in \mathscr{B}(\mathbb{R}^{d\times n})$.

Let $\mathscr{C}$ be the family of all finite-dimensional cylinder sets and $\mathscr{B}(E^{[0,\infty)})$ the σ-algebra it generates. This is small enough to be able to build probability measures on $\mathscr{B}(E^{[0,\infty)})$ using Carathéodory's Theorem (see B8.1). On the other hand

$\mathscr{B}(E^{[0,\infty)})$ only contains events which can be defined using at most countably many coordinates. In particular, the set

$$\left\{\omega \in E^{[0,\infty)} : \omega(t) \text{ is continuous}\right\}$$

is not $\mathscr{B}(E^{[0,\infty)})$-measurable.

We will have to do some work to show that many processes can be assumed to be continuous, or right continuous. The sample paths are then fully described by their values at times $t \in \mathbb{Q}$, which will greatly simplify the study of quantities of interest such as $\sup_{0\leq s\leq t}|X_s|$ or $\tau_0(\omega) = \inf\{t \geq 0 : X_t(\omega) > 0\}$.

A monotone class argument will tell us that a probability measure on $\mathscr{B}(E^{[0,\infty)})$ is characterised by its finite-dimensional distributions – so if we can take continuous paths, then we only need to find the probabilities of cylinder sets to characterise the distribution of the process.

In this section, we are going to provide a very general result about constructing continuous time stochastic processes and a criterion due to Kolmogorov which gives conditions under which there will be a version of the process with continuous paths.

Let $\mathscr{T}$ be the set of finite increasing sequences of non-negative numbers, i.e. $D \in \mathscr{T}$ if and only if $D = (t_1, t_2, \dots, t_n)$ for some n and $0 = 0 < t_1 < t_2 < \ldots < t_n$.

Suppose that for each $D \in \mathscr{T}$ of length n we have a probability measure P_D on $(E^n, \mathscr{B}(E^n))$. The collection $\{P_D : D \in \mathscr{T}\}$ is called a family of finite-dimensional (marginal) distributions.

Definition 4.24. *A family $\{P_D : D \in \mathscr{T}\}$ of finite dimensional distributions is called* consistent *if for any $D = (t_1, t_2, \dots, t_n) \in \mathscr{T}$ and $1 \leq j \leq n$*

$$\begin{aligned} P_D(A_1 \times A_2 \times \ldots \times A_{j-1} \times E \times A_{j+1} \times \ldots \times A_n) \\ = P_{D'}(A_1 \times A_2 \times \ldots \times A_{j-1} \times A_{j+1} \times \ldots \times A_n) \end{aligned}$$

where $A_i \in \mathscr{B}(E)$ and $D' := (t_1, t_2, \dots, t_{j-1}, t_{j+1}, \dots, t_n)$.

(In other words, if we integrate out over the distribution at the jth time point then we recover the corresponding marginal for the remaining lower dimensional vector.)

If we have a probability measure P on $\left(E^{[0,\infty)}, \mathscr{B}(E^{[0,\infty)})\right)$ then it defines a consistent family of marginals via

$$P_D(A) = P(\{\omega \in E^{[0,\infty)} : (\omega(t_1), \dots, \omega(t_n)) \in A\})$$

where $D = (t_1, t_2, \dots, t_n)$, $A \in \mathscr{B}(E^n)$, and we note that the set in question is in $\mathscr{B}(E^{[0,\infty)})$ as it depends on finitely many coordinates. But we would like to have a converse – if I give you P_D, does there exist a corresponding measure P?

Theorem 4.25 (Daniell–Kolmogorov Extension Theorem)**.** *Let $\{P_D : D \in \mathscr{T}\}$ be a consistent family of finite-dimensional distributions. Then there exists a probability* *measure P on $\left(E^{[0,\infty)}, \mathscr{B}(E^{[0,\infty)})\right)$ such that for any n, $D = (t_1, \ldots, t_n) \in \mathscr{T}$ and $A \in \mathscr{B}(E^n)$,*

$$P_D(A) = P\left[\left\{\omega \in E^{[0,\infty)} : (\omega(t_1), \ldots, \omega(t_n)) \in A\right\}\right]. \tag{16}$$

We will not prove this here, but notice that (16) defines P on the cylinder sets and so if we can establish countable additivity then the proof reduces to an application of Carathéodory's extension theorem [Paper A4, Paper B8.1]. Uniqueness is a consequence of the Monotone Class Lemma.

This is a remarkably general result, but it doesn't allow us to say anything meaningful about the paths of the process. For that we appeal to Kolmogorov's criterion.

Theorem 4.26 (Kolmogorov–Čentsov continuity criterion)**.** *Suppose that a stochastic process $(X_t : t \leq T)$ defined on $(\Omega, \mathscr{F}, \mathbb{P})$ satisfies*

$$\mathbb{E}[|X_t - X_s|^\alpha] \leq C|t-s|^{1+\beta}, \quad 0 \leq s,t \leq T \tag{17}$$

for some strictly positive constants α, β and C. Then there exists $\tilde{X}$, a modification of X, whose paths are γ-locally Hölder continuous $\forall \gamma \in (0, \beta/\alpha)$ a.s., i.e.

$$\sup_{s,t \in [0,T], s \neq t} \frac{|\tilde{X}_t - \tilde{X}_s|}{|t-s|^\gamma} < \infty \quad a.s. \tag{18}$$

In particular, the sample paths of $\tilde{X}$ are a.s. continuous (and uniformly continuous on $[0,T]$).

Remark 4.27. *Many more results and conditions in this direction are possible. See for example Cramér and Leadbetter,* Stationary and Related Stochastic Processes*, Wiley, 1967.*

Exercise 4.28. *Use the Kolmogorov continuity criterion to show that Brownian motion admits a modification which is locally Hölder continuous of order γ for any $0 < \gamma < 1/2$.*

4.2.2 Constructing Brownian motion on path spaces

We are now in a position to demonstrate the existence (construction) of Brownian motion.

We shall take $\Omega = E^{[0,\infty)}$ (where $E = \mathbb{R}^d$ with its Borel σ-algebra) as the sample space, with the σ-algebra $\mathscr{F}$ generated by the cylinder sets. Let μ be a probability measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$, and define the family $\{P_D : D \in \mathscr{T}\}$ of finite dimensional distributions

$$P_D\left[X_{t_i} \in A_i : i = 0, 1, \cdots, n\right] = \int_{A_0} \cdots \int_{A_n} \mu(dx_0) p_D(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

if $D : 0 = t_0 < t_1 < \cdots < t_n$, for any $A_0, \cdots, A_n \in \mathscr{B}(\mathbb{R}^d)$. Here $\{X_t : t \geq 0\}$ is the coordinate process on the path space Ω. Then clearly $\{P_D : D \in \mathscr{T}\}$ is consist, by Daniell–Kolmogorov's Theorem, there is a unique probability measure, denoted by $\mathbb{P}^\mu$, on $(\Omega, \mathscr{F})$, such that

$$\mathbb{P}^\mu [X_{t_i} \in A_i : i = 0, 1, \cdots, n] = P_D [X_{t_i} \in A_i : i = 0, 1, \cdots, n]$$

for any $D : 0 = t_0 < t_1 < \cdots < t_n$ and any $A_i \in \mathscr{B}(\mathbb{R}^d)$. This is equivalent to say that, for any non-negative or bounded measurable f and for any $0 = t_0 < t_1 < \cdots < t_n$ we have

$$\mathbb{P}^\mu [f(X_{t_0}, X_{t_1}, \cdots, X_{t_n})] = \int_{A_0} \cdots \int_{A_n} f(x_0, \cdots, x_n) \mu(dx_0) p_{\mathbf{t}}(x_1, \ldots, x_n) dx_1 \cdots dx_n.$$

In particular, for any $t > s \geq 0$

$$\begin{aligned} \mathbb{P}^\mu \left[|X_t - X_s|^4 \right] &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |y - x|^4 \mu(dx_0) p(s, x_0, x) p(t - s, x, y) dx dy \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |y|^4 \mu(dx_0) p(s, x_0, x) p(t - s, x, y + x) dx dy \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |y|^4 \mu(dx_0) p(t - s, 0, y) dy \\ &= C|t - s|^2 \end{aligned}$$

where

$$C = \int_{\mathbb{R}^d} |y|^4 p(1, 0, y) dy$$

is a constant depending only on d. According to the Kolmogorov continuity criterion, $t \mapsto X_t$ is continuous almost surely. Therefore $\mathbb{P}^\mu$ is supported on the space of continuous paths, $C([0,\infty), \mathbb{R}^d)$, that is, the restriction of $\mathbb{P}^\mu$ on $C([0,\infty), \mathbb{R}^d)$ is indeed a probability measure, denoted still by $\mathbb{P}^\mu$.

By construction, the coordinate process $X = \{X_t : t \geq 0\}$ on the continuous path space $C([0,\infty), \mathbb{R}^d)$ under the probability measure $\mathbb{P}^\mu$ is a d-dimensional Brownian motion with initial distribution μ. This construction gives the canonical realization of a d-dimensional Brownian motion with initial distribution μ.

We shall use $\mathbb{P}^x$ to denote $\mathbb{P}^{\delta_x}$ (where $x \in \mathbb{R}^d$) for simplicity, which is the law of d-dimensional Brownian motion started from x. Then, by Monotone Class Lemma, one can easily show that

$$\mathbb{P}^\mu(\cdot) = \int_{\mathbb{R}^d} \mathbb{P}^x(\cdot) \mu(dx).$$

The law of d-dimensional Brownian motion started from the origin 0, $\mathbb{P}^0$, is called the (d-dimensional) Wiener measure, denoted by some authors as $\mathbf{W}$.

5 Stochastic integration

At the beginning of the course we argued that whereas classically differential equations take the form

$$dX(t) = a(t, X(t))dt,$$

in many settings, the dynamics of the physical quantity in which we are interested may also have a random component and so perhaps take the form

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

We actually understand equations like this in the integral form:

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

If a is nice enough, then the first term has a classical interpretation. It is the second term, or rather a generalisation of it, that we want to make sense of now.

The first approach will be to mimic what we usually do for construction of the Lebesgue integral, namely work out how to integrate simple functions and then extend to general functions through passage to the limit. We will then provide a very slick, but not at all intuitive, approach that nonetheless gives us some 'quick wins' in proving properties of the integral.

In this chapter we shall work with a filtered complete probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ where $(\mathscr{F}_t)_{t \geq 0}$ satisfies the usual conditions, and $\mathscr{F}_\infty = \sigma\{\mathscr{F}_t : t \geq 0\}$. Recall that our original goal was to make sense of differential equations driven by 'rough' inputs. In fact, we will recast our differential equations as integral equations and so we must develop a theory that allows us to integrate with respect to 'rough' driving processes. The class of processes with which we work are called *semi-martingales*, and we shall specialise to the continuous ones.

We are going to start with functions for which the integration theory that we already know is adequate – these are called functions of finite variation.

Throughout, we assume that a filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ satisfying *the usual conditions* is given. By usual conditions we meant the followings are satisfied.

1) $(\Omega, \mathscr{F}, \mathbb{P})$ is a completed probability measure. Let $\mathscr{N}$ be the collection of all $\mathbb{P}$-null events.

2) $(\mathscr{F}_t)_{t \geq 0}$ is right-continuous, and $\mathscr{N} \subset \mathscr{F}_t$ for every $t \geq 0$.

A typical setting is the following. We are given a Brownian motion $B = (B_t)_{t \geq 0}$ on some probability space, where $(\mathscr{F}_t)_{t \geq 0}$ is the (completed) filtration generated by Brownian motion B, and $\mathscr{F} = \mathscr{F}_\infty$. For this case the filtration $(\mathscr{F}_t)_{t \geq 0}$ is continuous.

5.1 Stochastic integrals for processes of finite variation

There is in fact no new theory of stochastic integration with respect to processes with finite variations. The theory of Lebesgue's integration and Riemann-Stieltjes integration[54] may be applied for processes with finite variations. Any way, some notations and notions may be introduced.

Let $\rho : [0,\infty) \mapsto \mathbb{R}$ be a right-continuous, non-decreasing function, and $\rho(0) = \rho(0+) = \lim_{t\downarrow 0}\rho(t)$. Then there is a unique measure[55] m_ρ on $([0,\infty),\mathscr{B}([0,\infty))$ such that

$$m_\rho(\{0\}) = 0, \quad m_\rho((s,t]) = \rho(t) - \rho(s) \quad \text{for } t > s \geq 0.$$

If $w : [0,\infty) \mapsto \mathbb{R}$ is Borel measurable and w integrable with respect to m_ρ, then the (Lebesgue) integrals $\int_0^\infty w \mathrm{d}m_\rho$ and $\int_0^t w \mathrm{d}m_\rho$ (for every $t \geq 0$) is defined accordingly. These integrals shall be denoted by $\int_0^\infty w(s)\mathrm{d}\rho(s)$ and $\int_0^t w(s)\mathrm{d}\rho(s)$. In particular, if w is continuous then $\int_0^t w(s)\mathrm{d}\rho(s)$ coincides with the Riemann-Stieltjes integral and therefore, if w is continuous, then

$$\int_0^t w(s)\mathrm{d}\rho(s) = \lim_{|D|\to 0}\sum_{i\geq 1} w(t_{i-1})\left(\rho(t_i\wedge t) - \rho(t_{i-1}\wedge t)\right) \tag{19}$$

where the limit is taken over partitions $D : 0 = t_0 < t_1 < \cdots$ with $\lim_i t_i = \infty$ (i.e. $D \in \mathscr{P}$).

Similarly, if ρ itself is continuous, and if $w : [0,\infty) \mapsto \mathbb{R}$ is left-continuous on $(0,\infty)$, then the Lebesgue integral $\int_0^t w(s)\mathrm{d}\rho(s)$ exists for every $t \geq 0$ and (19) holds.

Suppose $A : [0,\infty) \mapsto \mathbb{R}$ has finite variations on any bounded intervals of $[0,\infty)$, and $V(A)$ denotes its variation, i.e.

$$V(A)(t) = \lim_{|D|\to 0}\sum_{i\geq 1}\left|A(t_i\wedge t) - A(t_{i-1}\wedge t)\right|.$$

Then both $V(A)$ and $\rho^{(A)} := V(A) - A$ are non-increasing on $[0,\infty)$. Moreover, if A is right-continuous (resp. continuous) on $[0,\infty)$ then is $V(A)$ too. If $w : [0,\infty) \mapsto \mathbb{R}$

[54] For Riemann-Stieltjes integration, you may consult Chapter 7, in T. M. Apostol: *Mathematical Analysis* (Second Edition), Addison-Wesley Pub. Company (1974).

[55] If ρ is a non-decreasing function on an interval (a,b), maybe unbounded, then right and left limits of ρ at every point $t \in (a,b)$ exist, denoted by $\rho(t+)$ and $\rho(t-)$. The functions $t \mapsto \rho(t+)$ is right continuous, and $t \mapsto \rho(t-)$ is left continuous on (a,b), and both are non-decreasing. These functions are called the right-continuous, respectiively left-continuous, modification of ρ, maybe denoted by ρ_+, resp. by ρ_-. The fundamental theorem in the theory of Lebesgue's integration says there is a unique measure m_ρ on the Borel $\mathscr{B}((a,b))$ of (a,b), satisfying that $m_{\rho_+}((s,t]) = \rho(t+) - \rho(s+)$ for any $b > t > s > a$. This measure m_{ρ_+} is called the Lebesgue-Stieltjes measure associated with ρ (or more precisely ρ_+) on (a,b). If a or/and b is finite, then this measure can be extended to be a measure on $[a,b)$ (or $[a,b]$) by assigning a mass at a (and at b) as you wish (which if courses lead to a measure on $[a,b)$ (or $[a,b]$). Here we assign the mass at 0 to be zero for non-decreasing function ρ on $[0,\infty)$.

is integrable with respect to both measures $m_{V(A)}$ and $m_{\rho^{(A)}}$, then

$$\int_0^t w(s)\mathrm{d}A(s) := \int_0^t w(s)\mathrm{d}V(A)(s) - \int_0^t w(s)\mathrm{d}\rho^{(A)}(s)$$

for every $t \geq 0$. If w is integrable with respect to $m_{V(A)}$, then we shall use the following notation:

$$\int_0^t w(s)|\mathrm{d}A(s)| := \int_0^t w(s)\mathrm{d}V(A)(s)$$

for every $t \geq 0$.

If A is right-continuous, and w is continuous, or if A is continuous and w is left-right continuous on $(0,\infty)$, then

$$\int_0^t w(s)\mathrm{d}A(s) = \lim_{D\in\mathscr{P},|D|\to 0} \sum_{i\geq 1} w(t_{i-1})\left(A(t_i \wedge t) - A(t_{i-1}\wedge t)\right)$$

for any $t \geq 0$.

The (deterministic) theory of Lebesgue's, and Riemann-Stieltjes' integration may be applied to ample paths of stochastic processes.

Definition 5.1. *An adapted right-continuous process $A = (A_t : t \geq 0)$ is called a* finite variation process *(or a* process of finite variation*) if $A_0 = 0$ and $t \mapsto A_t$ is (a function) of finite variation a.s..*

Proposition 5.2. *Let A be a finite variation process and K a progressively measurable process s.t.*

$$\forall t \geq 0,\ \forall \omega \in \Omega,\ \int_0^t |K_s(\omega)||dA_s(\omega)| < \infty.$$

Then $((K\cdot A)_t : t \geq 0)$, defined as $(K\cdot A)_t(\omega) := \int_0^t K_s(\omega)dA_s(\omega)$, is a finite variation process.

Proof. The right continuity is immediate from the deterministic theory, but we need to check that $(K\cdot A)_t$ is adapted (and hence progressive, by Proposition 3.29). For this we check that if $t > 0$ is fixed and $h : [0,t]\times\Omega \to \mathbb{R}$ is measurable with respect to $\mathscr{B}([0,t])\otimes\mathscr{F}_t$, and if

$$\int_0^t |h(s,\omega)||dA_s(\omega)| < \infty$$

for every $\omega \in \Omega$, then

$$\int_0^t h(s,\omega)dA_s(\omega)$$

is $\mathscr{F}_t$-measurable.

Fix $t > 0$. Consider first h defined by $h(s,\omega) = 1_{(u,v]}(s)1_\Gamma(\omega)$ for $(u,v] \subseteq [0,t]$ and $\Gamma \in \mathscr{F}_t$. Then

$$(h \cdot A)_t = 1_\Gamma(A_v - A_u)$$

is $\mathscr{F}_t$-measurable. By the Monotone Class Theorem, $(h \cdot A)_t$ is $\mathscr{F}_t$-measurable for any $h = 1_G$ with $G \in \mathscr{B}([0,t]) \otimes \mathscr{F}_t$, or, more generally, any bounded $\mathscr{B}([0,t]) \otimes \mathscr{F}_t$-measurable function h. If h is a general $\mathscr{B}([0,t]) \otimes \mathscr{F}_t$-measurable function satisfying

$$\int_0^t |h(s,\omega)||dA_s(\omega)| < \infty \quad \forall \omega \in \Omega,$$

then h is a point-wise limit, $h = \lim_{n\to\infty} h_n$, of simple functions with $|h| \geq |h_n|$. The integrals $\int h_n(s,\omega)dA_s(\omega)$ converge by the Dominated Convergence Theorem, and hence $\int_0^t h(s,\omega)dA_s(\omega)$ is also $\mathscr{F}_t$-measurable (as a limit of $\mathscr{F}_t$-measurable functions). In particular, $(K \cdot A)_t(\omega)$ is $\mathscr{F}_t$-measurable since by progressive measurability, $(s,\omega) \mapsto K_s(\omega)$ on $[0,t] \times \Omega$ is $\mathscr{B}([0,t]) \otimes \mathscr{F}_t$-measurable. □

5.2 Stochastic integral with respect to continuous square integrable martingales

Let us first define stochastic integrals with respect to a continuous square integrable martingale $M = (M_t)_{t\geq 0}$, so that $\mathbb{E}\left[M_t^2\right] < \infty$. We shall define stochastic for *simple processes* first. Let $\mathscr{L}_0$ be the collection of all simple bounded processes (adapted to $(\mathscr{F}_t)_{t\geq 0}$) of the form

$$\varphi_t = \varphi^{(0)} 1_{\{0\}}(t) + \sum_{i=0}^{\infty} \varphi^{(i)} \mathbf{1}_{(t_i,t_{i+1}]}(t), \quad t \geq 0, \tag{20}$$

for some $m \in \mathbb{N}$, for some finite partition: $0 \leq t_0 < t_1 < \ldots < t_i < \cdots$ with $t_i \to \infty$, where $\varphi^{(i)}$ are *bounded* $\mathscr{F}_{t_i}$-measurable random variables. The the stochastic integral $\varphi \cdot M$ of φ in (20) with respect to M is defined by

$$(\varphi \cdot M)_t := \sum_{i=0}^{\infty} \varphi^{(i)} (M_{t\wedge t_{i+1}} - M_{t\wedge t_i}), \quad t \geq 0. \tag{21}$$

We shall show that $\varphi \cdot M$ is a continuous square integrable martingale. Of course the continuity of $\varphi \cdot M$ is obvious as M is continuous. Let $M_t^{(i)} = \varphi^{(i)}(M_{t\wedge t_{i+1}} - M_{t\wedge t_i})$ for $t \geq 0$. Then $(\varphi \cdot M)_t := \sum_{i=0}^{\infty} M_t^{(i)}$ (which is, for each t, is a finite sum indeed).

Since $M = (M_t)_{t\geq 0}$ is a continuous square integrable, so is every $(M_{t_i\wedge t})_{t\geq 0}$, and $\langle M_{t_i\wedge\cdot}\rangle_t = \langle M\rangle_{t\wedge t_i}$. Therefore

$$\mathbb{E}\left[M_{t_i\wedge t}|\mathscr{F}_s\right] = M_{t_i\wedge s}$$

and

$$\mathbb{E}\left[M_{t_i\wedge t}^2 - \langle M\rangle_{t\wedge t_i}|\mathscr{F}_s\right] = M_{t_i\wedge s}^2 - \langle M\rangle_{s\wedge t_i}$$

for any $t > s \geq 0$ and for every i.

Moreover, from the definition, $M_t^{(i)} = 0$ for $t \le t_i$, $M_t^{(i)} = \varphi^{(i)}(M_{t\wedge t_{i+1}} - M_{t\wedge t_i})$ for $t > t_i$. In particular, $M_t^{(i)}$ is $\mathscr{F}_{t_{i+1}}$-measurable for every $t \ge 0$. By using these facts we are in a position to show the basic facts stated in the following lemmas.

Lemma 5.3. *For each i, $M^{(i)}$ is martingale (where $i = 0, 1, \ldots$), and $\varphi \cdot M$ is a continuous martingale with initial zero.*

Proof. This involves a core (but simple) idea in Itô's definition of stochastic integrals: $\varphi^{(i)}$ is $\mathscr{F}_{t_i}$-measurable, and therefore φ is an adapted process. For $t > s \ge 0$. If $s \le t_i$, then

$$\begin{aligned}
\mathbb{E}\left[M_t^{(i)}|\mathscr{F}_s\right] &= \mathbb{E}\left[\mathbb{E}\left[\varphi^{(i)}(M_{t\wedge t_{i+1}} - M_{t\wedge t_i})|\mathscr{F}_{t_i}\right]|\mathscr{F}_s\right] \\
&= \mathbb{E}\left[\varphi^{(i)}\mathbb{E}\left[M_{t\wedge t_{i+1}} - M_{t\wedge t_i}|\mathscr{F}_{t_i}\right]|\mathscr{F}_s\right] \\
&= 0 = M_s^{(i)}.
\end{aligned}$$

If $s > t_i$, then $\varphi^{(i)} \in \mathscr{F}_{t_i} \subset \mathscr{F}_s$ and therefore

$$\begin{aligned}
\mathbb{E}\left[M_t^{(i)}|\mathscr{F}_s\right] &= \varphi^{(i)}\mathbb{E}\left[M_{t\wedge t_{i+1}} - M_{t\wedge t_i}|\mathscr{F}_s\right] \\
&= \varphi^{(i)}(M_{s\wedge t_{i+1}} - M_{s\wedge t_i})
\end{aligned}$$

for any $t > s$. This completes the proof. □

Lemma 5.4. *$\varphi \cdot M$ is a continuous, square integrable martingale, and the quadratic variation process*

$$\langle \varphi \cdot M \rangle_t = \int_0^t \varphi_s^2 d\langle M \rangle_s \quad \textit{for every } t \ge 0.$$

Therefore

$$\mathbb{E}\left[|(\varphi \cdot M)_t|^2\right] = \mathbb{E}\left[\int_0^t \varphi_s^2 d\langle M \rangle_s\right] \tag{22}$$

for every $t \ge 0$, which is called Itô's isometry.

Proof. This is really the key fact in the definition of Itô's integrals. We divide the proof in several steps.

First, for $i \ne j$, the intervals $(t_i, t_{i+1}]$ and $(t_j, t_{j+1}]$ are disjoint, we show that $M_t^i M_t^j$ is a martingale, hence $\langle M^i, M^j \rangle_t = 0$. We may assume that $i < j$ without losing generality. Let $t > s \ge 0$.

(1) $t \le t_j$, then $M_t^{(j)} = M_s^{(j)} = 0$, so that $\mathbb{E}\left[M_t^{(i)} M_t^{(j)}|\mathscr{F}_s\right] = 0$.

(2) $t > t_j$ but $s \le t_j$, then $\mathscr{F}_s \subset \mathscr{F}_{t_j}$, and $M_t^{(i)}$ is $\mathscr{F}_{t_{i+1}}$ measurable (for every t). Since $i < j$, $t_{i+1} \le t_j$, thus $M_t^{(i)}$ is $\mathscr{F}_{t_j}$-measurable. Therefore

$$\begin{aligned}
\mathbb{E}\left[M_t^{(i)} M_t^{(j)}|\mathscr{F}_s\right] &= \mathbb{E}\left[\mathbb{E}\left(M_t^{(i)} M_t^{(j)}|\mathscr{F}_{t_j}\right)|\mathscr{F}_s\right] \\
&= \mathbb{E}\left[M_t^{(i)}\mathbb{E}\left(M_t^{(j)}|\mathscr{F}_{t_j}\right)|\mathscr{F}_s\right] \\
&= \mathbb{E}\left[M_t^{(i)} M_{t_j}^{(j)}|\mathscr{F}_s\right] = 0 = M_s^{(i)} M_s^{(j)}.
\end{aligned}$$

(3) $t > t_j$ and $s \geq t_j$, then $\mathscr{F}_{t_{i+1}} \subset \mathscr{F}_{t_j} \subset \mathscr{F}_s$ $M_t^{(i)}$ is $\mathscr{F}_s$ measurable and $M_t^{(i)} = M_s^{(i)}$, hence

$$\mathbb{E}\left[M_t^{(i)} M_t^{(j)} | \mathscr{F}_s\right] = M_s^{(i)} \mathbb{E}\left[M_t^{(j)} | \mathscr{F}_s\right] = M_s^{(i)} M_s^{(j)}.$$

Therefore $\{M_t^{(i)} M_t^{(j)}; t \geq 0\}$ is a continuous martingale for $i \neq j$.

Next we prove that, for any but fixed i, $\{(M_t^{(i)})^2 - A_t^{(i)}; t \geq 0\}$ is a martingale, where

$$A_t^{(i)} = \left(\varphi^{(i)}\right)^2 \left(\langle M \rangle_{t_{i+1} \wedge t} - \langle M \rangle_{t_i \wedge t}\right) \quad \text{ for } t \geq 0.$$

For simplicity, we set $N_t = (M_t^{(i)})^2 - A_t^{(i)}$ for $t \geq 0$. We notice that $A_t^{(i)} = 0$ for $t \leq t_i$, and $A_t^{(i)}$ is $\mathscr{F}_{t_{i+1}}$-measurable, and $A_t^{(i)} = A_s^{(i)}$ for $t > s \geq t_{i+1}$. Therefore $N_t = 0$ for $t \leq t_i$, N_t is $\mathscr{F}_{t_{i+1}}$-measurable for any t, and $N_t = N_s$ for any $t > s \geq t_{i+1}$.

Suppose $t > s \geq 0$.

(1) If $t \leq t_i$, then $N_t = N_s = 0$ so $\mathbb{E}[N_t | \mathscr{F}_s] = N_s$.

(2) If $t > t_i$ but $s \leq t_i$, then $N_s = 0$ and $(\varphi^{(i)})^2$ is $\mathscr{F}_{t_i}$-measurable, so that

$$\mathbb{E}[N_t | \mathscr{F}_s] = \mathbb{E}[\mathbb{E}(N_t | \mathscr{F}_{t_i}) | \mathscr{F}_s],$$

while

$$\begin{aligned}
\mathbb{E}(N_t | \mathscr{F}_{t_i}) &= (\varphi^{(i)})^2 \mathbb{E}\left(\left(M_{t_{i+1} \wedge t} - M_{t_i}\right)^2 - \left(\langle M \rangle_{t_{i+1} \wedge t} - \langle M \rangle_{t_i}\right) | \mathscr{F}_{t_i}\right) \\
&= (\varphi^{(i)})^2 \mathbb{E}\left(M_{t_{i+1} \wedge t}^2 + M_{t_i}^2 - 2 M_{t_{i+1} \wedge t} M_{t_i} - \langle M \rangle_{t_{i+1} \wedge t} + \langle M \rangle_{t_i} | \mathscr{F}_{t_i}\right) \\
&= (\varphi^{(i)})^2 \mathbb{E}\left(M_{t_{i+1} \wedge t}^2 - \langle M \rangle_{t_{i+1} \wedge t} - M_{t_i}^2 + \langle M \rangle_{t_i} - 2 M_{t_{i+1} \wedge t} M_{t_i} + 2 M_{t_i}^2 | \mathscr{F}_{t_i}\right) \\
&= (\varphi^{(i)})^2 \left(M_{t_{i+1} \wedge t_i}^2 - \langle M \rangle_{t_{i+1} \wedge t_i} - M_{t_i}^2 + \langle M \rangle_{t_i} - 2 M_{t_{i+1} \wedge t_i} M_{t_i} + 2 M_{t_i}^2\right) \\
&= 0
\end{aligned}$$

and therefore $\mathbb{E}[N_t | \mathscr{F}_s] = 0 = N_s$. Here we have used the fact that $\{M_{t_k \wedge t}^2 - \langle M \rangle_{t_k \wedge t}; t \geq 0\}$ is a stopped martingale, so it is a martingale, for every k.

(3) If $t > s > t_i$, then $(\varphi^{(i)})^2$ is $\mathscr{F}_{t_i}$-measurable, so we have

$$\begin{aligned}
\mathbb{E}(N_t | \mathscr{F}_s) &= (\varphi^{(i)})^2 \mathbb{E}\left(\left(M_{t_{i+1} \wedge t} - M_{t_i \wedge t}\right)^2 - \left(\langle M \rangle_{t_{i+1} \wedge t} - \langle M \rangle_{t_i \wedge t}\right) | \mathscr{F}_s\right) \\
&= (\varphi^{(i)})^2 \mathbb{E}\left(M_{t_{i+1} \wedge t}^2 - \langle M \rangle_{t_{i+1} \wedge t} - M_{t_i \wedge t}^2 + \langle M \rangle_{t_i \wedge t} - 2 M_{t_{i+1} \wedge t} M_{t_i} + 2 \langle M \rangle_{t_i} | \mathscr{F}_s\right) \\
&= (\varphi^{(i)})^2 \left(M_{t_{i+1} \wedge s}^2 - \langle M \rangle_{t_{i+1} \wedge s} - M_{t_i \wedge s}^2 + \langle M \rangle_{t_i \wedge s} - 2 M_{t_{i+1} \wedge s} M_{t_i} + 2 \langle M \rangle_{t_i}\right) \\
&= (\varphi^{(i)})^2 \left(\left(M_{t_{i+1} \wedge s} - M_{t_i \wedge s}\right)^2 - \left(\langle M \rangle_{t_{i+1} \wedge s} - \langle M \rangle_{t_i \wedge s}\right)\right) \\
&= N_s
\end{aligned}$$

which completes the proof that $M^{(i)}$ is a continuous, square integrable martingale with initial zero, and

$$\left\langle M^{(i)} \right\rangle_t = \left(\varphi^{(i)}\right)^2 \left(\langle M \rangle_{t \wedge t_{i+1}} - \langle M \rangle_{t_i \wedge t}\right)$$

for $t \geq 0$.

We show now that $\langle \varphi \cdot M \rangle_t = \int_0^t \varphi_s^2 d\langle M \rangle_s$. Let $A_t = \sum_{i=0}^{\infty} \left\langle M^{(i)} \right\rangle_t$ which is indeed a finite sum for every $t \geq 0$, and let $N_t = \langle \varphi \cdot M \rangle_t - A_t$ for $t \geq 0$. Then A is non-decreasing, continuous, adapted, and $A_0 = 0$. Since

$$(\varphi \cdot M)_t^2 = \sum_{i,j} M_t^{(i)} M_t^j$$

so that

$$(\varphi \cdot M)_t^2 - A_t = \sum_{i \neq j} M_t^{(i)} M_t^j + \sum_{i=0}^{\infty} \left((M_t^{(i)})^2 - \left\langle M^{(i)} \right\rangle_t \right).$$

Now every term on the right had side is a martingale, thus, for $t > s \geq 0$, we have

$$\begin{aligned}\mathbb{E}\left[(\varphi \cdot M)_t^2 - A_t | \mathscr{F}_s \right] &= \sum_{i \neq j} M_s^{(i)} M_s^j + \sum_{i=0}^{\infty} \left((M_s^{(i)})^2 - \left\langle M^{(i)} \right\rangle_s \right) \\ &= (\varphi \cdot M)_s^2 - A_s\end{aligned}$$

which implies that

$$\langle \varphi \cdot M \rangle_t = A_t = \sum_{i=0}^{\infty} \left\langle M^{(i)} \right\rangle_t = \sum_{i=0}^{\infty} \left(\varphi^{(i)} \right)^2 \left(\langle M \rangle_{t \wedge t_{i+1}} - \langle M \rangle_{t_i \wedge t} \right).$$

On the other hand, since

$$\varphi_t^2 = (\varphi^{(0)})^2 1_{\{0\}}(t) + \sum_{i=0}^{\infty} (\varphi^{(i)})^2 \mathbf{1}_{(t_i, t_{i+1}]}(t), \quad t \geq 0,$$

so that

$$\begin{aligned}\int_0^t \varphi_s^2 d\langle M \rangle_s &= \int_0^t (\varphi^{(0)})^2 1_{\{0\}}(s) d\langle M \rangle_s + \sum_{i=0}^{\infty} \int_0^t (\varphi^{(i)})^2 \mathbf{1}_{(t_i, t_{i+1}]}(s) d\langle M \rangle_s \\ &= \sum_{i=0}^{\infty} (\varphi^{(i)})^2 \int_0^t \mathbf{1}_{(t_i, t_{i+1}]}(s) d\langle M \rangle_s \\ &= \sum_{i=0}^{\infty} (\varphi^{(i)})^2 \int_{t_i \wedge t}^{t_{i+1} \wedge t} d\langle M \rangle_s \\ &= \sum_{i=0}^{\infty} (\varphi^{(i)})^2 \left(\langle M \rangle_{t \wedge t_{i+1}} - \langle M \rangle_{t_i \wedge t} \right) \\ &= \langle \varphi \cdot M \rangle_t .\end{aligned}$$

The Itô's isometry follows the fact the initial values of both $\varphi \cdot M$ and $\langle \varphi \cdot M \rangle$ are zero, so that

$$\mathbb{E}\left[(\varphi \cdot M)_t^2 - \int_0^t \varphi_s^2 d\langle M \rangle_s \right] = 0$$

which yields that

$$\mathbb{E}\left[\left(\int_0^t \varphi_s \mathrm{d}M_s\right)^2\right] = \mathbb{E}\left[\int_0^t \varphi_s^2 \mathrm{d}\langle M\rangle_s\right]$$

for every $t \geq 0$. $\square$

Lemma 5.5. *Suppose $N = (N_t)_{t\geq 0}$ is another continuous square integrable martingale, and ψ is another simple process, then*

$$\langle \varphi \cdot M, N\rangle_t = \int_0^t \varphi_s d\langle M, N\rangle_s$$

and

$$\langle \varphi \cdot M, \psi \cdot N\rangle_t = \int_0^t \varphi_s \psi_s d\langle M, N\rangle_s$$

for every $t \geq 0$.

The proof is similar. In fact

$$\langle \varphi \cdot M, N\rangle_t = \sum_{i=0}^{\infty} \langle M^i, N\rangle_t = \sum_{i=0}^{\infty} \varphi^{(i)} \left(\langle M, N\rangle_{t_{i+1}\wedge t} - \langle M, N\rangle_{t_i \wedge t}\right)$$

which yields the conclusion.

Exercise 5.6. *Let $M \in \mathscr{H}^{2,c}$. The mapping $\varphi \mapsto \varphi \cdot M$ is a linear map from $\mathscr{L}_0$ to $\mathscr{H}_0^{2,c}$. Moreover,*

$$\|\varphi \cdot M\|_{\mathscr{H}^2}^2 = \mathbb{E}\left[\int_0^{\infty} \varphi_t^2 d\langle M\rangle_t\right]. \tag{23}$$

*[*Hint: The proof is easy – we just need to show linearity. But given φ, $\psi \in \mathscr{L}_0$, we use a refinement of the partitions on which they are constant to write them as simple functions with respect to the same partition and the result is trivial.*]*

We shall now extend Itô's integrals to a large class of integrands by using Itô's isometry (22). That is, we shall extend the definition of Itô's integrals with respect to a continuous square integrable martingale M to those in the closure of simple processes under the distance induced by the Itô's isometry. Indeed the closure of $\mathscr{L}_0$ preserving Itô's isometry can be identified.

To this end we should revel an important that every simple process is by definition is left continuous on $(0,\infty)$, and therefore a simple process is more than progressive measurability. By taking closure, we expect this type of measurability has to be maintained.

Definition 5.7. *Let $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. The σ-algebra on $[0,\infty)\times\Omega$ generated by all adapted processes which are left continuous on $(0,\infty)$, is called the predictable σ-algebra* *on* $[0,\infty)\times\Omega$, *denote by* $\mathscr{P}([0,\infty)\times\Omega)$[56]. *A stochastic process* $\Phi = (\Phi_t)_{t\geq 0}$ *is predictable if the function* $\Phi(t,\omega) := \Phi_t(\omega)$ *on* $[0,\infty)\times\Omega$ *is* $\mathscr{P}([0,\infty)\times\Omega)$*-measurable.*

By definition, any left continuous, adapted process (hence any simple process) is predictable. Any predictable process is progressively measurable.

Definition 5.8. *Let* $\mathscr{L}_2(M)$ *denote the vector space of all predictable processes* $\Phi = (\Phi_t)_{\geq 0}$ *such that*

$$\mathbb{E}\left[\int_0^t \Phi_s^2 d\langle M\rangle_s\right] < \infty \quad \text{for any } t > 0.$$

For $\Phi \in \mathscr{L}_2(M)$, *define*

$$\|\Phi\|_{2;M} = \sum_{n=1}^{\infty}\frac{1}{2^n}\left(\|\Phi\|_{2,n;M}\wedge 1\right), \quad \text{where } \|\Phi\|_{2,T;M} = \sqrt{\mathbb{E}\left[\int_0^T \Phi_s^2 d\langle M\rangle_s\right]},$$

and its induced metric $\|\Phi-\Psi\|_{2;M}$ *for any* $\Phi,\Psi\in\mathscr{L}_2(M)$.

The key fact we need is the following fact.

Lemma 5.9. *The space* $\mathscr{L}_0$ *of simple, adapted and bounded processes is dense in* $\mathscr{L}_2(M)$ *under the metric*[57] $\|\Phi-\Psi\|_{2;M}$.

Proof. [The proof is not examinable.] If $\Phi\in\mathscr{L}_2(M)$, then $\left\|\Phi-\Phi 1_{\{|\Phi|\leq m}\right\|_{2;M}\to 0$ as $m\to\infty$. Therefore we prove this lemma for bounded $\Phi\in\mathscr{L}_2(M)$. Furthermore, by a monotone class argument, we only need to prove that every bounded, left continuous and adapted process $\Phi = (\Phi_t)_{t\geq 0}$ can be approximated by simple processes. Choose a sequence of partitions $D^{(n)} = \{t_i^{(n)}\}$ with $t_i^{(n)}\to\infty$ as $i\to\infty$ for each $n = 1,2,\ldots$, such that $|D^{(n)}|\to 0$ when $n\to\infty$. Define

$$\Phi_t^{(n)} = \Phi_0 1_{\{0\}}(t) + \sum_{i=0}^{\infty}\Phi_{t_i^{(n)}} 1_{(t_i^{(n)},t_{i+1}^{(n)}]}(t)$$

for $t\geq 0$ and $n\to\infty$. $\Phi^{(n)}\to\Phi$ point-wise when $n\to\infty$, as Φ is left continuous. Since Φ is bounded and $\Phi^{(n)}$ are bounded by the same bound of $|\Phi|$. Therefore by Lebesgue's bounded convergence theorem, one can conclude that $\left\|\Phi^{(n)}-\Phi\right\|_{2;M}\to 0$ as $n\to\infty$, which finishes the proof. □

We are now in a position to definite stochastic integral of $\Phi\cdot M$ for $\Phi\in\mathscr{L}_2(M)$ as the following.

[56] The predictable σ-algebra by definition depends on the underlying filtration $(\mathscr{F}_t)_{t\geq 0}$.

[57] In fact the metric may be a quasi-one, and we shall consider Φ and Ψ as the same element in $\mathscr{L}_2(M)$ if $\|\Phi-\Psi\|_{2;M} = 0$.

Choose a sequence of bounded simple processes $\Phi^{(n)} \in \mathscr{L}_0$ ($n = 1,2,\ldots$) such that $\left\|\Phi^{(n)} - \Phi\right\|_{2,M} \to 0$ as $n \to \infty$. Then $\Phi^{(n)} \cdot M$ are continuous square integrable martingales with initial zero, and, by Itô's isometry

$$\mathbb{E}\left[\left|(\Phi^{(n)} \cdot M)_t - (\Phi^{(m)} \cdot M)_t\right|^2\right] = \mathbb{E}\left[\int_0^t |\Phi_s^{(n)} - \Phi_s^{(m)}|^2 d\left\langle M\right\rangle_s\right] \to 0$$

as $n,m \to \infty$ for every $t > 0$, and therefore $\{\Phi^{(n)} \cdot M : n = 1,2,\ldots\}$ is a Cauchy sequence in $\mathscr{M}_2^c$. Therefore there exists a unique limit in $\mathscr{M}_2^c$, denoted by $\Phi \cdot M$, which is the unique continuous, square integrable martingale with initial zero, such that $\Phi^{(n)} \cdot M \to \Phi \cdot M$ as $n \to \infty$. $\Phi \cdot M$ is called the Itô integral of Φ against the continuous martingale M, which is independent of the choice of an approximation sequence $\Phi^{(n)}$.

The following facts follow immediately from the construction of Itô's stochastic integrals.

Proposition 5.10. *Let $D^{(n)} = \{t_i^{(n)}\}$ with $t_i^{(n)} \to \infty$ as $i \to \infty$ for each $n = 1,2,\ldots$, a sequence of partitions such that $|D^{(n)}| \to 0$. If $\Phi = (\Phi_t)_{t\geq 0}$ is* left continuous and adapted*, and*

$$\mathbb{E}\left[\int_0^t |\Phi_s|^2 d\left\langle M\right\rangle_s\right] < \infty \quad \textit{for every } t > 0.$$

Then

$$\int_0^t \Phi_s dM_s := (\Phi \cdot M)_t = \lim_{n\to\infty} \sum_i \Phi_{t_i^{(n)}} (M_{t_{i+1}^{(n)}\wedge t} - M_{t_i^{(n)}\wedge t}) \tag{24}$$

in probability, for every $t \geq 0$.

Proposition 5.11. *Itô's integrals possess the following properties.*

1) If $M = (M_t)_{t\geq 0}$ and $N = (N_t)_{t\geq 0}$ are continuous square integrable martingales, $\Phi \in \mathscr{L}_2(M)$ and $\Psi \in \mathscr{L}_2(N)$, then $\Phi \cdot M$ and $\Psi \cdot N$ are continuous square integrable martingales with initial zero, and

$$\left\langle \Phi \cdot M, \Psi \cdot N\right\rangle_t = \int_0^t \Phi_s \Psi_s d\left\langle M,N\right\rangle_s \quad \textit{for } t \geq 0.$$

Moreover $(\Phi \cdot M)_t(\Psi \cdot N)_t - \int_0^t \Phi_s \Psi_s d\left\langle M,N\right\rangle_s$ is a continuous martingale.

2) Itô's isometry holds

$$\mathbb{E}\left[\left(\int_0^t \Phi_s dM_s\right)^2\right] = \mathbb{E}\left[\int_0^t |\Phi_s|^2 d\left\langle M\right\rangle_s\right] \quad \textit{for } t \geq 0.$$

3) If $M = (M_t)_{t\geq 0}$ is a continuous square integrable martingale, $\Phi \in \mathscr{L}_2(M)$ and $\Psi \in \mathscr{L}_2(\Phi \cdot M)$, then

$$\Psi \cdot (\Phi \cdot M) = (\Psi \cdot \Phi) \cdot M.$$

As an immediate application, we have the following consequence.

Corollary 5.12. *Let $\Phi \in \mathscr{L}_2(M)$ and T be a stopping time. Then*

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

2) It holds that

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

Proof. 1) follows from the previous proposition immediately. 2) We only need to show this for left continuous and bounded $\Phi \in \mathscr{L}_2(M)$. Then (by using the notations in Proposition 5.10), we have

$$\begin{aligned}(\Phi \cdot M)_{t\wedge T} &= \lim_{n\to\infty} \sum_i \Phi_{t_i^{(n)}} (M_{t_{i+1}^{(n)}\wedge t\wedge T} - M_{t_i^{(n)}\wedge t\wedge T}) \\ &= \lim_{n\to\infty} \sum_i \Phi_{t_i^{(n)}} (M^T_{t_{i+1}^{(n)}\wedge t} - M^T_{t_i^{(n)}\wedge t})\end{aligned}$$

which gives that $(\Phi \cdot M)^T = \Phi \cdot M^T$. □

5.3 Stochastic integral with respect to Brownian motion

The theory can be applied to Brownian motion with initial zero. Let $B = (B_t)$ be a Brownian motion. Then $(B_t)^2 - t$ is a continuous martingale, so that $\langle B \rangle_t = t$. Therefore $\Phi \in \mathscr{L}_2(B)$ if it is predictable and

$$\mathbb{E}\left[\int_0^t |\Phi_s|^2 ds\right] < \infty \quad \text{for any } t > 0.$$

Then $\Phi \cdot B$ is a continuous square integral martingale with initial zero. If $\Phi = (\Phi_t)_{t\geq 0}$ is left continuous, then

$$\int_0^t \Phi_s dB_s = \lim_{D:0=t_0<\cdots<t_m=t, |D|\to 0} \sum_i \Phi_{t_i}(B_{t_{i+1}} - B_{t_i})$$

for every $t \geq 0$, in probability. Note we use the left hand value Φ_{t_i} on the interval $[t_i, t_{i+1}]$. This is the key idea in Itô's original approach to ensure the resulted integral is a martingale. If f is a Borel measurable function, then $\Phi_t = f(B_t)$ (for $t \geq 0$) is progressively measurable, and is predictable with respect to the so-called Brownian filtration. Then $\{f(B_t) : t \geq 0\} \in \mathscr{L}_2(B)$ if and only if

$$\begin{aligned}\mathbb{E}\left[\int_0^t |f(B_s)|^2 ds\right] &= \int_0^t \mathbb{E}\left[|f(B_s)|^2\right] ds \\ &= \int_0^t \frac{1}{\sqrt{2\pi s}} \int_{-\infty}^{\infty} f(x)^2 e^{-\frac{x^2}{2s}} dx ds < \infty.\end{aligned}$$

The Itô's integral, in the case f is continuous, is defined by

$$\int_0^t f(B_s) dB_s = \lim_{D:0=t_0<\cdots<t_m=t, |D|\to 0} \sum_i f(B_{t_i})(B_{t_{i+1}} - B_{t_i})$$

for $t \geq 0$. In particular

$$\begin{aligned}
\int_0^t B_s dB_s &= \lim_{D:0=t_0<\cdots<t_m=t,|D|\to 0} \sum_i B_{t_i}(B_{t_{i+1}} - B_{t_i}) \\
&= \lim_{D:0=t_0<\cdots<t_m=t,|D|\to 0} \sum_i \left[-\frac{1}{2}(B_{t_{i+1}} - B_{t_i})^2 + \frac{1}{2}\left(B_{t_{i+1}}^2 - \frac{1}{2}B_{t_i}^2 \right) \right] \\
&= \frac{1}{2}B_t^2 - \frac{1}{2}t
\end{aligned}$$

In other words

$$B_t^2 - t = 2\int_0^t B_s dB_s \quad \text{for } t \geq 0$$

which is a special case of Itô's formula. This is *not* what one would have predicted from classical integration theory (the extra term, called the correction term, here comes from the quadratic variation.

Even more strangely, it matters that in (24) we took the *left* endpoint of the interval for evaluating the integrand. On the problem sheet, you are asked to evaluate

$$\lim_{|D|\to 0} \sum_j B_{t_{j+1}} \left(B_{t_{j+1}} - B_{t_j} \right), \quad \text{and } \lim_{D\to 0} \sum \frac{B_{t_j} + B_{t_{j+1}}}{2} \left(B_{t_{j+1}} - B_{t_j} \right).$$

Each gives a different answer.

We can more generally define

$$\int_0^T f(B_s) \circ dB_s = \lim_{|D|\to 0} \sum \left(\frac{f(B_{t_j}) + f(B_{t_{j+1}})}{2} \right) \left(B_{t_{j+1}} - B_{t_j} \right).$$

This is the so-called *Stratonovich integral*, and has the advantage that from the point of view of calculations, the rules of Newtonian calculus hold true. From a modelling perspective however, it can be the wrong choice. For example, suppose that we are modelling the change in a population size over time and we use $[t_i, t_{i+1})$ to represent the $(i+1)$st generation. The change over (t_i, t_{i+1}) will be driven by the number of *adults*, so the population size at the *beginning* of the interval.

5.4 Stochastic integration with respect to continuous local martingales

We shall now extend Itô's theory of stochastic integration to continuous local martingales.

Let $M = (M_t)_{t\geq 0}$ be a continuous local martingale with initial zero. Suppose

$\Phi = (\Phi_t)_{t\geq 0}$ is a progressively measurable process[58] such that

$$\int_0^t |\Phi_s|^2 d\langle M\rangle_s < \infty \quad \text{almost surely} \tag{25}$$

for every $t > 0$. The collection of all such Φ is denoted by $\mathscr{L}_{2,\text{loc}}(M)$. Let

$$T^{(n)} = \inf\left\{ t \geq 0 : \int_0^t |\Phi_s|^2 d\langle M\rangle_s + |M_t| + \langle M\rangle_t \geq n \right\},$$

$n = 1, 2, \ldots$. $T^{(n)}$ are stopping times and our assumption above implies that $T^{(n)} \uparrow \infty$ when $n \to \infty$. For each n, $\Phi^{T^{(n)}}$ is bounded and $M^{T^{(n)}}$ is a bounded continuous martingale and $\left\langle M^{T^{(n)}} \right\rangle = \langle M\rangle^{T^{(n)}}$. Therefore $\Phi^{T^{(n)}} \in \mathscr{L}_2(M^{T^{(n)}})$ and the stochastic integral $\Phi^{T^{(n)}} \cdot M^{T^{(n)}}$ is a continuous square integrable martingale with initial zero. According to the properties of stochastic integrals with respect to continuous square integrable martingales, we have

$$\left(\Phi^{T^{(n)}} \cdot M^{T^{(n)}} \right)^{T^{(m)}} = \Phi^{T^{(m)}} \cdot M^{T^{(m)}}$$

for any $n \geq m$. Therefore we may construct a process, denoted by $\Phi \cdot M$, by $(\Phi \cdot M)_t = (\Phi^{T^{(n)}} \cdot M^{T^{(n)}})_t$ on $\{t \leq T^{(n)}\}$ for $n = 1, 2, \ldots$. Then $(\Phi \cdot M)^{T^{(n)}} = \Phi^{T^{(n)}} \cdot M^{T^{(n)}}$ for every $n = 1, 2, \ldots$. Since $T^{(n)}$ are stopping times and $T^{(n)} \uparrow \infty$, $\Phi \cdot M$ is a continuous local martingale with initial zero, and by definition

$$\langle \Phi \cdot M\rangle_t = \int_0^t |\Phi_s|^2 d\langle M\rangle_s \quad \text{almost surely}$$

for every $t \geq 0$. It follows that condition (25) is also necessary indeed to ensure that the resulted stochastic integral $\Phi \cdot M$ is a continuous local martingale.

Lemma 5.13. *Let $X^{(n)} = (X_t^{(n)})_{t\geq 0}$ ($n = 1, 2, \ldots$) be a sequence of continuous stochastic processes, and T_k ($k = 1, 2, \ldots$) be a sequence stopping times such that $T_k \uparrow \infty$ as $k \to \infty$. Suppose for each k, the stopped processes $(X^{(n)})^{T_k}$ converges to zero in probability, uniformly on any bounded interval, as $n \to \infty$. Then $X^{(n)} \to 0$ in probability, uniformly on any bounded interval[59], as $n \to \infty$.*

---

[58] Here we departed from the requirement that Φ should be predictable, due to the assumption that M is *continuous* local martingale. In fact, there is a predictable process $\tilde{\Phi}$ such that $\Phi = \tilde{\Phi}$ with respect to the measure induced by $\langle M\rangle$, or equivalently

$$\int_0^t |\Phi_s - \tilde{\Phi}_s|^2 d\langle M\rangle_s = 0 \quad \text{for all } t \geq 0, a.e.$$

Hence we can define stochastic integral $\Phi \bullet M$ to be $\tilde{\Phi} \bullet M$ naturally.

[59] We have used this technique in the proof Proposition 3.65.

Proof. For every $\delta > 0$ we have

$$\mathbb{P}\left[\sup_{s\in[0,t]}|X_s^{(n)}| > \delta\right] \le \mathbb{P}\left[\sup_{s\in[0,t]}|X_s^{(n)}| > \delta, T_k \ge t\right] + \mathbb{P}[T_k < t]$$
$$\le \mathbb{P}\left[\sup_{s\in[0,t]}|(X_s^{(n)})^{T_k}| > \delta\right] + \mathbb{P}[T_k < t],$$

so by letting $n \to \infty$ we obtain that

$$\limsup_{n\to\infty}\mathbb{P}\left[\sup_{s\in[0,t]}|X_s^{(n)}| > \delta\right] \le \mathbb{P}[T_k < t]$$

for any k. Sending $k \to \infty$ we conclude that

$$\lim_{n\to\infty}\mathbb{P}\left[\sup_{s\in[0,t]}|X_s^{(n)}| > \delta\right] = 0$$

for any $t > 0$, which completes the proof. □

Proposition 5.14. *Let $M = (M_t)_{t\ge 0}$ be a continuous local martingale and $\Phi = (\Phi_t)_{t\ge 0} \in \mathscr{L}_{2,loc}(M)$ be left-continuous on $(0,\infty)$. Then*

$$\int_0^t \Phi_s dM_s = \lim_{D:0=t_0<\cdots<t_m=t, |D|\to 0} \sum_i \Phi_{t_i}(M_{t_{i+1}} - M_{t_i}) \quad \textit{in probability}$$

for every $t \ge 0$.

Proof. For any processes $\phi = (\phi_t)_{t\ge 0}$, $X = (X_t)_{t\ge 0}$ and finite partition $D : 0 = t_0 < \cdots < t_i < \cdots$ with $t_i \to \infty$ we define

$$R_t^D(\phi, X) = \sum_{i=0}^{m-1} \phi_{t_i}(X_{t_{i+1}\wedge t} - X_{t_i\wedge t}).$$

Choose a sequence of partitions of $[0,\infty)$, $D^{(n)} = \{t_i^{(n)}\}$ with $t_i^{(n)} \to \infty$ as $i \to \infty$, where $n = 1, 2, \ldots$ Let $T^{(k)}$ be defined as above. Then $\Phi^{T^{(k)}}$ is left continuous, and therefore

$$(\Phi\cdot M)_t^{T^{(k)}} = (\Phi^{T^{(k)}}\cdot M^{T^{(k)}})_t = \lim_{n\to\infty} R_t^{D^{(n)}}(\Phi^{T^{(k)}}, X^{T^{(k)}})$$
$$= \lim_{n\to\infty}\left(R_t^{D^{(n)}}(\Phi, X)\right)^{T^{(k)}}$$

in $\mathscr{M}_2$, for every $t \ge 0$. By Doob's maximal inequality

$$(\Phi\cdot M)_t^{T^{(k)}} - \left(R_t^{D^{(n)}}(\Phi, X)\right)^{T^{(k)}}$$

converges to zero uniformly for t in any bounded interval, in probability. Now the conclusion follows from the previous lemma applying to the sequence of continuous processes

$$X^{(n)} = \Phi\cdot M - R^{D^{(n)}}(\Phi, X)$$

where $n = 1, 2, \ldots$. □

5.5 Stochastic integration with respect to continuous semimartingales

Naturally, we are going to define an integral with respect to a continuous semimartingale $X = X_0 + M + A$ as a sum of integrals w.r.t. the continuous local martingale M and w.r.t. the continuous, adapted, variational process A.

We shall use $\mathscr{L}_{\text{loc}}(X)$ to denote the space of those $\Phi = (\Phi_t)_{t\geq 0} \in \mathscr{L}_{2,\text{loc}}(M)$ such that

$$\int_0^t |\Phi_s||dA_s| < \infty \quad \text{for all } t > 0.$$

Then we define Itô's integral via

$$\int_0^t \Phi_s dX_s = \int_0^t \Phi_s dM_s + \int_0^t \Phi_s dA_s \quad \text{for } t \geq 0.$$

Then the stochastic integral $\left(\int_0^t \Phi_s dX_s\right)_{t\geq 0}$ is again a continuous semimartingale.

Proposition 5.15. *Let $X = X_0 + M + A$ be a continuous semimartingale. Suppose $\Phi = (\Phi_t)_{t\geq 0} \in \mathscr{L}_{loc}(X)$ is left continuous, then*

$$\int_0^t \Phi_s dX_s = \lim_{D:0=t_0<\cdots<t_m=t, |D|\to 0} \sum_i \Phi_{t_i}(X_{t_{i+1}} - X_{t_i}) \quad \textit{in probability,}$$

for every $t \geq 0$.

Stochastic dominated convergence

We should also like to know how our integral behaves under limits.

Proposition 5.16 (Stochastic Dominated Convergence Theorem)**.** *Let X be a continuous semimartingale and $\Phi^{(n)}$ a sequence in $\mathscr{L}_{loc}(X)$ with $\Phi_t^{(n)} \to 0$ as $n \to \infty$ for all t almost surely. Further suppose that $|\Phi_t^{(n)}| \leq \Psi_t$ for all n where $\Psi \in \mathscr{L}_{loc}(X)$. Then $\Phi^{(n)} \cdot X$ converges to zero in probability and, more precisely,*

$$\forall t \geq 0 \quad \sup_{s\leq t} \left| \int_0^s \Phi_u^{(n)} dX_u \right| \longrightarrow 0 \textit{ in probability as } n \to \infty.$$

Proof. Clearly we only need to treat the case where $X = M$ is a continuous local martingale. Let

$$T_k = \inf\{t \geq 0 : |M_t| + \langle M \rangle_t + \int_0^t |\Psi_s|^2 d\langle M \rangle_s \geq k\}$$

for $k = 1, 2, \ldots$. Then the stopping times $T_k \uparrow \infty$. By a similar argument as in the proof of Proposition 24, the conclusion follows from Lemma 5.13. □

6 Itô's formula and its applications

In this chapter we shall establish the major tool in stochastic analysis, namely, Itô's formula, and gives a couple of direct applications. The full power will be demonstrated in Papers C8.1 and C8.2, in your Part C next year.

6.1 Itô's formula and integration by parts

We already saw that the stochastic integral of Brownian motion with respect to itself did not behave as we would expect from calculus. So what are the analogues of integration by parts and the chain rule for stochastic integrals?

Proposition 6.1 (Integration by parts). *If X and Y are two continuous semimartingales then*

$$X_tY_t - X_0Y_0 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X,Y\rangle_t$$

for every $t \geq 0$.

Proof. Fix t and let $D^{(n)} = \{t_i^{(n)}\}$ (where for each n, $t_i^{(n)} \to \infty$ as $i \to \infty$) be a sequence of partitions of $[0,\infty)$ with mesh $|D^{(n)}| \to 0$ when $n \to \infty$. Using equality

$$X_tY_t - X_sY_s = X_s(Y_t - Y_s) + Y_s(X_t - X_s) + (X_t - X_s)(Y_t - Y_s)$$

we obtain that (where $t_i = t_i^{(n)} \wedge t$ for fixed but any n)

$$\begin{aligned} X_tY_t - X_0Y_0 =& \sum_i \Big(X_{t_i}(Y_{t_{i+1}} - Y_{t_i}) + Y_{t_i}(X_{t_{i+1}} - X_{t_i}) + (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i})\Big) \\ &\longrightarrow \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X,Y\rangle_t \quad \text{as } n \to \infty \end{aligned}$$

where the convergence is convergence in probability. □

Theorem 6.2 (Itô's formula). *Let $X^1,\dots,X^d$ be continuous semimartingales and $f : \mathbb{R}^d \to \mathbb{R}$ a C^2 function. Then $\{f(X_t^1,\dots,X_t^d) : t \geq 0\}$ is a continuous semimartingale and up to indistinguishability*

$$\begin{aligned} f(X_t^1,\dots,X_t^d) =& f(X_0^1,\dots,X_0^d) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x^i}(X_s^1,\dots,X_s^d)dX_s^i \\ &+ \frac{1}{2}\sum_{1\leq i,j\leq d} \int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s^1,\dots,X_s^d)d\langle X^i,X^j\rangle_s. \end{aligned} \tag{26}$$

In particular, for $d = 1$, we have

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d\langle X\rangle_s.$$

The last term on the right hand side is called the correction term (from The Fundamental Theorem in Calculus).

Proof. Let $X^i = X_0^i + M^i + A^i$, where M^i are continuous local martingales and A^i are continuous, adapted variational processes, $M_0 = A_0 = 0$. Let V^i denote the total variation process of A^i. Let

$$T_k = \inf\left\{ t \geq 0 : \sum_{i=1}^d \left(|M_t^i| + |X_t^i| + V_t^i + \langle M^i\rangle_t\right) \geq k \right\}$$

for $k = 1,2,\ldots$. Then T_k are stopping times and $T_k \uparrow \infty$. It is sufficient to prove (26) up to time τ_k for each k fixed but any. We will prove that the result holds for polynomials and then the full result follows by approximating C^2 functions by polynomials.

First note that it is obvious that the set of functions for which the formula holds is a vector space containing the functions $f \equiv 1$ and $f(x_1,\ldots,x_d) = x_i$ for $i \leq d$.

We now check that if (26) holds for two functions F and G, then it holds for the product FG. Integration by parts yields

$$F_t G_t - F_0 G_0 = \int_0^t F_s dG_s + \int_0^t G_s dF_s + \langle F,G\rangle_t. \tag{27}$$

By associativity of stochastic integration, and because (26) holds for G,

$$\int_0^t F_s dG_s = \sum_{i=1}^d \int_0^t F(X_s)\frac{\partial G_s}{\partial x^i} dX_s^i + \frac{1}{2}\sum_{1\leq i,j\leq d}\int_0^t F(X_s)\frac{\partial^2 G_s}{\partial x^i \partial x^j} d\langle X^i, X^j\rangle_s,$$

with a similar expression for $\int_0^t G_s dF_s$. Using the fact that (26) holds for F and G, we also have

$$\langle F,G\rangle_t = \sum_{i=1}^d \sum_{j=1}^d \int_0^t \frac{\partial F_s}{\partial x^i}\frac{\partial G_s}{\partial x^j} d\langle X^i, X^j\rangle_s.$$

Substituting these into (27), we obtain Itô's formula for FG. To pass to a general C^2 function f, using Taylor theorem, allowing us to approximate the second derivative of F uniformly on compacts by a polynomial (and hence F' and F are also uniformly approximated on compacts). Using the dominated convergence theorem (and the fact that everything is nicely bounded up to time T_k), we have the result up to time T_k, and then we send $k \to \infty$. □

6.2 Applications of Itô's formula

As a first application of this, suppose that M is a continuous local martingale and A is a continuous, adapted process of finite variation. Then $\langle M,A\rangle \equiv 0$ and applying Itô's formula with $X^1 = M$ and $X^2 = A$ yields

$$f(M_t,A_t) - f(M_0,A_0) = \int_0^t \frac{\partial f}{\partial m}(M_s,A_s)dM_s + \int_0^t \frac{\partial f}{\partial a}(M_s,A_s)dA_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial m^2}(M_s,A_s)d\langle M\rangle_s.$$

Note that this gives us the semimartingale decomposition of $F(M_t,A_t)$ and we can, for example, read off the conditions on F under which we recover a local martingale. In particular, taking $f(x,y) = \exp(\lambda x - \frac{\lambda^2}{2}y)$ with $X^1 = M$ and $X^2 = \langle M,M\rangle$, we obtain:

Proposition 6.3. *Let M be a continuous local martingale and $\lambda \in \mathbb{R}$. Then*

$$\mathscr{E}^{\lambda}(M)_t := \exp\left(\lambda M_t - \frac{\lambda^2}{2}\langle M\rangle_t\right), \quad t \geq 0, \tag{28}$$

is a continuous (non-negative) local martingale. In fact the same holds true for any $\lambda \in \mathbb{C}$ with the real and imaginary parts being local martingales.

Proof. Let $f(x,y) = \exp\left(\lambda x - \frac{\lambda^2}{2}y\right)$, and apply Itô's formula to $\mathscr{E}^{\lambda}(M)_t = f(M_t, \langle M\rangle_t)$. Computing the partial derivatives and simplifying gives:

$$\mathscr{E}^{\lambda}(M)_t = \mathscr{E}^{\lambda}(M)_0 + \int_0^t \frac{\partial f}{\partial x}(M_s, \langle M\rangle_s)dM_s$$

which is a (non-negative) continuous local martingale. □

Note that we have $\frac{\partial}{\partial x}f(x,y) = \lambda f(x,y)$ so that we could have written this as

$$\mathscr{E}^{\lambda}(M)_t = \mathscr{E}^{\lambda}(M)_0 + \lambda \int_0^t \mathscr{E}^{\lambda}(M)_s dM_s$$

or in 'differential form' as

$$d\mathscr{E}^{\lambda}(M)_t = \lambda \mathscr{E}^{\lambda}(M)_t dM_t$$

which shows $\mathscr{E}^{\lambda}(M)$ solves the stochastic exponential differential equation driven by M: $dY_t = \lambda Y_t dM_t$.

6.2.1 Lévy's characterization

Here is a beautiful application of exponential martingales:

Theorem 6.4 (Lévy's characterisation of Brownian motion)**.** *Let M be a continuous local martingale starting at zero. Then M is a standard Brownian motion if and only if $\langle M\rangle_t = t$ a.s. for all $t \geq 0$.*

Proof. We know that the quadratic variation of a Brownian motion B is given by $\langle B\rangle_t = t$.Suppose M is a continuous local martingale starting in zero with $\langle M\rangle_t = t$ a.s. for all $t \geq 0$. Then, by Proposition 6.3,

$$\exp\left(\mathrm{i}\xi M_t + \frac{\xi^2}{2}t\right), \quad t \geq 0$$

is a local martingale for any $\xi \in \mathbb{R}$ and, since it is bounded, it is a martingale. Let $0 \leq s < t$. We have

$$\mathbb{E}\left[\exp\left(\mathrm{i}\xi M_t + \frac{\xi^2}{2}t\right)\Big|\mathscr{F}_s\right] = \exp\left(\mathrm{i}\xi M_s + \frac{\xi^2}{2}s\right)$$

which we can rewrite as

$$\mathbb{E}\left[e^{\mathrm{i}\xi(M_t-M_s)}\middle|\mathscr{F}_s\right]=e^{-\frac{\xi^2}{2}(t-s)}. \tag{29}$$

In other words, $M_t - M_s$ has a normal distribution with mean zero and variance $t-s$. It follows also from (29) that for $A \in \mathscr{F}_s$,

$$\mathbb{E}\left[1_A e^{i\xi(M_t-M_s)}\right]=\mathbb{P}[A]\mathbb{E}\left[e^{i\xi(M_t-M_s)}\right],$$

so fixing $A \in \mathscr{F}_s$ with $\mathbb{P}[A] > 0$ and writing $\mathbb{P}_A = \mathbb{P}[\cdot \cap A]/\mathbb{P}[A]$ (which is a probability measure on $\mathscr{F}_s$) for the conditional probability given A, we have that $M_t - M_s$ has the same distribution under $\mathbb{P}$ as under $\mathbb{P}_A$ and so $M_t - M_s$ is independent of $\mathscr{F}_s$ and we have that M is an $\{\mathscr{F}_t\}_{t\geq 0}$-Brownian motion. □

So the quadratic variation is capturing all the information about M. This is surprising – recall that it is a special property of Gaussians that they are characterised by their means and the variance-covariance matrix, but in general we need to know much more. It also shows we didn't really need the Gaussian assumption in our definition of Brownian motion, it's guaranteed by the independence and variance assumptions.

6.2.2 Dambis–Dubins–Schwarz Theorem

It turns out that what we just saw for Brownian motion has a powerful consequence for all continuous local martingales – they are characterised by their quadratic variation and, in fact, they are all time changes of Brownian motion.

Theorem 6.5 (Dambis–Dubins–Schwarz Theorem)**.** *Let M be a continuous local martingale (with respect to the filtration $(\mathscr{F}_t)_{t\geq 0}$) with $M_0 = 0$ and $\langle M\rangle_\infty = \infty$ a.s. Let $\tau_s := \inf\{t \geq 0 : \langle M\rangle_t > s\}$. Then the process B defined by $B_s := M_{\tau_s}$, is a Brownian motion (with respect to the filtration $(\mathscr{F}_{\tau_t})_{t\geq 0}$) and $M_t = B_{\langle M\rangle_t}$ for every $t \geq 0$.*

Proof. Note that τ_s is the first hitting time of an open set (s,∞) for an adapted process $\langle M\rangle$ with continuous sample paths, and hence τ_s is a stopping time (recall that $(\mathscr{F}_t)_{t\geq 0}$ is right continuous). Further, $\langle M\rangle_\infty = \infty$ implies that $\tau_s < \infty$. The process $(\tau_s : s \geq 0)$ is non-decreasing and right-continuous (in fact $s \to \tau_s$ is the right-continuous inverse of $t \to \langle M\rangle_t$). Let $\mathscr{G}_s := \mathscr{F}_{\tau_s}$. Note that it satisfies the usual conditions. The process B is right continuous by continuity of M and right-continuity of τ. We have

$$\lim_{u\uparrow s} B_u = \lim_{u\uparrow s} M_{\tau_u} = M_{\tau_{s-}}.$$

But$[\tau_{s-}, \tau_s]$ is either a point or an interval of constancy of $\langle M\rangle$. The latter are known (exercise) to coincide a.s. with the intervals of constancy of M and hence

$M_{\tau_{s-}} = M_{\tau_s} = B_s$ so that B has a.s. continuous paths. To conclude that B is a $(\mathscr{G}_s)$-Brownian motion, by Lévy's theorem, it remains to show that (B_s) and $(B_s^2 - s)$ are $(\mathscr{G}_s)$-local martingales.

Note that M^{τ_n} and $(M^{\tau_n})^2 - \langle M \rangle^{\tau_n}$ are uniformly integrable martingales. Taking $0 \le u < s < n$ and applying the Optional Stopping Theorem we obtain

$$\mathbb{E}[B_s|\mathscr{G}_u] = \mathbb{E}[M_{\tau_s}^{\tau_n}|\mathscr{F}_{\tau_u}] = M_{\tau_u}^{\tau_n} = M_{\tau_u} = B_u$$

and

$$\begin{aligned}\mathbb{E}[B_s^2 - s|\mathscr{G}_u] &= \mathbb{E}\left[(M_{\tau_s}^{\tau_n})^2 - \langle M \rangle_{\tau_s}^{\tau_n}|\mathscr{F}_{\tau_u}\right] \\ &= (M_{\tau_u}^{\tau_n})^2 - \langle M \rangle_{\tau_u}^{\tau_n} = (M_{\tau_u})^2 - \langle M \rangle_{\tau_u} = B_u^2 - u,\end{aligned}$$

where we used continuity of $\langle M \rangle$ to write $\langle M \rangle_{\tau_u} = u$. It follows that B is indeed a $(\mathscr{G}_s)$-Brownian motion. Finally, $B_{\langle M \rangle_t} = M_{\tau_{\langle M \rangle_t}} = M_t$, again since the intervals of constancy of M and of $\langle M \rangle$ coincide a.s. so that $s \to \tau_s$ is constant on $[t, \tau_{\langle M \rangle_t}]$. $\square$

7 Appendix

(*Under Construction – I shall update the material through the lectures*). In this section, I shall include a few facts we use frequently in lectures, sometime without mentioning them explicitly.

7.1 Measures and integration

This part can be regarded a short summary for those in learned in Paper A4 and Paper 8.1 (measure theory part).

By a measure on a measurable space $(\Omega, \mathscr{F})$, where $\mathscr{F}$ is a σ-field of some subsets of Ω (in probability theory, Ω is usually called a sample space), is a function $\mu : \mathscr{F} \mapsto [0, \infty]$ such that $\mu(\emptyset) = 0$ and it is countably additive in the sense that

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

for any sequence of events $A_n \in \mathscr{F}$ which are disjoint.

Then $(\Omega, \mathscr{F}, \mu)$ is called a measure space. As a default, by a measure we always mean a σ-finite measure: there a sequence B_n such that $\bigcup_{n=1}^{\infty} B_n = \Omega$ and $\mu(B_n) = 1$. A probability on $(\Omega, \mathscr{F})$ is a measure, often denoted by $\mathbb{P}$, such that $\mathbb{P}(\Omega) = 1$.

As a standard procedure, given a measure space $(\Omega, \mathscr{F}, \mu)$ we can always enlarge the σ-algebra $\mathscr{F}$ to be a σ-algebra $\mathscr{F}^\mu$ and extend the definition (in a unique way) of measure μ to a measure (still denoted by μ) on $\mathscr{F}^\mu$ so that the new measure restricted on $\mathscr{F}$ coincides with μ. The procedure can be described as the following. First define the outer measure $\mu^\star$ induced by μ:

$$\mu^\star(E) := \inf\left\{\sum_{i=1}^{\infty} \mu(A_i) : \text{ where } A_i \in \mathscr{F} \text{ and } \bigcup_{i=1}^{\infty} A_i\right\}$$

for *any subset* $E \subset \Omega$. Clearly $\mu^\star(\emptyset) = 0$ and $\mu^\star$ is countably sub-additive:

$$\mu^\star\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu^\star(A_i)$$

for any sequence of subsets $A_i \subset E$. The next step is to identify the collection, denoted by $\mathscr{F}^\mu$, of all possible $\mu^\star$-measurable subsets. Namely we say E is $\mu^\star$-measurable if

$$\mu^\star(F) = \mu^\star(E \cap F) + \mu^\star(E^c \cap F) \quad \text{for any subset } F \subset \Omega.$$

A major achievement in the measure theory is the following:

1) $\mathscr{F}^\mu$ is a σ-algebra and $\mathscr{F} \subseteq \mathscr{F}^\mu$;

2) if $\mu^\star(E) = 0$, then $E \in \mathscr{F}^\mu$;

3) the outer measure $\mu^\star$ (which is defined for every subset of Ω) restricted on $\mathscr{F}^\mu$ is a measure;

4) $\mu^\star(A) = \mu(A)$ for every $A \in \mathscr{F}$, so $\mu^\star$ is really an extension of μ;

5) $(\Omega, \mathscr{F}^\mu, \mu^\star)$ is complete in the sense that, if $\mu^\star(E) = 0$, then $E \in \mathscr{F}^\mu$.

Therefore, with good reason, we shall still denote $\mu^\star$ restricted on $\mathscr{F}^\mu$ by μ. The measure space $(\Omega, \mathscr{F}^\mu, \mu)$ is called the completion of $(\Omega, \mathscr{F}, \mu)$. You only need to make completion of a measure space once, that is, $(\mathscr{F}^\mu)^{\mu^\star} = \mathscr{F}^\mu$.

Since the completion of a measure space is universal, as a default, a measure μ on $(\Omega, \mathscr{F})$ is *automatically extended to* its completion, so we do not need to worry about the measuriability of null sets[60].

It is very important to recognize that the *concept of measurability* (hence the concept of random variables) depends only on the σ-algebras involved, and is independent of measures, probabilities involved. You should review the concept of measurable mappings between two measurable spaces, which gives the most general concept of measurability, yet quite elementary[61].

The integration theory (Lebesgue's integration) is another standard construction over a measure space $(\Omega, \mathscr{F}, \mu)$. First define the collection of all non-negative simple $\mathscr{F}$-measurable (real) functions, denoted by $S(\mathscr{F})_+$, that is, a function $\phi : \Omega \mapsto \mathbb{R}$ belongs to $S(\mathscr{F})_+$ if $\phi = \sum_{i=1}^m c_i 1_{A_i}$ for some $m \in \mathbb{N}$, constants $c_i \geq 0$ and some $A_i \in \mathscr{F}$. Then the integral of ϕ against μ is defined by

$$\int_\Omega \phi d\mu = \sum_{i=1}^m c_i \mu(A_i)$$

[60] While in this procedure, we deal with only one σ-algebra with one measure. However we often need to handle several, even a family of σ-algebras, or several measures, in this kind of situations, it should be made clear which σ-algebra against which measure you are applying the construction of the completion.

[61] Given two measurable spaces $(\Omega, \mathscr{F})$ and $(E, \mathscr{E})$, a mapping $X : \Omega \mapsto E$ is measurable, with respect to σ-algebras $\mathscr{F}/\mathscr{E}$, if $X^{-1}(A) \in \mathscr{F}$ for every $A \in \mathscr{E}$. That is the pull-back σ-algebra $X^{-1}(\mathscr{E}) \subset \mathscr{F}$. This definition depends on both the σ-algebras on the initial space and the target space of the mapping.

which is independent of the representation of ϕ (as long as we insist that in a representation c_i are non-negative!) If f is measurable and non-negative, then its Lebesgue integral is defined by

$$\int_\Omega f d\mu = \sup\left\{\int_\Omega \phi d\mu : \text{ where } \phi \in S(\mathscr{F})_+ \text{ and } f \leq \phi\right\}.$$

The most fundamental fact, not trivial, is that the sup that defines the integral $\int_\Omega f d\mu$ is indeed additive in f, this justifies the definition of integration.

If f is measurable, then $f = f^+ - f^-$, where $f^+ = f \vee 0$ and $f^- = (-f) \vee 0$ which are measurable and non-negative. Then we say f is integrable with respect to μ (also called μ-integrable) if both integrals $\int_\Omega f^+ d\mu < \infty$ and $\int_\Omega f^- d\mu < \infty$, and $\int_\Omega f d\mu = \int_\Omega f^+ d\mu - \int_\Omega f^- d\mu$. The space of all inegrable functions on $(\Omega, \mathscr{F}, \mu)$ is denoted by $L^1(\Omega, \mathscr{F}, \mu)$. Since $|f| = f^+ + f^-$, therefore f is μ-integrable if 1) f is measurable; 2) $\int_\Omega |f| d\mu < \infty$.

The notation $\int_\Omega f d\mu$, as long as it makes sense, may be denoted also by $\int_\Omega f(x)\mu(dx)$, $\mu(f)$, $\mathbb{E}^\mu[f]$ etc.

Theorem 7.1. *Let $\{f_n : n = 1, 2, \ldots\}$ be a sequence of measurable functions.*

*1) (*Fatou's lemma*) If f_n are non-negative, then*

$$\int_\Omega \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_\Omega f_n d\mu.$$

In particular, if $\liminf_{n\to\infty} \int_\Omega f_n d\mu < \infty$, then $\liminf_{n\to\infty} f_n$ is integrable (hence finite almost surely).

*2) (*Levi, MCT*) If f_n are non-negative and $f_n \leq f_{n+1}$ (almost surely) for every n, then*

$$\int_\Omega \lim_{n\to\infty} f_n d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu.$$

*3) (*Dominated Convergence Theorem*) Suppose $f_n \to f$ as $n \to \infty$ almost surely, and there is an integral function g such that $|f_n| \leq g$ almost surely for every n, then $\int_\Omega f d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu$.*

DCT is useful, while we can do better, and a final result of convergence in L^1 can be achieved by the notion of uniform integrability.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. Then a family $\{\xi_\alpha : \alpha \in \Lambda\}$ of integrable (generalized real) random variables is uniformly integrable if

$$\sup_{\alpha \in \Lambda} \mathbb{E}\left[|\xi_\alpha| : |\xi_\alpha| \geq L\right] \to 0 \quad \text{as } L \to \infty.$$

This is equivalent to the following ε-δ definition. $\{\xi_\alpha : \alpha \in \Lambda\}$ is integrable if the following two conditions are satisfied:

(i) $\{\xi_\alpha : \alpha \in \Lambda\}$ is bounded in L^1, i.e. $\sup_\alpha \mathbb{E}[|\xi_\alpha|] < \infty$;

(ii) for every $\varepsilon > 0$, there is $\delta > 0$ such that for any $A \in \mathscr{F}$ such that $\mathbb{P}[A] \leq \delta$ we have

$$\mathbb{E}[|\xi_\alpha| : A] \leq \varepsilon \quad \text{for every } \alpha \in \Lambda.$$

The following facts are very useful.

1) Suppose that there is an integrable η such that $|\xi_\alpha| \leq \eta$ for every $\alpha \in \Lambda$, then $\{\xi_\alpha : \alpha \in \Lambda\}$ is uniformly integrable.

2) If there is $p > 1$, $\{\xi_\alpha : \alpha \in \Lambda\}$ is bounded in L^p-space, i.e. $\sup_\alpha \mathbb{E}[|\xi_\alpha|^p] < \infty$.

3) If $\xi_\alpha = \mathbb{E}[\xi|\mathscr{G}_\alpha]$ (for $\alpha \in \Lambda$) where ξ is integrable and $\mathscr{G}_\alpha$ are σ-algebras, then $\{\xi_\alpha : \alpha \in \Lambda\}$ is uniformly integrable.

Lemma 7.2. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let ξ_n, ξ be integrable.*

1) $\xi_n \to \xi$ in L^1 if and only if $\{\xi_n; n \geq 1\}$ is uniformly integrable and $\xi_n \to \xi$ in probability as $n \to \infty$.

*2) (*Scheffé's lemma*) If $\xi_n \to \xi$ almost surely as $n \to \infty$, then $\xi_n \to \xi$ in L^1 if and only if $\mathbb{E}[|\xi_n|] \to \mathbb{E}[|\xi|]$ as $n \to \infty$.*

Finally, for $p \geq 1$, we define $L^p(\Omega, \mathscr{F}, \mu)$ to be the space of all measurable functions f such that $|f|^p$ is integrable. For any measurable function f, define

$$\|f\|_{L^p(\Omega,\mathscr{F},\mu)} = \left(\int_\Omega |f|^p d\mu \right)^{\frac{1}{p}}$$

which is denoted by $\|f\|_p$ if the measure space in use is clear. Then a measurable function $f \in L^p(\Omega, \mathscr{F}, \mu)$ if and only if $\|f\|_p < \infty$. By identifying two functions which are equal almost surely as the same element in $L^p(\Omega, \mathscr{F}, \mu)$, then $L^p(\Omega, \mathscr{F}, \mu)$ equipped with the norm $\|\cdot\|_p$ is a complete normed space (Banach space).

7.2 Martngale inequalities, convergence theorem, in discrete-time

If f is a convex function on (a,b), that is, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$ for any $x, y \in (a,b)$ and $0 \leq \lambda \leq 1$, then f must be continuous on (a,b), and f has left-hand derivative $f'_-(x) = \sup_{y<x} \frac{f(y)-f(x)}{y-x}$ and right-hand derivative $f'_+(x) = \inf_{y>x} \frac{f(y)-f(x)}{y-x}$ at every $x \in (a,b)$. Moreover $f'_-(x) \leq f'_+(x)$ for every $x \in (a,b)$ and both function $x \mapsto f'_-(x)$ and $x \mapsto f'_+(x)$ are non-decreasing, and

$$f(x) \geq f'_+(x_0)(x - x_0) + f(x_0) \tag{30}$$

for any $x, x_0 \in (a,b)$ (and in fact here $f'_+(x_0)$ can be replace by any number $A \in [f'_-(x_0), f'_+(x_0)]$.[62]

[62]These properties of a convex function can be figured out by drawing a sketch, and observing that convexity means the "slop" $\frac{f(y)-f(x)}{y-x}$ (where $y < x$) is non-decreasing when the interval (x,y) moves towards in the right-hand direction.

Suppose now ξ is a random variable taking values in (a,b) on a probability space $(\Omega,\mathscr{F},\mathbb{P})$, and f is a convex function. Suppose both ξ and $f(\xi)$ are integrable. Let $\mathscr{G}\subset\mathscr{F}$ be any sub σ-algebra. Set $x_0=\mathbb{E}[\xi|\mathscr{G}]$, and $x=\xi$ in (30). We obtain

$$f(\xi)\geq f_+^{'}(\mathbb{E}[\xi|\mathscr{G}])(\xi-\mathbb{E}[\xi|\mathscr{G}])+f(\mathbb{E}[\xi|\mathscr{G}]).$$

Note that $f_+^{'}$ is non-decreasing, so it is Borel measurable, thus $f_+^{'}(\mathbb{E}[\xi|\mathscr{G}])$ is $\mathscr{G}$ measurable. Taking conditional expectation $\mathbb{E}[\cdot|\mathscr{G}]$ (which preverses the inequality $\geq$) both sides of the previous inequality, we therefore have

$$\begin{aligned}\mathbb{E}[f(\xi)|\mathscr{G}]&\geq\mathbb{E}\left[f_+^{'}(\mathbb{E}[\xi|\mathscr{G}])(\xi-\mathbb{E}[\xi|\mathscr{G}])|\mathscr{G}\right]+f(\mathbb{E}[\xi|\mathscr{G}])\\&=f_+^{'}(\mathbb{E}[\xi|\mathscr{G}])\mathbb{E}[(\xi-\mathbb{E}[\xi|\mathscr{G}])|\mathscr{G}]+f(\mathbb{E}[\xi|\mathscr{G}])\\&=f(\mathbb{E}[\xi|\mathscr{G}])\end{aligned}$$

This proves the Jensen inequality for conditional expectation.

Lemma 7.3. *Suppose f is convex on (a,b), and ξ is an integrable random variable taking values in (a,b). Suppose $f(\xi)$ is integrable , then*

$$f(\mathbb{E}[\xi|\mathscr{G}])\leq\mathbb{E}[f(\xi)|\mathscr{G}] \tag{31}$$

where $\mathscr{G}\subset\mathscr{F}$ is a σ-algebra.

You should know some examples of convex functions.

1) If f has twice derivatives on (a,b), then it is convex if and only if $f^{''}(x)\geq 0$ for $x\in(a,b)$.

2) If f is differentiable on (a,b), then it is convex if and only if $x\mapsto f^{'}(x)$ is non-decreasing.

3) $f(x)=|x|^p$ is convex on $(-\infty,\infty)$ if $p\geq 1$.

Applying Jensen's inequality with the definition of martingales (sub-martingales, super-martingales), we therefore have the flowing consequences.

Let $(\Omega,\mathscr{F},(\mathscr{F}_n),\mathbb{P})$ be a filtered probability space.

1) If $M=(M_n)$ is a martingale, and f is a convex function $\mathbb{R}$, and if $f(M_n)$ is integrable for every n, then $(f(M_n))$ is a sub-martingale (in particular, $n\mapsto\mathbb{E}[f(M_n)]$ is increasing).

2) If $X=(X_n)$ is a sub-martingale taking values in some interval (a,b), and if f is *convex and non-decreasing* on (a,b), and suppose $f(X_n)$ is integrable for each n, then $(f(X_n))$ is also a sub-martingale.

3) If $M=(M_n)$ is a martingale, $p\geq 1$, and $\mathbb{E}[|M_n|^p]<\infty$ for every n, then $\{|M_n|^p;n\geq 0\}$ is a non-negative sub-martingale.

4) If $X=(X_n)$ is a sub-martingale, then its positive part $\{X_n^+=X_n\vee 0;n\geq 0\}$ is a non-negative sub-martingale. If $X_n\ln^+(X_n)$ is integrable for every n, then $\{X_n\ln^+(X_n);n\geq 0\}$ is a sub-martingale, where $\ln^+(x)=\ln x$ for $x\geq 1$ and $\ln^+(x)=0$ if $x<1$, so that $x\ln^+(x)$ is convex.

The theory of martingales in discrete-time essentially consists of several fundamental theorems about sub-, super-martingales: Doob's optional (stopping) theorem, basic martingale inequalities, and Doob's martingale convergence theorem, which we shall state as the following for your revision for the paper B8.1.

First recall that a random variable $T : \Omega \mapsto \mathbb{Z}_+ \cup \{\infty\}$ is an $(\mathscr{F}_n)$-stopping time[63], if $\{T = n\} \in \mathscr{F}_n$ for every $n = 0, 1, \cdots$.

Theorem 7.4. *If $X = (X_n)$ is a sub-martingale and $T \geq S$ are two* bounded *stopping times, then $\mathbb{E}[X_T|\mathscr{F}_S] \geq X_S$. There are similar statements for martingales, and for super martingales. Therefore, sub-martingale (super-martingale, martingale) property is valid at bounded stopping times. In particular, a stopped martingale (resp. sub-martingale; super-martingale at a stopping time) is a martingale (resp. sub-martingale; super-martingale at a stopping time).*

Theorem 7.5. *If $X = (X_n)$ is a non-negative sub-martingale, then for every $N \in \mathbb{Z}_+$*

1) (Doob's maximal inequality)

$$\mathbb{P}\left[\max_{n \leq N} X_n \geq \lambda\right] \leq \frac{1}{\lambda}\mathbb{E}X_N$$

for every $\lambda > 0$;

2) For every $p > 1$

$$\mathbb{P}\left[\left(\max_{n \leq N} X_n\right)^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}|X_N|^p.$$

There is a discrete-time version for super-martingale.

Theorem 7.6. *[*Doob's maximal inequality*] Suppose $M = (M_n)_{n \geq 0}$ is a super-martingale w.r.t. a filtration $(\mathscr{F}_n)_{n \geq 0}$. Then (Doob's maximal inequality)*

$$\mathbb{P}\left[\sup_{k \leq N} M_k \geq \lambda\right] \leq \frac{1}{\lambda}\left(\mathbb{E}[M_0] - \mathbb{E}\left[M_N : \sup_{k \leq N} M_k < \lambda\right]\right);$$

$$\mathbb{P}\left[\inf_{k \leq N} M_k \leq -\lambda\right] \leq \frac{1}{\lambda}\mathbb{E}\left[-M_N : \inf_{k \leq N} M_k \leq -\lambda\right]$$

and

$$\mathbb{P}\left[\sup_{k \leq N} |M_k| \geq \lambda\right] \leq \frac{1}{\lambda}\left(\mathbb{E}[M_0] + 2\mathbb{E}\left[M_N^-\right]\right)$$

for any $\lambda > 0$ and $N \in \mathbb{N}$.

[63] The usefulness of the concept of stopping times lies in the fact that a stopping time T gives rise to a natural partition of the sample space Ω, consistence with a filtration, according to the values of a stopping time, namely, the partition $\{T = n\}$ (where $n = 0, 1, 2, \ldots$) together with $\{T = \infty\}$. The appearance of the concept of stopping times reduced the huge literature by many wisemen to something routine. This simple example demonstrates well the importance of formulating core notions in a subject of science.

Let us recall the definition of an ordered set of real numbers $\{x_0, x_1, \ldots, x_N\}$ of $N+1$ numbers. Let $a < b$ be two numbers. Let

$$T_0 = \inf\{k \geq 0 : x_k < a\},$$

$$T_1 = \inf\{k > T_0 : x_k > b\},$$

$$\ldots\ldots$$

$$T_{2j} = \inf\{k > T_{2j-1} : x_k < a\}$$

and

$$T_{2j+1} = \inf\{k > T_{2j} : x_k > b\}$$

inductively. The up-crossing number

$$U_a^b\left((x_i)_{i=0,\ldots,N}\right) := \max\left\{j : T_{2j-1} \leq N\right\}.$$

If $x = (x_n)_{n\geq 0}$ is a sequence, then

$$U_a^b\left((x_n), N\right) := U_a^b\left((x_i)_{i=0,\ldots,N}\right)$$

and

$$U_a^b\left((x_n)\right) := \lim_{N\to\infty} U_a^b\left((x_n), N\right)$$

which takes values in $\mathbb{Z}_+ \cup \{\infty\}$.

If $X = (X_n)_{n\geq 0}$ is a random sequence, then we define

$$U_a^b(X,N)(\omega) := U_a^b\left((X_n(\omega)), N\right) \quad \text{and } U_a^b(X)(\omega) := U_a^b\left((X_n(\omega))\right).$$

If $X = (X_n)$ is adapted to $(\mathscr{F}_n)$ then T_i are stopping times and

$$\left\{U_a^b(X,N) = j\right\} = \left\{j : T_{2j-1} \leq N < T_{2j}\right\}$$

for every $j \in \mathbb{Z}_+$. Therefore $U_a^b(X,N)$ hence $U_a^b(X)$ are non-negative random variables taking values in $\mathbb{Z}_+ \cup \{\infty\}$.

Lemma 7.7. *Let (x_n) be a sequence of real numbers. Then* $\lim_{n\to\infty} x_n$ *exists, ∞ or $-\infty$*[64]*, if and only if for every pair $a < b$ (where we can add that a, b are rationals) $U_a^b((x_n)) < \infty$.*

Theorem 7.8. *(*Doob's up-crossing lemma*) If $X = (X_n)$ is a super-martingale, and $N \in \mathbb{Z}_+$, then for any pair $a < b$ we have*

$$\mathbb{P}\left[U_a^b(X,N) > j\right] = \frac{1}{b-a}\mathbb{E}\left[(X_N - a)^- : U_a^b(X,N) = j\right]$$

[64]This assumption is equivalent to that $\liminf x_n = \limsup x_n$ (including the case the up/low limits are ∞ and $-\infty$.

for any $j \in \mathbb{Z}_+$*, and therefore*

$$\mathbb{P}\left[U_a^b(X,N)\right] = \frac{1}{b-a}\mathbb{E}\left[(X_N - a)^-\right].$$

Hence

$$\mathbb{P}\left[U_a^b(X)\right] = \frac{\sup_n \mathbb{E}\left[|X_n|\right] + a}{b-a}$$

(where $\sup_n \mathbb{E}\left[|X_n|\right]$ *maybe infinity of course).*

The martingale convergence theorems follow from Doob's up-crossing lemma immediately.

Theorem 7.9. *(*Doob's convergence theorem*) Let* $X = (X_n)$ *be a super-martingale.*

1) If $\sup_n \mathbb{E}\left[|X_n|\right] < \infty$*, i.e.* $\{X_n; n \geq 0\}$ *is bounded in* L^1*-space, then* $X_n \to X_\infty$ *almost surely as* $n \to \infty$*, and* X_∞ *is integrable. Moreover* $X_n \to X_\infty$ *in* L^1 *if and only if* $\{X_n; n \geq 0\}$ *is uniformly integrable. In the latter case,* $\mathbb{E}\left[X_\infty | \mathscr{F}_n\right] \leq X_n$ *(if* X *is a martingale,* $\mathbb{E}\left[X_\infty | \mathscr{F}_n\right] = X_n$*) for every* n*.*

2) Let $p > 1$*. Suppose* $\{X_n; n \geq 0\}$ *is bounded in* L^p*-space*[65]*:* $\sup_n \mathbb{E}\left[|X_n|^p\right] < \infty$*, then* $X_n \to X_\infty$ *almost surely and also in* L^p *as* $n \to \infty$*.*

There is a backward version of the martingale convergence theorem. Let us explain this difficult theorem for super-martingales.

The setting is slightly different though, that is, the time parameter range now is $\mathbb{Z}_- = \{\cdots, -2, -1, 0\}$ and we are dealing with a family of sub σ-algebras $\mathscr{F}_n$ (where $n = 0, -1, -2, \cdots$) such that $\mathscr{F}_m \subset \mathscr{F}_n$ for $m \leq n$, $m, n \in \mathbb{Z}_-$. An adapted and integrable random sequence $X = (X_n)_{n \in \mathbb{Z}_-}$ is a super-martingale with respect to $(\mathscr{F}_n)_{n \leq 0}$ if

$$\mathbb{E}\left[X_n | \mathscr{F}_{n-1}\right] \leq X_{n-1} \quad \text{for } n = 0, -1, -2, \cdots.$$

We may rewrite the definition in an equivalent way in terms of what we are familiar with. Let $N \geq 1$ be any but fixed. Define $\mathscr{G}_k = \mathscr{F}_{-(N-k)}$ and $Y_k = X_{-(N-k)}$, $(k = 0, 1, \cdots, N)$, so that $(\mathscr{G}_k)_{k=0,\cdots,N}$ is a filtration (in the ordinary sense). Then $X = (X_n)_{n \leq 0}$ is super-martingale with respect to $(\mathscr{F}_n)_{n \leq 0}$, is the same as saying that Y is a $(\mathscr{G}_k)$-super-martingale, for every N.

Theorem 7.10. *Let* $X = (X_n)_{n \leq 0}$ *be a super-martingale with respect to* $(\mathscr{F}_n)_{n \leq 0}$*. Suppose* $\lim_{n \to -\infty} \mathbb{E}\left[X_n\right] < \infty$*, then*

1) $\{X_n : n \leq 0\}$ *is uniformly integrable;*

2) $X_n \to X_{-\infty}$ *almost surely and in* L^1*, as* $n \to \infty$*.*

Proof. 1) The most useful part of this theorem is the uniform integrability, so let us provide a careful proof. The key assumption is the existence of the limit $\lim_{n \to \infty} \mathbb{E}\left[X_n\right]$. Since X is a super-martingale, so that $\mathbb{E}\left[X_n\right] \leq \mathbb{E}\left[X_{n-1}\right]$ for every $n = 0, -1, -2, \ldots, n \to \mathbb{E}\left[X_n\right]$ is monotone, and $\mathbb{E}\left[X_n\right] \geq \mathbb{E}\left[X_0\right]$ for all $n = 0, -1, -2,$

[65]This assumption as $p > 1$ implies that $\{X_n; n \geq 0\}$ is uniformly integrable.

..., therefore $\lim_{n\to-\infty}\mathbb{E}[X_n]$ exists or ∞. Therefore the assumption ensures that $\lim_{n\to-\infty}\mathbb{E}[X_n]$ exists and finite. Hence $\{\mathbb{E}[X_n] : n = 0,-1,-2,\ldots\}$ is a Cauchy sequence. Therefore for every $\varepsilon > 0$, there is N, such that for any $n,m \le -N$, we have $|\mathbb{E}(X_n) - \mathbb{E}(X_m)| < \frac{\varepsilon}{2}$. For every $C > 0$ and $n < -N$ we have

$$\begin{aligned}
\mathbb{E}[|X_n| : |X_n| \ge C] &= \mathbb{E}[X_n : X_n \ge C] - \mathbb{E}[X_n : X_n \le -C] \\
&= \mathbb{E}[X_n] - \mathbb{E}[X_n : X_n < C] - \mathbb{E}[X_n : X_n \le -C] \\
&\le \mathbb{E}[X_n] - \mathbb{E}[X_{-N} : X_n < C] - \mathbb{E}[X_{-N} : X_n \le -C] \\
&= \mathbb{E}[X_n] - \mathbb{E}[X_{-N}] - \mathbb{E}[X_{-N} : X_n \ge C] - \mathbb{E}[X_{-N} : X_n \le -C] \\
&= \mathbb{E}[X_n] - \mathbb{E}[X_{-N}] - \mathbb{E}[X_{-N} : |X_n| \ge C]
\end{aligned}$$

where the inequality follows from the super-martingale property:

$$\mathbb{E}[X_n : X_n < C] \ge \mathbb{E}[X_{-N} : X_n < C]$$

and

$$\mathbb{E}[X_n : X_n \le -C] \ge \mathbb{E}[X_{-N} : X_n \le -C]$$

for any $n \le -N$. Since $\mathbb{E}[X_n] - \mathbb{E}[X_{-N}] < \frac{\varepsilon}{2}$ for $n \le -N$, so that

$$\mathbb{E}[|X_n| : |X_n| \ge C] < \frac{\varepsilon}{2} - \mathbb{E}[X_{-N} : |X_n| \ge C]$$

for all $n \le -N$ and C. Now

$$\begin{aligned}
\mathbb{P}[|X_n| \ge C] &\le \frac{1}{C}\mathbb{E}\left(X_n^+ + X_n^-\right) = \frac{1}{C}\mathbb{E}\left(X_n + 2X_n^-\right) \\
&= \frac{1}{C}\left(\mathbb{E}[X_n] + 2\mathbb{E}\left[X_n^-\right]\right) \\
&\le \frac{1}{C}\left(\lim_{m\to-\infty}\mathbb{E}[X_m] + 2\mathbb{E}\left[X_0^-\right]\right)
\end{aligned}$$

for all $n \le -N$, here we have used the facts that $\lim_{m\to-\infty}\mathbb{E}[X_m] = \sup_k \mathbb{E}[X_k]$ and X^- is a backward sub-martingale. It follows that

$$\mathbb{P}[|X_n| \ge C] \to 0$$

as $C \to \infty$, uniformly in $n \le -N$. Since X_{-N} (single integrable random variable) is uniformly integrable, hence there $C_0 > 0$ such that

$$\mathbb{E}[|X_{-N}| : |X_n| \ge C] \le \frac{\varepsilon}{2} \quad \text{for all } C > C_0$$

and for all $n \le -N$ (cf. Question 2, Problem Sheet 1). It follows that

$$\mathbb{E}[|X_n| : |X_n| \ge C] < \varepsilon \quad \text{for all } n \le -N \quad \text{and } C > C_0.$$

Since $\{X_{-N}, X_{-N+1}, \cdots, X_0\}$ is a finite family of integrable random variables, so it is uniformly integrable. Hence there is $C_1 \ge C_0$, such that

$$\mathbb{E}[|X_n| : |X_n| \ge C] < \varepsilon \quad \text{for all } n = 0,-1,\cdots,-N \quad \text{and } C > C_0.$$

Putting together we deduce that

$$\lim_{C\to\infty}\sup_{n=0,-1,\cdots}\mathbb{E}\left[|X_n| : |X_n|\geq C\right]=0$$

and therefore $\{X_n : n=0,-1,-2,\cdots\}$ is uniformly integrable.

2) The proof of the convergence follows from the up-crossing estimate. Since, by Doob's up-crossing lemma

$$U_a^b((X_{-N},\cdots,X_0))\leq\frac{\mathbb{E}\left[|X_0|+|a|\right]}{b-a}$$

for every N, and $a<b$, so that

$$U_a^b(X)\leq\frac{\mathbb{E}\left[|X_0|+|a|\right]}{b-a}<\infty.$$

Therefore $\lim_{n\to-\infty}X_n=X_{-\infty}$ almost surely (but $X_{-\infty}$ maybe ∞ or $-\infty$). However X is uniformly integrable, so that $X_{-\infty}$ must be integrable and the convergence is in L^1 as well. □

Let us now review the optional stopping time theorem in discrete setting. First of all, we know it is easy to devise this theorem for bounded stoping times.

Theorem 7.11. *Let $X=(X_n)_{n\geq 0}$ be a martingale (resp. a super-martingale) on a filtered probability space $(\Omega,\mathscr{F},(\mathscr{F}_n),\mathbb{P})$, and $S\leq T$ be two bounded stopping times, then $\mathbb{E}\left[X_T|\mathscr{F}_S\right]=X_S$.*

This theorem can be extended to uniformly integrable martingales (resp. super-martingale which is closed from right side).

Theorem 7.12. *Let $X=(X_n)_{n\geq 0}$ be a uniformly integrable martingale[66] on a filtered probability space $(\Omega,\mathscr{F},(\mathscr{F}_n),\mathbb{P})$, and S,T be two stopping times. Then X_T is integrable and $\mathbb{E}\left[X_T|\mathscr{F}_S\right]=X_{T\wedge S}$, where $X_\infty=\lim_{n\to\infty}X_n$.*

Proof. [A good exercise for revising the optional stopping time theorem]. For every n but fixed, we apply the stopping time theorem to the following. $\mathscr{G}_k=\mathscr{F}_k$ if $k\leq n$, $\mathscr{G}_k=\mathscr{F}_\infty$ for $k\geq n+1$, $\tilde{X}_k=X_k$ for $k\leq n$ and $\tilde{X}_k=X_\infty$ for $k\geq n+1$. Then $\tilde{X}$ is a uniformly integrable martingale. Let $\tilde{S}=S$ if $S\leq n$ and $\tilde{S}=n+1$ if $S>n$. Then $\tilde{S}$ is a bounded stopping time, so that

$$\mathbb{E}\left[\tilde{X}_{n+1}|\mathscr{G}_{\tilde{S}}\right]=\mathbb{E}\left[X_\infty|\mathscr{G}_{\tilde{S}}\right]=\tilde{X}_{\tilde{S}}.$$

While by definition, $\mathscr{G}_{\tilde{S}}=\mathscr{F}_{S_n}$ and $\tilde{X}_{\tilde{S}}=X_{S_n}$, where $S_n=S1_{\{S\leq n\}}+\infty 1_{\{S>n\}}$. Hence

$$\mathbb{E}\left[X_\infty|\mathscr{F}_{S_n}\right]=X_{S_n}$$

66 According to Doob's martingale convergence theorem, this condition implies that $X_\infty=\lim_{n\to\infty}X_n$ almost surely and in L^1.

for every n. On the other hand it can be verified that $\mathscr{F}_S \cap \{S = S_n\} = \mathscr{F}_{S_n} \cap \{S = S_n\}$, so that

$$\mathbb{E}[X_\infty|\mathscr{F}_S]\,1_{\{S=S_n\}} = \mathbb{E}[X_\infty|\mathscr{F}_{S_n}]\,1_{\{S=S_n\}} = X_{S_n}1_{\{S=S_n\}}.$$

Letting $n \to \infty$, and using the fact that $S_n \to S$ to obtain that bounded,

$$\mathbb{E}[X_\infty|\mathscr{F}_S] = X_S.$$

Therefore X_S is integrable. We also have $\mathbb{E}[X_\infty|\mathscr{F}_T] = X_T$, hence, if $T \geq S$, then

$$\mathbb{E}[X_T|\mathscr{F}_S] = \mathbb{E}[\mathbb{E}[X_\infty|\mathscr{F}_T]\,\mathscr{F}_S] = \mathbb{E}[X_\infty|\mathscr{F}_S] = X_S.$$

In general

$$\begin{aligned}\mathbb{E}[X_T|\mathscr{F}_S] &= \mathbb{E}\left[X_T 1_{\{T\leq S\}} + X_{T\vee S}1_{\{T>S\}}|\mathscr{F}_S\right] \\ &= X_T 1_{\{T\leq S\}} + \mathbb{E}[X_{T\vee S}|\mathscr{F}_S]\,1_{\{T>S\}} \\ &= X_T 1_{\{T\leq S\}} + X_S 1_{\{T>S\}} = X_{T\wedge S}\end{aligned}$$

which completes the proof. □

Similarly, for super-martingales we have

Theorem 7.13. *Suppose $X = (X_n)_{n\geq 0}$ is a super-martingale, and suppose there is an integrable, $\mathscr{F}_\infty$ measurable random variable X_∞ such that*[67] *$X_n \geq \mathbb{E}[X_\infty|\mathscr{F}_n]$ for every n. Let $S \leq T$ be two stopping times. Then X_T and X_S are integrable and $\mathbb{E}[X_T|\mathscr{F}_S] \leq X_S$*

Proof. Let $Y_n = \mathbb{E}[X_\infty|\mathscr{F}_n]$ (for $n \geq 0$) which is a uniformly integrable martingale, and $Y_\infty = \lim Y_n = 0$. Let

$$Z_n := X_n - Y_n = X_n - \mathbb{E}[X_\infty|\mathscr{F}_n] \geq 0.$$

Thus $Z = (Z_n)$ is a non-negative super-martingale. By stopping time theorem for bounded stopping times we have

$$\mathbb{E}[Z_{S\wedge n}] \leq \mathbb{E}[Z_0] \quad \text{for any } n$$

Letting $n \to \infty$, Fatou's lemma implies that $\mathbb{E}[Z_S] \leq \mathbb{E}[Z_0]$ so Z_S is integrable, and therefore $X_S = Z_S + Y_S$ is integrable. Applying stopping time theorem for bounded stopping times (like in the proof of the previous theorem)

$$Z_{S_n} \geq \mathbb{E}[Z_{T_n}|\mathscr{F}_{S_n}]$$

where $S_n = S1_{\{S\leq n\}} + \infty 1_{\{S>n\}}$. Since

$$\mathbb{E}[Z_{T_n}|\mathscr{F}_S]\,1_{\{S=S_n\}} = \mathbb{E}[Z_{T_n}|\mathscr{F}_{S_n}]\,1_{\{S=S_n\}} \leq Z_S 1_{\{S=S_n\}}$$

[67] That is, X is a right-hand side closed super-martingale.

for every n, letting $n \to \infty$ and using Fatou's lemma

$$\mathbb{E}[Z_T|\mathscr{F}_S] \leq \lim_{n\to\infty} \mathbb{E}[Z_{T_n}|\mathscr{F}_S] 1_{\{S=S_n\}} \leq Z_S.$$

Since $\mathbb{E}[Y_T|\mathscr{F}_S] = Y_S$, so that

$$\mathbb{E}[X_T|\mathscr{F}_S] = \mathbb{E}[Y_T + Z_T|\mathscr{F}_S] \leq Y_S + Z_S = X_S$$

which completes the proof. □

Remark. Under the conditions in the theorem, i.e. $X = (X_n)_{n\geq 0}$ is a right-hand side closed super-martingale: $\mathbb{E}[X_n|\mathscr{F}_m] \leq X_m$ for any $m \leq n \in \mathbb{Z}_+ \cup \{\infty\}$, then $\{X_n^- : n \geq 0\}$ is uniformly integrable (which is also sub-martingale). In fact

$$X_n \geq \mathbb{E}[X_\infty|\mathscr{F}_n] = \mathbb{E}\left[X_\infty^+ - X_\infty^-|\mathscr{F}_n\right] \geq -\mathbb{E}\left[X_\infty^-|\mathscr{F}_n\right]$$

so that

$$X_n^- \leq \mathbb{E}\left[X_\infty^-|\mathscr{F}_n\right] \quad \text{for } n \geq 0.$$

Therefore $\{X_n^- : n \geq 0\}$ is a uniformly integrable sub-martingale.

7.3 Lévy's construction of Brownian motion

This will be not covered in lectures, for your readings [*Not Examinable*].

There is a beautiful direct construction of Brownian motion by Lévy, improved by Ciesielski.

The following lemma establishes some classical and useful properties of normal distributions. Its proof is left as an exercise.

Lemma 7.14. *i. Let Z, Z' be independent random variables with $Z \sim N(\mu, \Sigma)$, $Z' \sim N(\mu', \Sigma')$. Then $Z + Z' \sim N(\mu + \mu', \Sigma + \Sigma')$. Equivalently, their densities satisfy the convolution property*

$$\int_{\mathbb{R}^d} \phi_{(\mu,\Sigma)}(y)\phi_{(\mu',\Sigma')}(x-y)dy = \phi_{(\mu+\mu',\Sigma+\Sigma')}(x).$$

ii. If $Z_i \sim N(\mu_i, \Sigma_i)$ is a sequence of independent normal random variables such that $\mu^* = \sum_{i\in\mathbb{N}} \mu_i$ and $\Sigma^* = \sum_{i\in\mathbb{N}} \Sigma_i$ exist (i.e. the sums converge), then the sequence of partial sums $\sum_{i=1}^n Z_i$ converges in (L^2, and hence in) probability to a random variable with distribution

$$\sum_{i\in\mathbb{N}} Z_i \sim N(\mu^*, \Sigma^*).$$

iii. If the pair (Z, Z') is a multivariate normal random variable, then Z and Z' are normal, and are independent if and only if their covariance is zero, that is, $E[(Z-\mu)(Z'-\mu')^\top] = 0$.

We begin with a countable family $\{Z_m\}$ of identically distributed random variables with $Z_m \sim N(0, I_d)$ for all m. Let $D_n = \{k2^{-n} : k, n \in \mathbb{Z}^+\}$, so that $D_n \subset D_{n+1}$, $D_0 = \mathbb{Z}^+$ and $\cup_n D_n$ is the set of Dyadic rationals. For simplicity of notation, let $\{Z_m\}$ be indexed by $m \in \cup_n D_n$ and $Z_0 := 0$.

We proceed as follows: First, we determine the value of the nth approximation X^n (of Brownian motion) on the points D_n. Second, we use linear interpolation to define X_t^n for all values of t. This gives us a sequence of paths which we shall show converge.

To fix the values of X_t^n for $t \in D_n$, we define

$$X_t^0 = \sum_{\{k \in D_0 : k < t\}} Z_k.$$

Next, for every $n > 0$, define $X_t^n = X_t^{n-1}$ for all $t \in D_{n-1}$. For $t \in D_n \setminus D_{n-1}$, let

$$X_t^n = X_t^{n-1} + 2^{-(n/2+1)} Z_t. \tag{32}$$

We now linearly interpolate between these points $\{X_t^n\}_{t \in D_n}$. Formally, we can write the interpolation step as

$$X_t^n = X_{\lfloor t \rfloor_n} + \frac{t - \lfloor t \rfloor_n}{\lceil t \rceil_n - \lfloor t \rfloor_n} (X_{\lceil t \rceil_n} - X_{\lfloor t \rfloor_n}),$$

where $\lfloor t \rfloor_n = \max\{s \in D_n : s \leq t\}$, $\lceil t \rceil_n = \min\{s \in D_n : s \geq t\}$. The use of linear interpolation is not vital to the construction, as we shall see (taking right-continuous step functions $X_t^n := X_{\lfloor t \rfloor_n}^n$ would work just as well for proving the existence of a limit, but would not immediately give continuity). We now seek to show that these paths converge, in a sufficiently strong sense, to a Brownian motion.

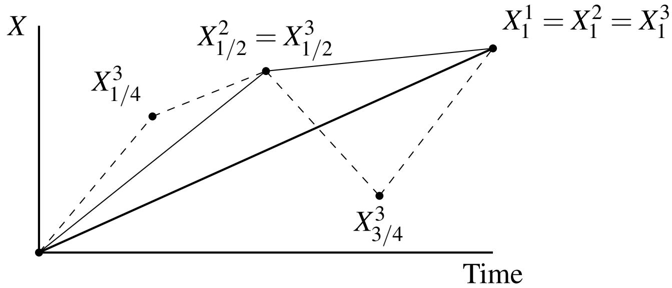

Figure 1: Three steps in Lévy's construction

Lemma 7.15. *Let $\{X^n\}_{n \in \mathbb{N}}$ be a sequence of a.s. continuous functions which converge uniformly on compacts in probability to a process X, that is, for any $\varepsilon > 0$,*

$$\lim_n \mathbb{P}\left[\sup_{s \in [0,t]} \|X_s^n - X_s\| < \varepsilon\right] = 1$$

for all t. Then X is also continuous.

Proof. For fixed t, taking a sub-sequence in n, we can assume that the convergence is almost sure, that is,

$$\mathbb{P}\left[\lim_n \sup_{s\in[0,t]} \|X_s^n - X_s\| = 0\right] = 1$$

Fixing ω, this is a statement of uniform convergence of $X^{n_j} \to X$, and the continuity of the limit is classical, as for any $\varepsilon > 0$, we can find $\delta, m > 0$ such that

$$\begin{aligned}
\|X_s - X_{s+\delta}\| &\leq \|X_s^{n_m} - X_s\| + \|X_{s+\delta}^{n_m} - X_{s+\delta}\| + \|X_s^{n_m} - X_{s+\delta}^{n_m}\| \\
&\leq 2 \sup_{r\in[0,t]} \|X_r^{n_m} - X_r\| + \|X_s^{n_m} - X_{s+\delta}^{n_m}\| \\
&\leq 3\varepsilon.
\end{aligned}$$

□

Remark 7.16. *The uncountable supremum in the statement of Lemma 7.15 is measurable, as our functions are continuous (so the supremum could equally be taken over the rationals, and suprema over countable sets are always measurable).*

Theorem 7.17. *The processes X^n defined in* (32) *converge a.s. uniformly on compacts to a process X. In its natural filtration, the limit is a Brownian motion starting at zero.*

Proof. Convergence. We first show that the processes converge. We consider the case where X is a Brownian motion in two dimensions, as this implies all other cases by the triangle inequality, and is notationally simpler. From our construction, we can see that

$$\sup_{s\in[0,t]} \|X_s^n - X_s^{n+1}\| = \max_{\{s\in D_{n+1}\setminus D_n : s<t\}} \|2^{-(n/2+1)} Z_s\|.$$

The set $\{s \in D_{n+1} \setminus D_n : s < t\}$ contains at most $t2^n$ elements, and the Z_s are independent $N(0, I_d)$ random variables. It is standard that $\|Z_s\|^2$ has a χ^2-distribution with $d = 2$ degrees of freedom, so if $F(x) := \mathbb{P}\left[\|Z_s\|^2 \leq x\right]$ is the distribution function of $\|Z_s\|^2$ we have

$$\begin{aligned}
&\mathbb{P}\left[\sup_{s\in[0,t]} \|X_s^n - X_s^{n+1}\| > \varepsilon\right] = \mathbb{P}\left[\max_{\{s\in D_{n+1}\setminus D_n : s<t\}} \|Z_s\| > 2^{n/2+1}\varepsilon\right] \\
&\leq \sum_{\substack{\{s\in D_{n+1}\setminus D_n, \\ s<t\}}} \mathbb{P}\left[\|Z_s\| > 2^{n/2+1}\varepsilon\right] = t2^n(1 - F(2^{n+2}\varepsilon^2)).
\end{aligned}$$

By changing into polar coordinates, it is easy to show that $F(x) = 1 - e^{-x/2}$ (this simple form is the reason we chose $d = 2$). Therefore,

$$\mathbb{P}\left[\sup_{s\in[0,t]} \|X_s^n - X_s^{n+1}\| > \varepsilon\right] \leq t2^n \exp(-2^{n+1}\varepsilon^2).$$

In particular,

$$\mathbb{P}\left[\sup_{s\in[0,t]}\|X_s^n - X_s^{n+1}\| > n^{-3}\right] \le t2^n \exp(-2^{n+1}n^{-6}).$$

Taking N large enough that $N\log(2) - 2^{N+1}N^{-6} < -N$, for all $n > N$ we have

$$\mathbb{P}\left[\sup_{s\in[0,t]}\|X_s^n - X_s^{n+1}\| > n^{-3}\right] \le te^{-n}.$$

By the Borel–Cantelli Lemma, as this sequence is summable we have

$$\mathbb{P}\left[\sup_{s\in[0,t]}\|X_s^n - X_s^{n+1}\| > n^{-3} \text{ for infinitely many } n\right] = 0.$$

In particular, with probability one, taking N sufficiently large, for all $n \ge N$,

$$\sup_{s\in[0,t]}\|X_s^n - X_s^{n+1}\| \le n^{-3}$$

and by the triangle inequality, recalling that $\sum_n n^{-2} = \pi^2/6$, for $N < n < m$,

$$\sup_{s\in[0,t]}\|X_s^n - X_s^m\| \le \sum_{j=n}^{m-1}\sup_{s\in[0,t]}\|X_s^j - X_s^{j+1}\| \le \frac{\pi^2}{6n}.$$

Therefore, with probability one, the processes X^n are converging uniformly on the interval $[0,t]$. By Lemma 7.15, X is a continuous process.

X is a Brownian Motion. We now need to show that X is a Brownian motion in its natural filtration, that is, that the increment $X_t - X_s$ is normally distributed and independent of $\mathscr{F}_s = \sigma(X_u, u \le s)$. First note that for s,t with $t \in D_n \setminus D_{n+1}$ and $\lceil s \rceil_n < t$, the random variable Z_t is not involved in the construction of X_s. Hence, as X generates the filtration and the $\{Z_u\}_{u\in\cup_n D_n}$ are independent, we see that Z_t is independent of $\mathscr{F}_s$.

It is clear that if s,t are integers with $s < t$, then

$$X_t - X_s = X_t^0 - X_s^0 = \sum_{\{k\in D_0: s<k<t\}} Z_k \sim N(0,(t-s)I_d).$$

Furthermore, in this case $X_t - X_s$ is independent of $\mathscr{F}_s$, as $Z_k = Z_{\lceil k \rceil_0}$ is independent of $\mathscr{F}_s$ for all $s < k$.

Now suppose that the result holds for $s,t \in D_n$.Then we see that for any $u \in D_{n+1} \setminus D_n$,

$$X_u - X_{\lfloor u \rfloor_n} = \frac{X_{\lceil u \rceil_n} - X_{\lfloor u \rfloor_n}}{2} + 2^{-(n/2+1)}Z_u \sim N(0, 2^{-(n+1)}I_d)$$

which is independent of $\mathscr{F}_{\lfloor u \rfloor_n}$. Similarly,

$$X_{\lceil u \rceil_n} - X_u = \frac{X_{\lceil u \rceil_n} - X_{\lfloor u \rfloor_n}}{2} - 2^{-(n/2+1)} Z_u \sim N(0, 2^{-(n+1)} I_d),$$

which is independent of $\mathscr{F}_{\lfloor u \rfloor_n}$. Therefore, for any $s, t \in D_{n+1}$,

$$X_t - X_s = (X_t - X_{\lfloor t \rfloor_n}) + (X_{\lfloor t \rfloor_n} - X_{\lceil s \rceil_n}) + (X_{\lceil s \rceil_n} - X_s),$$

which is the sum of three independent normal random variables, so

$$X_t - X_s \sim N(0, (t-s) I_d).$$

The first two terms are independent of $\mathscr{F}_{\lceil s \rceil_n} \supseteq \mathscr{F}_s$. We know the last term is independent of $\mathscr{F}_{\lfloor s \rfloor_n}$, and we can compute

$$\mathbb{E}\big[(X_{\lceil s \rceil_n} - X_s)(X_s - X_{\lfloor s \rfloor_n})^\top\big] = 0$$

so $(X_{\lceil s \rceil_n} - X_s)$ is independent of the increment $X_s - X_{\lfloor s \rfloor_n}$, as uncorrelated Gaussian random variables are independent. As we can write

$$\mathscr{F}_s = \mathscr{F}_{\lfloor s \rfloor_n} \vee \sigma(X_s - X_{\lfloor s \rfloor_n}) \vee \sigma\big(Z_u; u \in \,]\lfloor s \rfloor_n, s[\,\big),$$

we see that $X_{\lceil s \rceil_n} - X_s$ is independent of $\mathscr{F}_s$. Therefore $X_t - X_s$ is normally distributed and independent of $\mathscr{F}_s$, as desired.

Finally, for any $s < t$ we can find sequences $s_n \downarrow s$, $t_n \uparrow t$ with $s_n, t_n \in D_n$ and $s_0 \leq t_0$. Then $X_{t_n} - X_{s_n} \sim N(0, (t_n - s_n) I_d)$, and by continuity of X we see

$$X_t - X_s = X_{t_0} - X_{s_0} + \sum_{n=1}^{\infty} \big(X_{t_n} - X_{t_{n-1}} - X_{s_n} + X_{s_{n-1}}\big) \sim N(0, (t-s) I_d).$$

All the terms in this sum are independent of $\mathscr{F}_s$, as required. As $X_0 = 0$ by construction, we see that X is a Brownian motion starting at zero, in its natural filtration. □