Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Matemática
Primer Semestre de 2018

Stochastic Calculus Notes

Rodrigo Bazaes
rebazaes@uc.cl

June 17, 2018

Abstract

These notes in stochastic calculus are based on [1].

1 Stochastic Processes

1.1 Some Notation

We denote by:

1. $\mathbb{R}_+ := [0, \infty)$,
2. $\mathbb{Q}_+ := \mathbb{R}_+ \cap \mathbb{Q}$,
3. $\mathbb{Z}_+ := \{0, 1, 2, \cdots\}$,
4. $\mathbb{N} := \{1, 2, \cdots\}$

We will always be in some probability space $(\Omega, \mathcal{F}, P)$,and without loss of generality, we can assume that this measure space is complete.

Definition 1.1. *A filtration on a probability space $(\Omega, \mathcal{F}, P)$ is a collection of $\sigma-$algebras $\{\mathcal{F}_t\}_{t\in\mathbb{R}_+} = \{\mathcal{F}_t\}$ such that $\mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}$ if $0 \leq t \leq s < \infty$.*
We can define

$$\mathcal{F}_\infty := \sigma\left(\bigcup_{0\leq t<\infty} \mathcal{F}_t\right) \subset \mathcal{F} \tag{1.1}$$

WLOG,we can assume that each $\mathcal{F}_t$ contains all elements $A \in \mathcal{F}$ such that $P(A) = 0$, because we can replace $\mathcal{F}_t$ by

$$\overline{\mathcal{F}}_t := \{B \in \mathcal{F} : \text{there exists } A \in \mathcal{F}_t \text{ such that } P(A \triangle B) = 0\} \tag{1.2}$$

We call this filtration the *complete*, or *augmented* $\sigma-$algebra

Definition 1.2. *A stochastic process* $X = \{X_t : t \in \mathcal{I}\}$ *is a collection of random variables in the same probability space* $(\Omega, \mathcal{F}, P)$*,where* I *is some index set.Usually,the index set will be* $\mathbb{R}_+$*.Note that we can view* X *as a function on* $\mathbb{R}_+ \times \Omega$*,and we will write both* $X_t(\omega) = X(t, \omega)$*.If we don't mention the index set,it's assumed to be* $\mathbb{R}_+$*.*

If the random variables take their values in a space S, we say that the process is $S-$valued.Of course, we need a σ-algebra in this space. Usually, S will be a metric space, so the Borel $\sigma-$ algebra can be chosen.In the subsequent, we will only consider $S = \mathbb{R}^d$,unless otherwise specified.

Definition 1.3. *A process* $X = \{X_t : t \in R_+\}$ *is adapted to the filtration* $\mathcal{F}_t$ *if for all* $0 \leq t < \infty, X_t$ *is* $\mathbb{F}_t$ *measurable.Note that the smallest filtration such that* X *is adapted it is his natural filtration*

$$\mathcal{F}_t^X := \sigma\{X_s : 0 \leq s \leq t\}$$

Definition 1.4. *A process* X *is measurable if it is a measurable function*

$$X : (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+) \times (\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}_+))$$

Furthermore, we say that X *is progressively measurable if the restriction of the function* X *to* $[0, T] \times \Omega$ *is* $\mathcal{B}([0, T]) \times \mathcal{F}_T$ *measurable for each* T*.*

Remark 1.5. *If the process* X *is progressively measurable, then it is adapted, but the reverse does not hold.*

Definition 1.6. *If* X, Y *are stochastic processes in the same probability space, we say that are indistinguishable if there exist* $\Omega_0 \subset \Omega$ *with* $P(\Omega_0) = 1$ *and* $X_t(\omega) = Y_t(\omega)$ *for all* $t \in \mathbb{R}_+, \omega \in \Omega_0$*.*
We say that Y *is modification of* X *if for each* $t, P(X_t(\omega) = Y_t(\omega)) = 1$

Definition 1.7. *We say that* $X \stackrel{d}{=} Y$ *in distribution, if for all measurable set (that this has sense),* $P(X \in A) = P(Y \in A)$*.*
In fact, the following weaker statement is equivalent: for $\{t_1, t_2, \cdots, t_n\} \subset \mathbb{R}_+, P(X_{t_1} \in$

$A_1, \cdots X_{t_n} \in A_n) = P(Y_{t_1} \in A_1, \cdots Y_{t_n} \in A_n)$ *if* $B_1, \cdots B_n$ *are measurable sets where the last equality has sense.*

Remark 1.8. *If the space* $(\Omega, \mathcal{F}, P)$ *and the filtration* $\{\mathcal{F}_t\}$ *are complete,*X *is adapted to this filtration,and* Y *is a modification of* X*,then* Y *is also adapted.To see this,take* $B \in \mathcal{B}(\mathbb{R}^d)$*,and write* $A := \{Y_t \in B\}, C := \{X_t \in B\}, D_1 := (A \cap C^c), D_2 := (A^c \cap C)$*.Note that* $P(D_1) = P(D_2) = 0$*,so* $D_1, D_2 \in \mathcal{F}_t$*.Then,*

$$A = C \cup (A \cap C^c) \setminus (A^c \cap C) = C \cup D_1 \setminus D_2$$

Therefore, $A \in \mathcal{F}_t$

Definition 1.9. *A stopping time is a random variable* $\tau : \Omega \to [0, \infty]$ *such that the event* $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$ *for all* $t \in \mathbb{R}_+$*.*
If τ, σ *are stopping times, then* $\tau \wedge \sigma := \min\{\tau, \sigma\}$ *is also a stopping time.Also,* $\sigma, \vee \tau := \max\{\tau, \sigma\}$ *is a stopping time*

If τ is a stopping time, we define the stopping time σ-algebra $\mathcal{F}_\tau$ by

$$\mathcal{F}_\tau := \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \quad \forall t \in \mathbb{R}_+\}$$

We define the stopped process X_τ by $X_\tau(\omega) := X_{\tau(\omega)}(\omega)$ on the event $\{\tau < \infty\}$.

Remark 1.10. *When we deal with stopping times, sometimes we will be comparing both infinites, so we may assume that* $\infty \leq \infty, \infty = \infty$*, but it's false that* $\infty < \infty$

Lemma 1.11. *Let* σ, τ *stopping times, and* X *a stochastic process.*

i) For $A \in \mathcal{F}_\sigma$*,the events* $A \cap \{\sigma \leq \tau\}$ *and* $A \cap \{\sigma < \tau\}$ *lie in* $\mathcal{F}_\tau$*. In particular,* $\sigma \leq \tau$ *implies* $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$

ii) Both $\tau, \tau \wedge \sigma$ *are* $\mathcal{F}_\tau$ *measurable.The events* $\{\sigma \leq \tau, \}, \{\sigma < \tau\}, \{\sigma = \tau\}$ *lie in both* $\mathcal{F}_\sigma$ *and* $\mathcal{F}_\tau$

iii) If the process X *is progressively measurable,then* $X(\tau)$ *is* $\mathcal{F}_\tau$ *measurable on the event* $\{\tau < \infty\}$

Proof.

i) Let $A \in \mathcal{F}_\sigma$. We want to show that $(A \cap \sigma \leq \tau) \cap \{\tau \leq t\} \in \mathcal{F}_t$.Note that if $\sigma \leq \tau, \tau \leq t$,then $\sigma \leq t, \sigma \wedge t \leq \tau \wedge t, \tau \leq t$.Conversely, if $\sigma \leq t, \sigma \wedge t \leq \tau \wedge t, \tau \leq t$, then $\tau \wedge t = \tau, \sigma \wedge t = \sigma$,so $\sigma \leq \tau$.Hence,

$$(A \cap \{\sigma \leq \tau\}) \cap \{\tau \leq t\} = (A \cap \{\sigma \leq t\}) \cap \{\sigma \wedge t \leq \tau \wedge t\} \cap \{\tau \leq t\}$$

As $A \in \mathcal{F}_\sigma$,the first element in the right side of the equation is in $\mathcal{F}_t$.
For the second element,note that both $\tau \wedge t, \sigma \wedge t$ are $\mathcal{F}_t$ random variables.To see this, take $u \in \mathbb{R}$. If $u \geq t, \{\sigma \wedge t \leq u\} = \Omega$,and if $u < t$,then $\{\sigma \wedge t \leq u\} = \{\sigma \leq u\} \in \mathcal{F}_u \subset \mathcal{F}_t$.
The final term is clearly $\mathcal{F}_t$ measurable We conclude that $(A \cap \{\sigma \leq \tau\}) \in \mathcal{F}_\tau$. This proves that if $\sigma \leq \tau$,then $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$.
To show that $A \cap \{\sigma < \tau\} \in \mathcal{F}_\tau$,we write

$$A \cap \{\sigma < \tau\} = \bigcup_{n \geq 1} A \cap \{\sigma + \frac{1}{n} \leq \tau\}$$

Note that $\sigma \leq \sigma + \frac{1}{n}$,so by the last part, $A \in \mathcal{F}_{\sigma+1/n}$.Also, $\sigma + \frac{1}{n}$ is also a stopping time. Thus, using again the last part we conclude that $A \cap \{\sigma + \frac{1}{n} \leq \tau\} \in \mathcal{F}_\tau$

ii) To show that τ is $\mathcal{F}_\tau$ measurable, we need to prove that $\{\tau \leq s\} \in \mathcal{F}_\tau \leftrightarrow \{\tau \leq s\} \cap \{\tau \leq t\} \in \mathcal{F}_t$.But

$$\{\tau \leq s\} \cap \{\tau \leq t\} = \{\tau \leq t \wedge s\} \in \mathcal{F}_t$$

Using this,we conclude that $\sigma \wedge \tau$ is $\mathcal{F}_{\sigma \wedge \tau}$ measurable, so also $\mathcal{F}_\tau$ measurable.
Using $A = \Omega$ in part i), we have that $\{\sigma \leq \tau\}, \{\sigma < \tau\} \in \mathcal{F}_\tau$. Taking differences, $\{\sigma = \tau\} \in \mathcal{F}_\tau$. By symmetry, we can also prove these statements for $\mathcal{F}_\sigma$.

iii) We first claim that $\omega \to X(\tau(\omega) \wedge t, \omega)$ is $\mathcal{F}_t$ measurable. To see this, we write the map as a composition

$$\omega \to (\tau(\omega) \wedge t, \omega) \to X(\tau(\omega) \wedge t, \omega)$$

For the first composition, we note that $\tau(\omega) \wedge t$ is measurable from $(\Omega, \mathcal{F}_t)$ into $([0,t] \times \Omega, \mathcal{B}([0,t]) \times \mathcal{F}_t)$ by part i). Clearly the map $\omega \to \omega$ is measurable,so the whole map is measurable.
To see the second composition $(s, \omega) \to X(s, \omega)$,and this map is measurable from

$([0,t]\times\Omega, \mathcal{B}([0,t])\times\mathcal{F}_t)$ into $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ by the progressively measurable condition on X.Therefore, we have shown that $\{X_{\tau\wedge t}\in B\}\in\mathcal{F}_t$ for every $B\in\mathcal{B}(\mathbb{R}^d)$, and so

$$\{X_\tau \in B, \tau < \infty\} \cap \{\tau \le t\} = \{X_{\tau\wedge t}\in B\} \cap \{\tau \le t\} \in \mathcal{F}_t$$

This concludes the proof.

□

1.2 Regularity properties of paths

The following are properties of the function $X_t(\omega)$ when $\omega\in\Omega$ is fixed.

Definition 1.12.

i) A stochastic process X is continuous, if for each $\omega\in\Omega$, the function $X_t(\omega):\mathbb{R}_+\to\mathbb{R}^d$ is continuous.Analogous result is for left-continuous and right-continuous properties.

ii) An $\mathbb{R}^d-$valued process X is right continuous with left limits (abbreviated cadlag from now) if the following holds for every $\omega\in\Omega$:

$$X_t(\omega) = \lim_{s\searrow t} X_s(\omega) \text{ for every } t\in\mathbb{R}_+$$

$$X_{t-}(\omega) := \lim_{s\nearrow t} X_s(\omega) \text{ exists in } \mathbb{R}^d \text{ for all } t>0$$

Analogously, we define a left continuous with right limits (caglad).

iii) X is a finite variation process (FV process) if for each $\omega\in\Omega$, the path $t\to X_t(\omega)$ has bounded variation on each compact interval $[0,T]$.

Note that we can assume that these properties hold for a.e $\omega\in\Omega$, because we can construct a process $\tilde{X}$ with the required property and $X,\tilde{X}$ are indistinguishable But, this cannot be applied for a modification of X.

Lemma 1.13. *Let X be adapted to the filtration $\{\mathcal{F}_t\}$, and suppose that X is either left or right continuous. Then X is progressively measurable.*

Proof. We only prove the right continuous case. Fix $T < \infty$. Define on $[0,T] \times \Omega$ the function

$$X_n(t,\omega) := X(0,\omega) 1_{\{}0\}(t) + \sum_{k=0}^{2^n-1} X\left(\frac{(k+1)T}{2^n}, \omega\right) 1_{(kT2^{-n},(k+1)T2^{-n}}(t)$$

is a sum of products of $\mathcal{B}([\prime, \mathcal{T}]) \otimes \mathcal{F}_t$ measurable functions, so also X_n. As X is right continuous, then $X_n(t,\omega) \to X(t,\omega)$ when $n \to \infty$.Thus, X is also $\mathcal{B}([0,T]) \otimes \mathcal{F}_t$ measurable when restricted to $[0,T] \times \Omega$. □

Lemma 1.14. *Suppose X, Y are right-continuous processes defined on the same probability space.Suppose $P(X_t = Y_t) = 1$ for all t in some dense countable subset $S \subset \mathbb{R}_+$. Then X, Y are indistinguishable. The same conclusion holds under the assumption of left-continuity if $0 \in S$*

Proof. Let $\Omega_0 := \bigcap_{s \in S} \{\omega : X_s(\omega) = Y_s(\omega)\}$. Note that Ω_0 is countable intersection of measure 1 sets,so $P(\Omega_0) = 1$.Now if $t \in \mathbb{R}_+$,we approximate by elements in S, and using the right continuity, we conclude that $X_t(\omega) = Y_t(\omega)$ if $\omega \in \Omega_0$, for all $t \in \mathbb{R}_+$.Hence, X, Y are indistinguishable. In the left-continuous case, we cannot approximate 0 by the left, so we assume $0 \in S$. □

We introduce the notion of limit in the filtrations. If $\{\mathcal{F}_t\}$ is a filtration, we define

$$\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$$

$\{\mathcal{F}_{t+}\}$ is a new filtration , and $\mathcal{F}_t \subset \mathcal{F}_{t+}$.If $\mathcal{F}_t = \mathcal{F}_{t+}$,we say that $\{\mathcal{F}_t\}$ is right-continuous.Note that by definition, $\{\mathcal{F}_{t+}\}$ is right-continuous.

Definition 1.15. *We say that the filtration $\{\mathcal{F}_t\}$ satisfies the usual conditions if it is right continuous and complete.*

Note that because $\{\mathcal{F}_{t+}\}$ is bigger or equal than $\{\mathcal{F}_t\}$, the first contains more stopping times. In fact,we have the following:

Lemma 1.16. *If τ is a $[0,\infty]$ random variable ,then it's a stopping time with respect to $\{\mathcal{F}_{t+}\}$ if and only if $\{\tau < t\} \in \mathcal{F}_t$ for all $t \in \mathbb{R}_+$*

Proof. Suppose that τ is an $\{\mathcal{F}_{t+}\}$ stopping time.Then, for $n \in \mathbb{N}$,

$$\{\tau \leq t - n^{-1}\} \in \mathcal{F}_{(t-n^{-1})+} \subset \mathcal{F}_t$$

So, $\{\tau < t\} = \bigcup_{n\in\mathbb{N}}\{\tau \leq t - n^{-1}\} \in \mathcal{F}_t$.
Conversely, if $\{\tau < t\} \in \mathcal{F}_t$, then for ,then for $m \in \mathbb{N}$,

$$\{\tau \leq t\} \bigcap_{n\geq m} \{\tau < t + \frac{1}{m}\} \in \mathcal{F}_{t+1/m}$$

Thus, $\{\tau \leq t\} \in \bigcap_{m\in\mathbb{N}} \mathcal{F}_{t+1/m} = \mathcal{F}_{t+}$ $\square$

Now we introduce one important random time.

Definition 1.17. *Given a set H,the hitting time of the set H is*

$$\tau_H(\omega) := \inf\{t \geq 0 : X_t(\omega) \in H\}$$

Also, the first entry time is the same definition as above, if we put $t > 0$ in the infimum.

We prove that the hitting time is in fact a stopping time.

Lemma 1.18. *Let X be a process adapted to the filtration $\{\mathcal{F}_t\}$ and assume X is left or right- continuous If G is an open set, then τ_G is a stopping time with respect to $\mathcal{F}_{t+}$*

Proof. WLOG, X is right continuous.Note that $\tau_G(\omega) < t$ iff $X_s(\omega) \in G$ for some $s \in [0, t)$ iff $X_q(\omega) \in G$ for some rational $q \in [0, t)$ (because X is right continuous).Thus we have

$$\{\tau_G < t\} = \bigcup_{q\in Q_+\cap[0,t)} \{X_q \in G\} \in \sigma\{X_s : 0 \leq s < t\} \subset \mathcal{F}_t$$

And the proof is complete by Lemma 1.16. $\square$

Note that in the last lemma, we need the right/left continuity because we have to write the event $\{\tau_G < t\}$ as a countable union of sets, and this is guaranteed by the assumptions.

Example 1.19. *Assuming that X is continuous doest not permit us to conclude that $\{\tau_G \leq t\} \in \mathcal{F}_t$.Take $G = (b, \infty)$ for some $b > 0$,and consider two paths*

$$X_s(\omega_0) = X_s(\omega_1) = bs \quad \text{for } 0 \leq s \leq 1$$

while

$$\text{for } s \geq 1 \begin{cases} X_s(\omega_0) = bs \\ X_s(\omega_1) = b(2-s) \end{cases}$$

We have $\tau_G(\omega_0) = 1, \tau_G(\omega_1) = \infty$*.As* ω_0, ω_1 *agree on* $X_s, 0 \leq s \leq 1$*,then they are together or outside any event in* $\mathcal{F}_1^X$*. But* $\omega_0 \in \{\tau_G \leq 1\}, \omega_1 \notin \{\tau_G \leq 1\}$ *,thus* $\{\tau_G \leq 1\} \notin \mathcal{F}_t^X$

There is an alternative way to register arrival into a set. For a process X, let $X[s,t] = \{X_u : s \leq u \leq t\}$, and $\overline{X[s,t]}$ its closure. For a set H,define

$$\sigma_H := \inf\{t \geq 0 : \overline{X[s,t]} \cap H \neq \emptyset\}$$

Note that for a cadlag path,

$$\overline{X[0,t]} = \{X(u) : 0 \leq u \leq t\} \cup \{X(u-) : 0 < u \leq t\}$$

Lemma 1.20. *Suppose* X *is a cadlag process adapted to* $\{\mathcal{F}_t\}$ *and* H *is a closed set.Then* σ_H *is a stopping time.*

Proof. Fix $t \geq 0$.First we claim that

$$\{\sigma \leq t\} = \{X(0) \in H\} \cup \{X(s) \in H \text{ or } X(s-) \in H \text{ for some } s \in (0,t]\}$$

the contention $\supset$ is clear. To prove the other contention, suppose that $\sigma_H \leq t$.If the event on the right does not happen,by the definition of infimum, for each $k \in \mathbb{N}$ there exists $t < u_k \leq t + 1/k$ such that either $X(u_k) \in H$ or $X(u_k-) \in H$.As $u_k \to t$, the by the right continuity of X, both $X(u_k), X(u_k-)$ converge to $X(t) \in H$ because H is closed. This is a contradiction. That completes the equality.
Now let

$$H_n := \{y : \text{ there exists } x \in H \text{ such that } |x-y| < n^{-1}\}$$

Let $U := ([0,t] \cap \mathbb{Q}) \cup \{t\}$. Now we claim

$$\{X(0) \in H\} \cup \{X(s) \in H \text{ or } X(s-) \in H \text{ for some } s \in (0,t]\} = \bigcap_{n=1}^{\infty} \bigcup_{q \in U} \{X(q) \in H_n\} \in \mathcal{F}_t$$

To show this, note that if $X(s) = y \in H$ for some $s \in [0,t]$ or $X(s-) = y \in H$ for some $s \in (0,t]$,we can find a sequence $q_j \in U$ such that $X(q_j) \to y$. and then $X(q_j) \in H_n$ for all large enough j. Conversely, suppose that $q_n \in U$ such that $X(q_n) \in H_n$ for all n.Extract a convergent subsequence $q_n \to s$. By the cadlag property, a further subsequence of $X(q_n)$ converges to either $X(s)$ or $X(s-)$, and in any case, belongs to H because it's a closed set. □

Observe that if X has continuous paths, then the definitions of τ, σ coincide for a closed set H.

Corollary 1.21. *Assume X is continuous and H is closed. Then τ_H is a stopping time.*

1.3 Quadratic Variation

Let Y be a stochastic process. Given a partition $\pi = \{0 = t_0 < t_1 < \cdots < t_{m(\pi)} = t\}$ of $[0,t]$, we consider the sum

$$\sum_{i=0}^{m(\pi)-1} (Y_{t_{i+1}} - Y_{t_i})^2$$

We say that these sums converge to the random variable $[Y]_t$ in probability as $mesh(\pi) := \max(t_{i+1} - t_i) \to 0$ if for each $\epsilon > 0$ there exist $\delta > 0$ such that

$$P\left\{ \left| \sum_{i=0}^{m(\pi)-1} (Y_{t_{i+1}} - Y_{t_i})^2 - [Y]_t \right| \geq \epsilon \right\} \leq \epsilon \tag{1.3}$$

for all partitions π with $mesh(\pi) \leq \delta$.We express the limit as

$$\lim_{mesh(\pi) \to 0} \sum_i (Y_{t_{i+1}} - Y_{t_i})^2 = [Y]_t \text{ in probability} \tag{1.4}$$

Definition 1.22. *The quadratic variation process $[Y] = \{[Y]_t : t \in \mathbb{R}_+\}$ of a stochastic process Y is a process such that $[Y]_0 = 0$, the paths $t \to [Y]_t(\omega)$ are nondecreasing for all ω and the limit (1.4) holds for all $t \geq 0$.*

Definition 1.23. *Let X, Y two processes in the same space. The (quadratic) covariance process $[X,Y] := \{[X,Y]_t : t \in \mathbb{R}_+\}$ is defined by*

$$[X,Y] := [\frac{1}{2}(X+Y)] - [\frac{1}{2}(X-Y)] \tag{1.5}$$

provided that both quadratic variations on the right exist in the sense of equation (1.4).

Using the identify $ab = \frac{1}{4}(a+b)^2 - \frac{1}{4}(a-b)^2$ applied to $a = X_{t_{i+1}} - X_{t_i}, b = Y_{t_{i+1}} - Y_{t_i}$ we conclude that for each $t \in \mathbb{R}_+$,

$$\lim_{mesh(\pi)\to 0} \sum_i (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i}) = [X,Y]_t \tag{1.6}$$

in probability.Now,using the identities

$$ab = \frac{1}{2}((a+b)^2 - a^2 - b^2) = \frac{1}{2}(a^2 + b^2 - (a-b)^2)$$

we obtain the almost surely identities

$$[X,Y]_t = \frac{1}{2}\left([X+Y]_t - [X]_t - [Y]_t\right) \tag{1.7}$$

$$[X,Y]_t = \frac{1}{2}\left([X]_t + [Y]_t - [X-Y]_t\right) \tag{1.8}$$

Note that using (1.5) it's clear that $[Y,Y] = [Y]$

Proposition 1.24. *Suppose X and Y are cadlag processes,and $[X,Y]$ exists as in (1.5). Then there exists a cadlag modification of $[X,Y]$.Fo any $t, \triangle[X,Y]_t = (\triangle X_t)(\triangle Y_t)$ almost surely, where*

$$\triangle Z(t) := Z(t) - Z(t-)$$

is the jump of the process Z at t.

Proof. It's enough consider the case $X = Y$ by the definition of $[X,Y]$.Note that if $t < u$, and $[X,Y]_u, [X,Y]_t$ satisfy (1.5),then

$$[X,Y]_u - [X.Y]_t = \lim_{mesh(\pi)\to 0} \sum_i (X_{s_{i+1}} - X_{s_i})(Y_{s_{i+1}} - Y_{s_i}) \tag{1.9}$$

where the limit is in probability and taken over partitions $\{s_i\}$ of $[t,u]$ as the mesh tend to 0.Using this, given $\epsilon, \delta > 0$.Fix $t < u$,and pick $\eta > 0$ such that

$$P\left\{\left|[X]_u - [X]_t - \sum_{i=0}^{m(\pi)-1} (X_{t_{i+1}} - X_{t_i})^2 < \epsilon\right|\right\} > 1 - \delta \tag{1.10}$$

whenever $\pi = \{t = t_0 < t_1 < \cdots < t_{m(\pi)} = u\}$ is a partition of $[t,u]$ with mesh$(\pi) <$ η.Keeping t_1 fixed, refine π in $[t_1, u]$ so that

$$P\left\{\left|[X]_u - [X]_{t_1} - \sum_{i=0}^{m(\pi)-1} (X_{t_{i+1}} - X_{t_i})^2 < \epsilon\right|\right\} > 1 - \delta \tag{1.11}$$

So, if both equations are satisfied, with probability at least $1-2\delta$,

$$
\begin{aligned}
[X]_u - [X]_t &\le \sum_{i=0}^{m(\pi)-1} (X_{t_{i+1}} - X_{t_i})^2 + \epsilon \\
&= (X_{t_1} - X_t)^2 + \sum_{i=1}^{m(\pi)-1} (X_{t_{i+1}} - X_{t_i})^2 + \epsilon \\
&\le (X_{t_1} - X_t)^2 + [X]_u - [X]_{t_1} + 2\epsilon
\end{aligned}
$$

This is equivalent to say that

$$[X]_{t_1} \le [X]_t + (X_{t_1} - X_t)^2 + 2\epsilon$$

Note that this argument works for any $t_1 \in (t, t+\eta)$. By monotonicity, $[X]_{t+} \le [X]_{t_1}$,so for all these t_1,

$$P\{[X]_{t+} \le [X]_t + (X_{t_1} - X_t)^2 + 2\epsilon\} > 1 - 2\delta$$

We can pick $\eta > 0$ such that for $t_1 \in (t, t+\eta)$ by right continuity,

$$P\{(X_{t_1} - X_t)^2 < \epsilon\} > 1 - \delta$$

We obtain the estimate

$$P\{[X]_t \le [X]_{t+} \le [X]_t + 3\epsilon\} > 1 - 3\delta$$

Since ϵ, δ are arbitrary, it follows that $[X]_{t+} = [X]_t$ almost surely. It can be proved (PROVE) that the process $[X]_{t+}$ has cadlag paths,and therefore we can choose a version of $[X]$ with cadlag paths.

To prove the second claim, we consider the partition π from (1.10).Let $s = t_{m(\pi)-1}$.Keeping s fixed, refine π in $[t, s]$ in such a way with probability at least $1-2\delta$,

$$
\begin{aligned}
[X]_u - [X]_t &\le \sum_{i=0}^{m(\pi)-1} (X_{t_{i+1}} - X_{t_i})^2 + \epsilon \\
&= (X_u - X_s)^2 + \sum_{i=0}^{m(\pi)-2} (X_{t_{i+1}} - X_{t_i})^2 + \epsilon \\
&\le (X_u - X_2)^2 + [X]_s - [X]_t + 2\epsilon
\end{aligned}
$$

That implies that $[X]_u - [X]_s \leq (X_u - X_2)^2 + 2\epsilon$.Note that by monotonicity, $\triangle[X]_u \leq [X]_u - [X]_s$, and thus

$$P\{\triangle[X]_u \leq (X_s - X_u)^2 + 2\epsilon\} > 1 - 2\delta$$

As $s \in (u - \eta, u)$ is arbitrary, we can pick η small enough so that for all such s and with probability at least $1 - \delta$,

$$|(X_u - X_2)^2 - (\triangle X_u)^2| < \epsilon$$

And , as δ, ϵ are arbitrary, conclude that almost surely,

$$\triangle[X]_u \leq (\triangle X_u)^2$$

Similar calculations again with $s = t_{m(\pi)-1}$ give us with probability at least $1 - 2\delta$

$$[X]_u - [X]_s \geq (X_u - X_s) - \epsilon \geq (\triangle X_u)^2 - 2\epsilon$$

taking s close enough to u such that $P\{\triangle[X]_u \geq [X]_u - [X]_s - \epsilon\} > 1 - \delta$, we deduce that with probability at least $1 - 3\delta$ we have

$$\triangle[X]_u \geq (\triangle X_u)^2 - 3\epsilon$$

Taking $\epsilon, \delta \to 0$,concludes the equality. □

We conclude that if Y is a cadlag, then $[Y]$is an increasing process:

Definition 1.25. *An increasing process $A = \{A_t : 0 \leq t < \infty\}$ is an adapted process such that, for almost every $\omega, A_0(\omega) = 0$ and $s \to A_s(\omega)$ is non-decreasing and right-continuous.Monotonicity implies the existence of the lefts limits, so in particular, an increasing process is cadlag.*

We continue with two inequalities:

Lemma 1.26. *Suppose the processes below exist. Then, for fixed t*

$$|[X, Y]_t| \leq [X]_t^{1/2}[Y]_t^{1/2} \quad a.s. \tag{1.12}$$

and more generally for $0 \leq s < t$

$$|[X, Y]_t - [X, Y]_s| \leq ([X]_t - [X]_s)^{1/2}([Y]_t - [Y]_s)^{1/2} \quad a.s \tag{1.13}$$

Furthermore,

$$|[X]_t - [Y]_t| \leq [X-Y]_t + 2[X-Y]_t^{1/2}[Y]_t^{1/2} \quad a.s \tag{1.14}$$

In the cadlag case the inequalities are valid simultaneously at all $s < t \in \mathbb{R}_+$

Proof. The last affirmation follows from the fact that the inequalities are valid simultaneously for all rational times almost surely, and applying limits we deduce the result.
The first inequality follows from Cauchy-Schwartz
For the second one, we recall (1.9).We have that

$$[X,Y]_t - [X,Y]_s = \lim_{mesh(\pi)\to 0} \sum_i (X_{t_{i+1}} - X_{t_i})(X_{t_{i+1}} - X_{t_i})$$

with π partition of $[s,t]$.Now again apply Cauchy-Schwartz
For the last inequality, we use the identity $a^2 - b^2 = (a-b)^2 + 2(a-b)b$ applied to the increments of X and Y,obtaining

$$[X] - [Y] = [X-Y] + 2[X-Y,Y]$$

Now we apply (1.12) to $[X-Y,Y]$ to conclude. □

In our applications, $[X,Y]$ will be a cadlag process. Because this process is the difference of two increasing processes, $[X,Y]_t$ is a bounded variations (BV) function on any compact time interval. We can consider the Lebesgue -Stieltjes measure $\Lambda_{[X,Y]}$ defined by

$$\Lambda_{[X,Y]}(a,b] := [X,Y]_b - [X,Y]_a, \quad 0 \leq a < b < \infty$$

When the origin is included in the time interval, we assume $[X,Y]_{0-} = 0$. We derive a Cauchy-Schwartz" type of inequality:

Proposition 1.27 (Kunita-Watanabe inequality)**.** *Fix ω such that $[X],[Y]$ and [X,Y] exist and are right-continuous on the interval $[0,T]$.Then for any $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable bounded function G and H on $[0,T]\times\Omega$,*

$$\left|\int_{[0,T]} G(t,\omega)H(t,\omega)d[X,Y]_t(\omega)\right| \leq \left\{\int_{0,T} G(t,\omega)^2 d[X]_t(\omega)\right\}^{1/2} \left\{\int_{0,T} H(t,\omega)^2 d[Y]_t(\omega)\right\}^{1/2} \tag{1.15}$$

The integrals are of Lebesgue-Stieltjes type with respect to t, and ω is fixed.

Proof. Note that once ω is fixed,we can consider functions with parameter t. If g,h are step functions, it's easy to deduce (1.15) using (1.12) and the classical Cauchy-Schwartz inequality.

If g,h are arbitrary bounded Borel functions on $[0, T]$, and pick $0 < C < \infty$ such that $|g| \leq C, |h| \leq C$.Let $\epsilon > 0$,and define the Borel measure

$$\mu := \Lambda_{[X]} + \Lambda_{[Y]} + |\Lambda_{[X,Y]}|$$

on $[0, T]$.Here, $\Lambda_{[X]}$ is the positive Lebesgue-Stieltjes measure of the function $t \to [X]_t$,same to $\Lambda_{[Y]}$, and $|\Lambda_{[X,Y]}|$ is the positive total variation measure of the signed Lebesgue-Stieltjes measure $\Lambda_{[X,Y]}$.Now, take $\tilde{g}, \tilde{h}$ steps functions with $|\tilde{g}| \leq C, |\tilde{h}| \leq C$ and

$$\int (|g - \tilde{g}| + |h - \tilde{h}|) d\mu < \frac{\epsilon}{2C}$$

On the one hand we have

$$\begin{aligned}\left|\int_{[0,T]} ghd[X,Y]_t - \int_{[0,T]} \tilde{g}\tilde{h}d[X,Y]_t\right| &\leq \int_{[0,T]} |gh - \tilde{g}\tilde{h}|d|\Lambda_{[X,Y]}| = \int_{[0,T]} |g(h - \tilde{h}) + \tilde{h}(\tilde{g} - g)d|\Lambda_{[X,Y]}| \\ &\leq C \int_{[0,T]} |(\tilde{g} - g)|d|\Lambda_{[X,Y]}| + C \int_{[0,T]} |(\tilde{h} - h)|d|\Lambda_{[X,Y]}| \leq \epsilon\end{aligned}$$

On the other hand,

$$\begin{aligned}\left|\int_{[0,T]} g^2 d[X]_t - \int_{[0,T]} \tilde{g}^2 d[X]_t\right| &\leq \int_{[0,T} |g^2 - \tilde{g}^2| d[X]_t \\ &\leq 2c \int_{[0,T]} |g - \tilde{g}| d[X]_t \leq \epsilon\end{aligned}$$

And the same applies to h. Finally, putting all together and using the result for steps functions, we conclude that

$$\left|\int_{[0,T]} ghd[X,Y]_t|\right| \leq \epsilon + \left(\epsilon + \int_{[0,T]} g^2 d[X]_t\right)^{1/2} \left(\epsilon + \int_{[0,T]} g^2 d[Y]_t\right)^{1/2}$$

We take $\epsilon \to 0$,and finally we apply he result to $g(t))G(t, \omega), h(t) = H(t, \omega)$. □

1.4 Path spaces and Markov processes

Consider a $\mathbb{R}^d$-valued stochastic process, and we want an space of measure U such that the map $X : \Omega \to U$ be measurable.Note that for each $\omega \in \Omega, X(\omega)$ is a function from $\mathbb{R}_+ \to$

$\mathbb{R}^d$,therefore U is a space of functions.The path regularity determines the path space.Some examples are

i) Without further assumptions, U will be the space $(\mathbb{R}^d)^{[0,\infty)}$,the space of functions from $\mathbb{R}_+ \to \mathbb{R}^d$ with the product $\sigma-$algebra $\mathcal{B}(\mathbb{R}^d)^{\otimes[0,\infty)}$

ii) If X is an $\mathbb{R}^d-$valued cadlag process, then we can consider the path space $D = D_{\mathbb{R}^d}([0,\infty)$,the space of $\mathbb{R}^d-$valued cadlag functions on $[0,\infty)$, with $\sigma-$algebra generated by the coordinate projections $\xi \to \xi(t)$ from D to $\mathbb{R}^d$.The Skorohod metric makes D a complete,separable metric space, and the Borel $\sigma-$algebra is the generated by the coordinate projections.We denote this $\sigma-$algebra by $\mathcal{B}_D$

iii) If X is an $\mathbb{R}^d-$continuous valued process, then X maps into $C = C_{\mathbb{R}^d}([0,\infty))$.This space is metrized by

$$r(\eta,\zeta) := \sum_{k=1}^{\infty} 2^{-k}(1 \wedge \sup_{0\le t\le k} |\eta(t) - \zeta(t|), \quad \eta,\zeta \in C \tag{1.16}$$

This is the metric of uniform convergence on compact sets . With is metric, C is a complete and separable space, and its Borel $\sigma-$algebra is generated by the coordinate projections.Note that C is a subspace of D, and in fact the notions of convergence and measurability in C coincide with such notions in D.

Generating the $\sigma-$algebra using coordinate projections guarantees that X is a measurable mapping .We can define then the distribution of the process in the path space. For example, if X is a cadlag, define $\mu(B) = P(X \in B)$ for $B \in \mathcal{B}_D$.If we consider the probability space $(D, \mathcal{B}_D, \mu)$, and define the process $\{Y_t\}$ in this space via the mapping $Y_t(\omega) = \omega(t), \omega \in D$. Then the precesses X and Y have the same distribution. In fact

$$P(X \in B) = \mu(B) = \mu(\omega \in D : \omega \in B) = \mu(\omega \in D : Y(\omega) \in B) = \mu(Y \in B)$$

The two most important classes of stochastic processes are martingales and Markov processes.We assume that the process $X = \{X_t\}$ is adapted to the filtration $\{\mathcal{F}_t\}$.

i) Let X be a real-valued process. Then X is a *martingale* with respect to $\{\mathcal{F}_t\}$ if for each t, X_t is integrable and for all $s < t$,

$$E[X_t|\mathcal{F}_s] = X_s \tag{1.17}$$

ii) An $\mathbb{R}^d$-valued process X satisfies the *Markov property* with respect to $\{\mathcal{F}_t\}$ if

$$P[X_t \in B|\mathcal{F}_s] = P[X_t \in B|X_s] \text{ for all } s < t \text{ and } B \in \mathcal{B}_{\mathbb{R}^d} \tag{1.18}$$

We discuss the most basic example of discrete-time processes.

Example 1.28. *(Random Walk) Let $X_0, X_1, \cdots$ be an i.i.d sequence of random variables. Define the partial sums by $S_0 = 0$ and for $n \geq 1, S_n = X_1 + \cdots X_n$. Then S_n is a Markov chain,i.e., a discrete-time Markov process. If $E(X_1) = 0$,then $\{S_n\}$ is a martingale. The filtration used here is the natural $\sigma-$algebra $\mathcal{F}_n = \sigma\{X_1, \cdots, X_n\}$*

In the next section we will discuss martingales in more detail. Now, we focus on Markov processes. In particular, the Markov property and the strong Markov property.
In the last example, we can consider the family of processes changing the initial point , $S_0 = x, S_n = x + X_1 + \cdots + X_n$. We will consider probability distributions on a path space to define such family.
On the path space D, the coordinate variables are defined by $X_t(\omega) := \omega(t), \omega \in D$, and the natural filtration $\mathcal{F}_t = \sigma(X_s : s \leq t)$. We also write $X(t)$ when the subscript are not convenient.The shift maps $\theta_s : D \to D$ are defined by $\theta_s(\omega)(t) := \omega(t + s)$. That is, the path $\theta_s\omega$ has its time origin translated to s and the path before s is deleted.For an event $A \in \mathcal{B}_D$,the inverse image

$$\theta_s^{-1}A = \{\omega \in D : \theta_s\omega \in A\}$$

represents the event "A happens starting at time s"

Definition 1.29. *An $\mathbb{R}^d-$valued Markov process is a collection $\{P^x : x \in \mathbb{R}^d\}$ of probability measures on D with the properties:*

a) $P^x(\omega \in D : \omega(0) = x) = 1$

b) For each $A \in \mathcal{B}_D$, the function $x \to P^x(A)$ is measurable on $\mathbb{R}^d$

c) $P^x[\theta_t^{-1}A|\mathcal{F}_s](\omega) = P^{X_s(\omega)}(A)$ for P^x-almost every $\omega \in D$,for every $x \in \mathbb{R}^d$ and $A \in \mathcal{B}_D$

The first requirement says that P^x is the distribution with initial state x.The second requirement is because if you want that your initial state X_0 is distributed by a measure

μ, then $P^\mu(A) = \int P^x \mu(dx)$. The third requirement is the Markov property.Note that the event may depend on the entire process. But we could have defined c) first with event of the type $A = \{X_s \in B\}$, then to finite-dimensional events, and finally to $\mathcal{B}_D$.We write E^x for the expectation with respect to P^x. To prove (1.18), we use the fact that if $\mathcal{A}, \mathcal{B}$ are two $\sigma-$algebras such that $\mathcal{A} \subset \mathcal{B}$ and $E(X|\mathcal{B})$ is $\mathcal{A}$-measurable,

$$E(X|\mathcal{A}) = E(X|\mathcal{B}) \tag{1.19}$$

.In our case, we write

$$P^x(X_t \in B|\mathcal{F}_s)(\omega) = P^x(\theta_s^{-1}\{X_{t-s} \in B\}|\mathcal{F}_s) = P^{X_s(\omega)}(X_{t-s} \in B) = f(X_s(\omega))$$

Using b),we know that $f(x)$ is measurable, so $P^x(X_t \in B|\mathcal{F}_s)(\omega)$ is X_s measurable. As $\sigma(X_s) \subset \mathcal{F}_s$, we can apply the fact mentioned above to conclude (1.18).
The item c) in the definition says that if we know that the process is in state y at time t, then regardless of the past, the future of the process behaves exactly as a new process started from y.

Definition 1.30. *The transition probability of the Markov process is defined by*
$p(t, x, B) := P^x(X_t \in B)$.

For each $t \in \mathbb{R}_+$, p(t,x,B) is a measurable function of $x \in \mathbb{R}^d$ and a Borel probability measure in the set argument $B \in \mathcal{B}_{\mathbb{R}^d}$. It gives the conditional probability in (1.18), regardless of the initial distribution:

$$P^\mu(X_{s+t} \in B|\mathcal{F}_s)(\omega) = p(t, X_s(\omega), B). \tag{1.20}$$

To check this, take $A \in \mathcal{F}_s, B \in \mathcal{B}_{\mathbb{R}^d}$.

$$\begin{aligned}
E^\mu[1_A A\{X_{s+t} \in B\}] &= \int E^x[1_A 1\{X_{s+t} \in B\}]\mu(dx) \\
&= \int E^x[1_A(\omega) P^x(X_{s+t} \in B|\mathcal{F}_s)(\omega)]\mu(dx) \\
&= \int E^x[1_A(\omega) P^x(\theta_s^{-1}\{X_t \in B\}|\mathcal{F}_s)(\omega)]\mu(dx) \\
&\overset{c)}{=} \int E^x[1_A(\omega) P^{X_s(\omega)}(X_t \in B)]\mu(dx) \\
&= \int E^x[1_A(\omega) p(t, X_s(\omega), B)]\mu(dx) \\
&= E^\mu[1_A p(t, X_s, B)]
\end{aligned}$$

When we used the property c) in Definition 1.29.That proves the claim.

Proposition 1.31. *For* $0 = s_0 < s_1 < \cdots < s_n$ *and a bounded Borel function* f *on* $\mathbb{R}^{d(n+1)}$,

$$\begin{aligned}
E^\mu[f(X_{s_0}, X_{s_1}, \cdots X_{s_n})] = \int \cdots \int \mu(dx_0) p(s_1, x_0, dx_1) p(s_2 - s_1, x_1, dx_2) \\
\cdots p(s_n - s_{n-1}, x_{n-1}, dx_n) f(x_0, x_1, \cdots, x_n)
\end{aligned} \tag{1.21}$$

Proof. Note that if bounded and Borel on $\mathbb{R}^d$, we have

$$E^\mu(f(X_{s+t})|\mathcal{F}_s)(\omega) = \int f(y) p(t, X_s(\omega), dy) \tag{1.22}$$

because if $f = 1_B$, this is simply (1.20) and the rest is standard. To check the proposition,we proceed by induction in n. It's enough to consider functions of the form $f(x_0, x_1, \cdots, x_n) = f_0(x_0) f_1(x_1) \cdots f_n(x_n)$, because they are dense in the bounded Borel functions on $\mathbb{R}^{d(n+1)}$. If $n = 1$, we write

$$\begin{aligned}
E^\mu[f(X_0, X_{s_1})] = E^\mu[f_0(X_0) f_1(X_{s_1})] &= E^\mu[f_0(X_0) E^\mu\{f_1(X_{s_1}|\mathcal{F}_0)\}] \\
&= E^\mu[f_0(X_0) E^\mu\{f_1(X_{s_0+s_1-s_0}|\mathcal{F}_0)\}] \\
&\overset{(1.22)}{=} E^\mu[f_0(X_0) \int f(x_1) p(s_1, X_0(\omega), dx_1)] \\
&= \int \int f_0(x_0) f(x_1) p(s_1, X_0(\omega), dx_1) \mu(dx_0)
\end{aligned}$$

That proves the case $n = 1$.Now we suppose that the claim is true for $n = k \geq 1$.To prove the case $n+1$, we proceed similarly,

$$E^\mu[f_0(X_0)f_1(X_{s_1})\cdots f_n(X_{s_n})f_{n+1}(X_{s_{n+1}})] = E^\mu[f_0(X_0)f_1(X_{s_1})\cdots f_n(X_{s_n})E^\mu[f_{n+1}(X_{s_{n+1}})]|\mathcal{F}_{s_n}]$$
$$= E^\mu[f_0(X_0)f_1(X_{s_1})\cdots f_n(X_{s_n})\int f(x_{n+1})p(s_{n+1}-s_n, X_{s_n}, dx_{n+1})]$$
$$= \int\cdots\int \mu(dx_0)p(s_1,x_0,dx_1)p(s_2-s_1,x_1,dx_2)\cdots p(s_{n+1}-s_n,x_n,dx_{n+1})f(x_0,x_1,\cdots,x_n,x_{n+q})$$

in the last step we applied the induction hypothesis to the function

$$f(x_0,x_1,\cdots x_n) = f_0(x_0)\cdots f_n(x_n)\int f(y)p(s_{n+1}-s_n,x_n,dy)$$

$\square$

We can define a family of operators $S(t)$ that acts in the space of bounded and measurable functions g on the state space by the relation

$$S(t)g(x) := E^x[g(X_t)] \tag{1.23}$$

And $S(0)g = g$ is the identity operator. This family satisfies the *semi-group property*,

$$S(t+s) = S(t)S(s)$$

To prove this, we use the Markov property,

$$\begin{aligned} S(s+t)g(x) &= E^x[g(X_{s+t})] = E^x[E^x g(X_{s+t})|\mathcal{F}_s] \\ &\overset{(1.18)}{=} E^x[E^{X_s}g(X_t)] = E^x[S(t)g(X_s)] \\ &= S(s)S(t)g(x) \end{aligned}$$

We want to introduce the notion of strong Markov property,when instead of conditioning in a fixed time, you use a stopping time. We need a preliminary notion.

Definition 1.32. *A Markov process $\{P^x\}$ is called Feller process if for all $t \geq 0$, and $x_j, x \in \mathbb{R}^d, g \in C_b(\mathbb{R}^d)$*

$$x_j \to x \textit{ implies } E^{x_j}[g(X_t)] \to E^x[g(X_t)] \tag{1.24}$$

Theorem 1.33. *Let $\{P^x\}$ be a Feller process with state space $\mathbb{R}^d$. Let $Y(s,\omega)$ be a bounded, jointly measurable function of $(s,\omega) \in \mathbb{R}_+ \times D$ and τ a stopping time on D. Let the initial state x be arbitrary. Then on the event $\{\tau < \infty\}$ the equality*

$$E^x[Y(\tau, X \circ \theta_\tau)|\mathcal{F}_\tau](\omega) = E^{\omega(\tau)}[Y(\tau(\omega), X)] \tag{1.25}$$

holds for P^x-almost every path ω.Equation (1.25) is the strong Markov property.

Remark 1.34. *The expectation on the right in (1.25) should read*

$$E^{\omega(\tau)}[Y(\tau(\omega), X)] = \int_D Y(\tau(\omega), \tilde{\omega}) P^{\omega(\tau)}(d\tilde{\omega})$$

Proof. Let $A \in \mathcal{F}_\tau$.We have to prove that

$$E^x[1_A 1_{\{\tau<\infty\}} Y(\tau, X \circ \theta_\tau)] = \int_{A\cap\{\tau<\infty\}} E^{\omega(\tau)}[Y(\tau(\omega), X)] P^x(d\omega) \tag{1.26}$$

Step 1. We assume that all the finite values of τ can be arranged in an increasing sequence $t_1 < t_2 < \cdots$.Then

$$\begin{aligned}
E^x[1_A 1_{\{\tau<\infty\}} Y(\tau, X \circ \theta_\tau)] &= \sum_n E^x[1_A 1_{\{\tau=t_n\}} Y(\tau, X \circ \theta_\tau)] \\
&= \sum_n E^x[1_A 1_{\{\tau=t_n\}} Y(t_n, X \circ \theta_{t_n})] \\
&= \sum_n E^x[1_A 1_{\{\tau=t_n\}} E^x\{Y(t_n, X \circ \theta_{t_n})|\mathcal{F}_{t_n}\}(\omega)] \\
&\overset{(1.18)}{=} \sum_n E^x[1_A 1_{\{\tau=t_n\}} E^{\omega(t_n)}\{Y(t_n, X)\}] \\
&= E^x[1_A 1_{\{\tau<\infty\}} E^{\omega(\tau)}\{Y(\tau(\omega), X)\}]
\end{aligned}$$

Step 2. If τ is any stopping time, we approximate by stopping times from Step 1. We define $\tau_n := 2^{-n}([2^n\tau]+1)$.Note that $\{\tau < \infty\} = \{\tau_n < \infty\}$ and $\tau_n \searrow \tau$ as $n \to \infty$.Since $\tau_n > \tau, A \in \mathcal{F}_\tau \subset \mathcal{F}_{\tau_n}$. The possible finite values of τ_n are $\{2^{-n}k : k \in \mathbb{N}\}$ and by Step 1 we know that

$$E^x[1_A 1_{\{\tau<\infty\}} Y(\tau_n, X \circ \theta_{\tau_n})] = \int_{A\cap\{\tau<\infty\}} E^{\omega(\tau_n)}[Y(\tau_n(\omega), X)] P^x(d\omega) \tag{1.27}$$

We want to take limits when $n \to \infty$.We will consider a special type of function Y as in (1.31). Take Y of the type

$$Y(s,\omega) = f_0(s) \cdot \prod_{i=1}^{m} f_i(\omega(s_i))$$

where $0 \leq s_1 < \cdots < s_m$ are time points and $f_0, f_1, \cdots f_m$ are bounded continuous functions.On the left of (1.27) we have inside the expectation (Recall that $X(\omega) = \omega$)

$$\begin{aligned} Y(\tau_n, \theta_{\tau_n}\omega) &= f_0(\tau_n) \cdot \prod_{i=1}^{m} f_i(\omega(\tau_n + s_i)) \\ &\overset{n\to\infty}{\to} f_0(\tau) \prod_{i=1}^{m} f_i(X \circ \omega(\tau + s_i)) = Y(\tau, \theta_\tau(\omega)) \end{aligned}$$

Where we used the right continuity of the path ω. Finally an application of the dominated convergence theorem gives the convergence of the left side of (1.27).
On the right side of (1.27) we have

$$\begin{aligned} E^{\omega(\tau_n)}[Y(\tau_n(\omega), X)] &= \int Y(\tau_n(\omega), \tilde{\omega}) P^{\omega(\tau_n)}(d\tilde{\omega}) \\ &= f_0(\tau_n(\omega)) \int \prod_{i=1}^{m} f_i(\tilde{\omega}(s_i)) P^{\omega(\tau_n)}(d\tilde{\omega}) \end{aligned}$$

The proof of the convergence in this case is an extension of the Feller property to functions of this type,and the proof follows similarly to the proof of (1.31).Then using dominated convergence we obtain the convergence in the right side of (1.27). Finally, we conclude applying standard extension theorem for all bounded and measurable functions as in the claim of the theorem.

□

Let's look an example of this:

Example 1.35. *Let $\tau = \inf\{t \geq 0 : X_t = x$ or $X_{t-} = z\}$ be the first hitting time of point z, which is a stopping time by Lemma 1.20.Suppose that the process is continuous, so that $P^x(C) = 1$ for all x.Suppose further that $P^x(\tau < \infty) = 1$.Therefore $P^x(X_\tau = z) = 1$.We will use (1.25) to prove that for fixed $t > 0, P^x(X_{\tau+t} \in A) = P^z(X_t \in A)$.Take $Y(\omega) = 1\{\omega(t) \in$*

$A\}$, so the function does not depend on time. We have

$$P^x(X_{\tau+t} \in A) = E^x[Y \circ \theta_\tau] = E^x[E^x(Y \circ \theta_\tau)|\mathcal{F}_\tau]$$
$$= E^x[E^{X(\tau)}Y] = E^x[E^zY] = E^zY = P^z(X_t \in A)$$

Remark 1.36. *In both Markov Properties we used the filtration $\{\mathcal{F}_t\}$. We will see that Brownian Motion satisfies the Markov property under $\{\mathcal{F}_{t+}\}$, and then also the strong Markov property under the same filtration, because the last proof is based on the fact that the Markov property is satisfied.*

1.5 Brownian motion

We define one the most important stochastic processes,the Brownian motion.

Definition 1.37. *Consider some probability space $(\Omega, \mathcal{F}, P)$,let $\mathcal{F}_t$ be a filtration and $B = \{B_t : t \in \mathbb{R}_+\}$ and adapted real-valued stochastic process. Then B is a one-dimensional Brownian motion with respect to $\{\mathcal{F}_t\}$ if it has these two properties:*

i) For almost every ω, the path $t \to B_t(\omega)$ is continuous.

ii) For $0 \leq s < t, B_t - B_s$ is independent of $\mathcal{F}_s$ and has normal distribution with mean zero and variance $t - s$.
If furthermore

iii) $B_0 \equiv 0$ almost surely.

then B is a standard Brownian motion. Because the dependence of the filtration ,we call B an $\{\mathcal{F}_t\}$-Brownian motion; the same occurs with the par $\{B_t, \mathcal{F}_t : 0 \leq t < \infty\}$

Point ii) in the definition says that if Z is bounded and $\mathcal{F}_s$ random variable and h is a bounded Borel function on $\mathbb{R}$,

$$E[Z \cdot h(B_t - B_s)] = E(Z) \cdot \frac{1}{\sqrt{2\pi(t-s)}} \int_{\mathbb{R}} h(x) exp\left\{-\frac{x^2}{2(t-s)}\right\} dx$$

By an inductive argument,for any $0 \leq s_0 < s_1 < \cdots < s_n$, the increments

$$B_{s_1} - B_{s_0}, B_{s_2} - B_{s_1}, \cdots, B_{s_n} - B_{s_{n-1}}$$

are independent random variables, and independent of $\mathcal{F}_{s_0}$.Furthermore, the joint distribution of the increments in the same as the joint distribution

$$B_{t+s_1} - B_{t+s_0}, B_{t+s_2} - B_{t+s_1}, \cdots, B_{t+s_n} - B_{t+s_{n-1}}$$

for all $t \geq 0$.These two properties say that Brownian motion has stationary,independent increments.

Definition 1.38. *A d-dimensional standard Brownian motion is an $\mathbb{R}^d-$value process $B_t = (B_t^1, \cdots B_t^d)$ where each component is a standard Brownian motion. This is equivalent to*

i) $B_0 = 0$ almost surely

ii) For almost every ω, the path $t \to B_t(\omega)$ is continuous.

iii) For $0 \leq s < t, B_t - B_s$ in independent of $\mathcal{F}_s$, and has multivariate normal distribution with mean zero and covariance matrix $(t-s)I_{d\times d}$

If we want to construct a Brownian motion with initial distribution μ, take a standard Brownian motion $(\tilde{B}_t, \tilde{\mathcal{F}}_t)$ and a $\mu-$distributed random variable X independent of $\tilde{\mathcal{F}}_\infty$, and define $B_t = \tilde{B}_t + X$.The filtration is $\mathcal{F}_t = \sigma(X, \tilde{\mathcal{F}}_t)$. Recall that if $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are three $\sigma-$algebras in the same probability space and $\sigma(\mathcal{B}, \mathcal{C})$ in independent of $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ are independent ,then $\mathcal{C}$ is independent of $\sigma(\mathcal{B}, \mathcal{C})$.In our case, $\mathcal{A} = \sigma(X), \mathcal{B} = \tilde{\mathcal{F}}_s, \mathcal{C} = \sigma(B_t - B_s)$, and use that $B_t - B_s = \tilde{B}_t - \tilde{B}_s$ is independent of $\mathcal{F}_s$ and X is independent of $\tilde{\mathcal{F}}_s$.Conversely, if a process B_t satisfies i) and ii) from Definition 1.37, then $\tilde{B}_t = B_t - B_0$ is a standard Brownian motion, independent of B_0.

I am going to omit the construction of a Brownian motion,but a complete proof is found in [1].The path space in this case is $C = C_{\mathbb{R}}[0, \infty)$,and $B_t(\omega) = \omega(t)$ are the coordinate projections on C,and $\mathcal{F}_t^B = \sigma(B_s : 0 \leq s \leq t)$ be the filtration generated by the coordinate process.We state the result for completeness.

Theorem 1.39. *There exists a Borel probability measure P^0 on the path space $C = C_{\mathbb{R}}([0, \infty))$ such that the process $B = \{B_t : t \in \mathbb{R}_+\}$ on the probability space $(C, \mathcal{B}(\mathcal{C}), P^0)$ is a standard one-dimensional Brownian motion with respect to the filtration $\{\mathcal{F}_t^B\}$*

The construction give the following regularity property on paths. Fix $0 < \gamma < \frac{1}{2}$.For P^0-almost surely $\omega \in C$,

$$\sup_{0\leq s<t\leq T} \frac{|B_t(\omega) - B_s(\omega)|}{|t-s|^\gamma} < \infty \quad \text{for all } T < \infty \tag{1.28}$$

1.5.1 Brownian motion as a martingale and a strong Markov process

We will prove some classical properties of Brownian motion in the one-dimensional case. The multivariate case usually follows from the previous one. Recall that a Brownian motion satisfies i) and ii) in Definition 1.37, and is called standard if $B_0 = 0$

Proposition 1.40. *Suppose $B = \{B_t\}$ is Brownian motion with respect to the filtration $\{\mathcal{F}_t\}$ on $(\Omega, \mathcal{F}, P)$.Then B_t and $B_t^2 - t$ are martingales respect to $\{\mathcal{F}_t\}$*

Proof. For $s < t, E[B_t|\mathcal{F}_s] = E[B_s + B_t - B_s|\mathcal{F}_s] = B_s + E[B_t - B_s] = B_s$,where we used that $B_t - B_s$ is independent of $\mathcal{F}_s$.Similarly,

$$\begin{aligned} E[B_t^2 - t|\mathcal{F}_s] &= E[(B_s + B_t - B_s)^2 - s + (s-t)|\mathcal{F}_s] \\ &= B_s^2 - s + 2B_sE[(B_t - B_s)|\mathcal{F}_s] + E[(B_t - B_s)^2 - (t-s)] = B_s^2 - s \end{aligned}$$

where we used that $Var(B_t - B_s) = t - s$ □

Proposition 1.41. *Suppose $B = \{B_t\}$ is Brownian motion with respect to the filtration $\{\mathcal{F}_t\}$ on $(\Omega, \mathcal{F}, P)$.*

i) We can assume that $\mathcal{F}_t$ contains the null sets .Furthermore, $B = \{B_t\}$ is also a Brownian motion with respect to the right-continuous filtration $\{\mathcal{F}_{t+}\}$

ii) Fix $x \in \mathbb{R}_+$, and define $Y_t = B_{s+t} - B_s$.Then the process $Y = \{Y_t : t \in \mathbb{R}_+\}$ in independent of $\{\mathcal{F}_{s+}\}$ and it is a standard Brownian motion with respect to the filtration $\{\mathcal{G}_t\}$ defined by $\mathcal{G}_t = \mathcal{F}_{(s+t)+}$

Proof. In (1.2) we saw how complete the filtration.The adaptedness is not changed, so it's only necessary to check the independence of the increments $B_t - B_s$ with $\overline{\mathcal{F}_t}$.If $G \in \mathcal{F}$ has $A \in \mathcal{F}_s$ such that $P(A \triangle G) = 0$, then $P(G \cap H) = P(A \cap H)$ for any event H.This implies the independence of $B_t - B_s$ with $\overline{\mathcal{F}_t}$.

To conclude the proof, we consider $s \geq 0$ fixed, and $0 = t_0 < t_1 < t_2 < \cdots < t_n$, and for $h \geq 0$ we write

$$\xi(h) := (B_{s+h+t_1} - B_{s+h}, B_{s+h+t_2} - B_{s+h+t_1}, \cdots B_{s+h+t_n} - B_{s+h+t_{n-1}})$$

and let Z be a bounded $\mathcal{F}_{s+}$-measurable random variable.Clearly,for $h > 0$,Z is $\mathcal{F}_{s+h}$-measurable, and is independent of $\xi(h)$.Take f be a bounded continuous function on $\mathbb{R}^d$. Applying path continuity, independence and the stationarity of Brownian motion,

$$\begin{aligned}
E[Z \cdot f(\xi(0))] &= \lim_{h \searrow 0} E[Z \cdot f(\xi(h))] = \lim_{h \searrow 0} EZ \cdot E[f(\xi(h))] \\
&= EZ \cdot E[f(\xi(0))] = EZ \cdot E[f((B_{s+t_1} - B_s, B_{s+t_2} - B_{s+t_1}, \cdots B_{s+t_n} - B_{s+t_{n-1}}))] \\
&= EZ \cdot E[f(\xi(0))] = EZ \cdot E[f((B_{t_1} - B_0, B_{t_2} - B_{t_1}, \cdots B_{t_n} - B_{t_{n-1}}))]
\end{aligned}$$

This equality extends by density to bounded and Borel f.

The last equality implies that Y is independent of $\{\mathcal{F}_{s+}\}$,so B is a Brownian motion respect to $\{\mathcal{F}_{t+}\}$.Note that $\xi(0) = (Y_{t_1}, Y_{t_2} - Y_{t_1}, \cdots, Y_{t_n} - Y_{t_{n-1}})$ and $\eta := (Y_{t_1}, Y_{t_2}, \cdots Y_{t_n})$ is function of $\xi(0)$,we conclude the independence of η with $\{\mathcal{F}_{s+}\}$.But the election of points $0 = t_0 < t_1 < \cdots < t_n$ implies that the entire process Y is independent of $\{\mathcal{F}_{s+}\}$,and the last member of the equation above shows that given $\mathcal{F}_{s+}$, Y has the distribution of a standard Brownian motion.Finally, the independence of $Y_{t_2} - Y_{t_1}$ and $\mathcal{G}_{t_1}$ is the same as the independence of $B_{s+t_2} - B_{s+t_1}$ and $\mathcal{F}_{(s+t_1)+}$,which was already proven. That concludes the proof. □

Using i) and ii) in the last theorem we deduce that Y is a standard Brownian motion,independent of $\overline{\mathcal{F}_{t+}}$.

Observe that in the calculation in the proof ot the last proposition, it's only used the right continuity, so we can apply the last result to more general processes.

Lemma 1.42. *Suppose that $X = \{X_t : t \in \mathbb{R}_+\}$ is a right-continuous process adapted to the filtration $\{\mathcal{F}_t\}$ and for all $s < t$ the increments $X_t - X_s$ are independent of $\mathcal{F}_s$.Then $X_t - X_s$ is independent of $\overline{\mathcal{F}_{s+}}$*

Now we prove some results in the canonical space C.Here, $B_t(\omega) = \omega(t)$ for $\omega \in C$.For $x \in \mathbb{R}$, P^x denote the probability measure on C under which $B = \{B_t\}$ is a Brownian motion started at x.The expectation E^x satisfies

$$E^x(H) = E^0[H(x+B)]$$

for any function H $\mathcal{B}_C$ measurable and bounded.

Again, we consider the shifts $\{\theta_s : s \in \mathbb{R}_+\}$,and act on B as $\theta_s B = \{B_{t+s} : t \geq 0\}$

Using Proposition 1.41, B is also a Brownian motion relative to the filtration $\mathcal{F}_+^B = \cap_{s:s>t}\mathcal{F}_s^B$.It can be proved that the $\sigma-$algebras $\mathcal{F}_{t+}^B, \mathcal{F}_t^B$ are different.However, we will see that only differs by null sets.

Now we prove the Markov property respect to $\{\mathcal{F}_{t+}^B\}$

Proposition 1.43. *Let H be a bounded $\mathcal{B}_C$-measurable function on C.*

i) $E^x[H]$is Borel measurable function of x.

ii) For each $x \in \mathbb{R}$,

$$E^x[H \circ \theta_s | \mathcal{F}_{s+}^B](\omega) = E^{B_s(\omega)}[H] \quad \textit{for } P^x - \textit{almost every } \omega \tag{1.29}$$

In particular, the family $\{P^x\}$ on C satisfies Definition 1.29 of a Markov process with respect to the filtration $\{\mathcal{F}_{t+}^B\}$

Proof.

i) It's enough to prove that the map $x \rightarrow P^x(F)$ is measurable for every closed set F. In that case, we can extend this result to any F measurable, so the result is valid for $H = 1_F$, and with this we can extend to every measurable bounded function H.To show the claim for closed sets, consider H a continuous bounded function on C. Take $x_j \rightarrow x$ in $\mathbb{R}$. Using the continuity of H and dominated convergence,

$$E^{x_j}[H] = E^0[H(x_j + B)] \rightarrow E^0[H(x+B)] = E^x[H]$$

so $E^x(H)$ is a continuous function of x, an thus measurable.Now,approximate 1_F with F closed by continuous functions H_n.Therefore, $E^x(1_F) = \lim_{n\rightarrow\infty} E^x[H_n]$ is also Borel measurable in x.

ii) Write $\theta_s B = B_s + Y$, with $Y_t = B_{t+s} - B_s$. Let Z be a bounded $\mathcal{F}^B_{s+}-$measurable random variable. By ii) in Proposition 1.41, Y is a standard Brownian motion, independent of (Z, B_s),because the latter is $\mathcal{F}^B_{s+}$ measurable. We have

$$E^x[Z \cdot H(\theta_s B)] = E^x[Z \cdot H[B_s + Y]] = \int_C E^x[Z \cdot H[B_s + \zeta]]P^0(d\zeta)$$

Here we used the following fact:

If $X : \Omega \to U, Y : \Omega \to V$ are random variables in the same probability space,U,V measure spaces, and $f : U \times V \to \mathbb{R}$ is bounded and measurable. If X, Y are independent and μ is the distribution of Y,the

$$E[f(X,Y)] = \int_V E[f(X,y)]\mu(dy)$$

In our case, P^0 is the distribution of Y because it's a standard Brownian motion.We have

$$\begin{aligned}\int_C E^x[Z \cdot H[B_s + \zeta]]P^0(d\zeta) &= E^x\{Z \cdot \int_C H[B_s + \zeta]P^0(d\zeta)\} \\ &= E^x[Z \cdot E^{B_s}(H)]\end{aligned}$$

We proved that $E^x[Z \cdot H(\theta_s B)] = E^x[Z \cdot E^{B_s}(H)]$ for every $\mathcal{F}^B_{s+}$ measurable random variable Z.Now take $Z = 1_A$ with A an $\mathcal{F}^B_{s+}$ measurable set,and this is the definition of conditional expectation.

□

Proposition 1.44. *Let H be a bounded $\mathcal{B}_C$-measurable function on C. Then for any $x \in \mathbb{R}$ and $0 \le s < \infty$,*

$$E^x[H|\mathcal{F}^B_{s+}] = E^x[H|\mathcal{F}^B_s] \quad P^x - \textit{almost surely} \tag{1.30}$$

Proof. Suppose first H is of the type

$$H(\omega) = \prod_{i=1}^n 1_{A_i}(\omega(t_i))$$

for some $0 \le t_1 < t_2 < \cdots < t_n$ and $A_i \in \mathcal{B}_\mathbb{R}$.We separate those factors where $t_i \le s$,so $H = H_1 \cdot H_2 \circ \theta_s$, where H_1 is $\mathcal{F}_s$-measurable.Thus

$$E^x[H|\mathcal{F}^B_{s+}] = H_1 E^x[H_2 \circ \theta_s|\mathcal{F}^B_{s+}] \stackrel{(1.29)}{=} H_1 E^{B_s}[H_2]$$

and $E^x[H|\mathcal{F}^B_{s+}]$ is a $\mathcal{F}^B_s$-measurable function. As $\mathcal{F}^B_s \subset \mathcal{F}^B_{s+}$, we conclude the claim by (1.19). To conclude,apply the standard machinery to extend this to all bounded $\mathcal{B}_C$-measurable functions. □

Corollary 1.45. *If $A \in \mathcal{F}^B_{t+}$ then there exists $B \in \mathcal{F}^B_t$ such that $P^x(A \triangle B) = 0$.*

Proof. Let $Y = E^x(1_A|\mathcal{F}^B_t)$.Note that by (1.30) $1_A = E^x(1_A|\mathcal{F}^B_{t+}) = Y$ P^x-almost surely.Take $B = \{Y = 1\} \in \mathcal{F}^B_t$.We claim that $A \triangle B \subset \{1_A \neq Y\}$.To see this, if $\omega \in B \setminus A, Y(\omega) = 1 \neq 0 = 1_A(\omega)$. If $\omega \in A \setminus B, 1_A(\omega) = 1 \neq Y(\omega)$.That proves the claim, and also the corollary. □

Corollary 1.46. *(Blumenthal's 0-1 Law)Let $x \in \mathbb{R}$.Then for $A \in \mathcal{F}^B_{0+}, P^x(A) \in \{0, 1\}$.*

Proof. Note that $A \in \mathcal{F}^B_0 = \sigma(B_0)$, then $A = \{B_0 \in G\}$ for some $\mathcal{B}_C-$measurable set, so $P^x(A) = P^x(B_0 \in G) = 1_G(x) \in \{0, 1\}$.Then $\mathcal{F}^B_0$ satisfies the 0-1 law under P^x.In particular, for every $X \in L^1(P^x), E^x[X|\mathcal{F}^B_0] = E^x[X]$. The following equalities are valid P^x-almost surely for $A \in \mathcal{F}^B_{0+}$:

$$1_A = E^x(1_A|\mathcal{F}^B_{0+}) = E^x(1_A|\mathcal{F}_0) = P^x(A)$$

Thus there exists points $\omega \in C$ such that $1_A(\omega) = P^x(A)$,and so the only possible values for $P^x(A)$ are 0 or 1. □

From the 0-1 law we deduce the following fact: for almost every ω,given $\epsilon > 0$, there exists some $t \in (0, \epsilon)$ such that $B_t(\omega) = 0$

Corollary 1.47. *Define $\sigma := \inf\{t > 0 : B_t > 0\}, \tau := \inf\{t > 0 : B_t < 0\}$ and $T_0 := \inf\{t > 0 : B_t = 0\}$*

Proof. We show that $\{\sigma = 0\}, \{\tau = 0\} \in \mathcal{F}^B_{0+}$.Write

$$\{\sigma = 0\} = \bigcap_{m=n}^{\infty} \{B_q > 0 \text{ for some rational } q \in (0\frac{1}{m})\} \in \mathcal{F}^B_{\frac{1}{n}}$$

as $n \in \mathbb{N}$ is arbitrary, we deduce the claim, and the same is for τ.As each variable B_t is a centered Gaussian,

$$P^0(\sigma \leq \frac{1}{m}) \geq P^0(B_{1/m} > 0) = \frac{1}{2}$$

Taking $m \to \infty, P^0(\sigma = 0) \geq \frac{1}{2}$, and using the 0-1 law, $P(\sigma = 0) = 0$.The same argument applies for τ.

Finally, take ω such that $\sigma(\omega) = \tau(\omega) = 0$.There exists some $\epsilon > 0$ and $t, s \in (0, \epsilon)$ with $B_t(\omega) > 0, B_s(\omega) < 0$ and by continuity, $B_l(\omega) = 0$ for some $l \in (0, \epsilon)$.This says that $T_0(\omega) \leq \epsilon$.As ϵ is arbitrary, $T_0(\omega) = 0$ □

Definition 1.48. *The transition probabilities of Brownian motion is a normal distribution given by*

$$p(t, x, A) := P^x(B_t \in A) = \int_A p(t, x, y)dy \quad \text{for } A \in \mathcal{B}_{\mathbb{R}}$$

with

$$p(t, x, y) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}} \tag{1.31}$$

The transition probability density of Brownian motion is called Gaussian kernel.

Recall Theorem 1.33. To prove that Brownian motion is a strong Markov process, we only need to check that is a Feller process.

Proposition 1.49. *Brownian motion is a Feller process, and therefore a strong Markov process under the filtration $\{\mathcal{G}_t\}$,where $\mathcal{G}_t := \mathcal{F}^B_{t+}$.*

Proof. For $g \in C_b(\mathbb{R})$,

$$E^x[g(B_t)] = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} g(x + y)e^{-y^2/2t}dy$$

and the continuity in x follows from dominated convergence theorem. □

The strong Markov property for Brownian motion can be seen as follow: if τ is a stopping time, on the event $\{\tau < \infty\}$, $\tilde{B}_t := B_{t+\tau} - B_\tau$ is a standard Brownian motion,independent of $\mathcal{G}_\tau$.To show this, for $g \in C_b(\mathbb{R})$,let $h(\omega) := g(\omega - \omega(0))$. Then

$$E^x[h] = E^0[h(x + B)] = E^x[h(x + \omega)] = E^0[g(\omega)] = E^0[g]$$

Therefore

$$E^x[g(\tilde{B})|\mathcal{G}_\tau](\omega) = E^x[h \circ \theta_\tau|\mathcal{G}_\tau](\omega) = E^{\omega(\tau)}[h] = E^0[g]$$

This says that under $\mathcal{G}_\tau$, $\tilde{B}$ is a standard Brownian motion. Also,we have $E^x[g(\tilde{B})] = E^0[g]$, so if Z is $\mathcal{G}_\tau$- measurable,

$$E^x[g(\tilde{B})Z] = E^x[E^0(g)Z] = E^x[Z]E^0[g] = E^x[Z]E^x[g(\tilde{B})]$$

and it gives us the independence.

Remark 1.50. *Both Markov properties holds for d-dimensional Brownian motion.*

An important application of the strong Markov property is the *reflection principle*.Define the *running maximum* of Brownian motion by

$$M_t := \sup_{0 \le s \le t} B_s \tag{1.32}$$

Proposition 1.51. *(Reflection principle) Let $a \le b$ and $b > 0$ be real numbers. Then*

$$P^0(B_t \le a, M_t \ge b) = P^0(B_t \ge 2b - a) \tag{1.33}$$

Each inequality can be strict or weak by taking limits,and using the continuity of B_t.

Proof. Let $\tau_b := \inf\{t \ge 0 : B_t = b$.Note that by continuity, $M_t \ge b \leftrightarrow \tau_b \le t$.Define

$$Y(s, \omega) := 1\{s \le t, \omega(t-s) \ge 2b - a\} - 1\{s \le t, \omega(t-s) \le a\}$$

By symmetry of Brownian motion, for any $s \le t$,

$$E^b[Y(s, B)] = P^b(B_{t-s} \ge 2 - a) - P^b(B_{t-s} \le a)$$

$$P^0(B_{t-s} \ge b - a) - P^0(B_{t-s} \le a - b) = 0$$

We will use that $b = B_{\tau_b}(\omega)$, and X is the identity operator on C.Recall the equation (1.25):

$$\begin{aligned}
0 &= E^0[1\{\tau_b \le t\}E^b(Y(\tau_b(\omega), X))] \\
&= E^0[1\{\tau_b \le t\}E^0(Y(\tau_b, X \circ \theta_{\tau_b})|\mathcal{F}_{\tau_b})(\omega)] \\
&= E^0[1\{\tau_b \le t\}Y(\tau_b, X \circ \theta_{\tau_b})] \\
&= P^0(\tau_b \le t, B_t \ge 2b - a) - P^0(\tau_b \le t, B_t \le a) \\
&= P^0(B_t \ge 2b - a) - P^0(\tau_b \le t, B_t \le a)
\end{aligned}$$

In the third equality we used the definition of conditional expectation $(1\{\tau_b \le t\})$ is $\mathcal{F}_\tau$-measurable. In the fourth line we used the definition of Y to obtain

$$Y(\tau_b, B \circ \tau_b) = 1\{\tau_b \le t, B_t \ge 2b - a\} - 1\{\tau_b \le t, B_t \le a\}$$

In the last equality we used that $2b - a \ge b$ and $B_t \ge 2b - a$ implies $\tau_b \le t$. $\square$

Corollary 1.52. *M_t has the same distribution as $|B_t|$*

Proof. We take $a = b > 0$ in (1.33):

$$P^0(M_t \geq b) = P^0(B_t \leq b, M_t \geq b) + P^0(B_t > b, M_t \geq b)$$
$$\overset{(1.33)}{=} P^0(B_t \geq b) + P^0(B_t > b) = 2P^0(B_t \geq b) = P^0(|B_t| \geq b)$$

□

1.5.2 Path regularity of Brownian motion

The following is a consequence of the construction of Brownian motion:

Theorem 1.53. *Fix $0 < \gamma < \frac{1}{2}$.The following is true almost surely for Brownian motion:for every $T < \infty$ there exists a finite constant $C(\omega)$ such that*

$$|B_t(\omega) - B_s(\omega)| \leq C(\omega)|t - s|^\gamma \quad \textit{for all } 0 \leq s, t \leq T \tag{1.34}$$

We prove the opposite case for $\gamma > 1/2$: there is not local Hölder continuity in that case.

Theorem 1.54. *Let B be a Brownian motion.For finite positive reals γ, C and ϵ define the event*

$$G(\gamma, C, \epsilon) := \{\textit{there exists } s \in \mathbb{R}_+ \textit{ such that } |B_t - B_s| \leq C|t - s|^\gamma \textit{ for all } t \in [s - \epsilon, s + \epsilon]\}$$

Then if $\gamma > 1/2, P(G(\gamma, C, \epsilon)) = 0$ for all positive C and ϵ.

Proof. Fix $\gamma > 1/2$.WLOG we can assume that B is a standard Brownian motion,because the increments of B_t and $B_t - B_0$ are the same.Define

$$H_k(C, \epsilon) := \{\text{there exists } s \in [k, k+1] \text{ such that } |B_t - B_s| \leq C|t - s|^\gamma \text{ for all } t \in [s - \epsilon, s + \epsilon] \cap [k, k+1]\}$$

Note that $G(\gamma, C, \epsilon) \subset \cup_k H_k(C, \epsilon)$,so it's enough to show that $P(H_k(C, \epsilon)) = 0$ for each k.As $Y_t = B_{k+t} - B_k$ is a standard Brownian motion,then $P(H_k(C, \epsilon)) = P(H_0(C, \epsilon))$,it suffices to check the case $k = 0$.

Fix $m \in \mathbb{N}$ with $m(\gamma - \frac{1}{2}) > 1$.Let $\omega \in H_0(C, \epsilon)$ and pick $s \in [0, 1]$ so that $H_0(C, \epsilon)$ is satisfied.Consider n large enough so $m/n < \epsilon$. Imagine partitioning $[0, 1]$ into intervals of length $1/n$.Let

$$X_{n,k} := \max\{|B_{(j+1)/n} - B_{j/n}| : k \leq j \leq k + m - 1\} \text{ for } 0 \leq k \leq n - m$$

The point s has to lie in some interval $[\frac{k}{n}, \frac{k+m}{n}]$ for some $0 \le k \le n-m$.For such k,

$$\begin{aligned}|B_{(j+1)/n} - B_{\frac{j}{n}}| &\le |B_{(j+1)/n} - B_s| + |B_s - B_{j/n}| \\ &\le C(|\frac{j+1}{n} - s|^\gamma + |s - \frac{j}{n}|^\gamma) \le 2C(\frac{m}{n})^\gamma\end{aligned}$$

for all $k \le j \le k+m-1$ (because all the point in $[\frac{k}{n}, \frac{k+m}{n}]$are at distance at most $m/n < \epsilon$ from s).Thus $X_{n,k} \le 2C(\frac{m}{n})^\gamma$.We have then $P(H_0(C,\epsilon)) \le P(\cup_{k=0}^{n-m}\{X_{n,k} \le 2C(\frac{m}{n})^\gamma\})$.Has the increments of Brownian motion are independent and stationary, and using that $B_t \sim t^{1/2}B_1$,

$$\begin{aligned}P(H_0(C,\epsilon)) &\le nP(X_{n,0} \le 2C(\frac{m}{n})^\gamma) = n\prod_{j=0}^{m-1} P\left(|B_{(j+1)/n} - B_{j/n}| \le 2C(\frac{m}{n})^\gamma\right) \\ &= nP\left(|B_{1/n}| \le 2C(\frac{m}{n})^\gamma\right)^m = nP(|B_1| \le 2Cn^{1/2-\gamma}m^\gamma)^m \\ &= n\left(\frac{1}{\sqrt{2\pi}}\int_{-2Cn^{1/2-\gamma}m^\gamma}^{2Cn^{1/2-\gamma}m^\gamma} e^{-x^2/2}dx\right)^m \le n\left(\frac{1}{\sqrt{2\pi}}4Cn^{1/2-\gamma}m^\gamma\right)^m \\ &\le K(m)n^{1-m(1/2-\gamma)}\end{aligned}$$

Finally, recall that $m(1/2-\gamma) > 1$,so this term converges to zero when $n \to \infty$ □

Corollary 1.55. *The following is true almost surely for Brownian motion: the path $t \to B_t(\omega)$ is not differentiable at any time point.*

Proof. If $t \to B_t(\omega)$ is differentiable at some point s, then the limit

$$\xi = \lim_{t\to s}\frac{B_t(\omega) - B_s(\omega)}{t-s}$$

If an integer M satisfies $M > |\xi| + 1$,we can find another integer k such that for all $t \in [s-k^{-1}, s+k^{-1}]$,

$$-M \le \frac{B_t(\omega) - B_s(\omega)}{t-s} \le M$$

this implies

$$|B_t(\omega) - B_s(\omega)| \le M|t-s|$$

□

so $\omega \in G(1, M, k^{-1})$. So if $B_t(\omega)$ is differentiable any point , then $\omega \in \cup_M \cup_k G(1, M, k^{-1})$, and this set has zero measure.

Corollary 1.56. *Brownian motion has not bounded variation paths almost surely.*

Proof. Any BV function is difference of two nondecreasing functions.As these functions are differentiable in Lebesgue-almost every point, we deduce that Brownian paths are not of bounded variation for every interval. $\square$

Although Brownian motion is not of BV, its quadratic variation is finite.

Proposition 1.57. *Let B be a Brownian motion. For any partition π of $[0,t]$,*

$$E\left[\left(\sum_{i=0}^{m(\pi)-1}(B_{t_{i+1}}-B_{t_i})^2-t\right)^2\right]\le 2t mesh(\pi) \tag{1.35}$$

In particular,

$$\lim_{mesh(\pi)\to 0}\sum_{i=0}^{m(\pi)-1}(B_{t_{i+1}}-B_{t_i})^2=t \quad in\ L^2(P) \tag{1.36}$$

If we have a sequence of partitions π^nsuch that $\sum_n mesh(\pi^n)<\infty$,then the convergence above holds almost surely along this sequence

Proof. We will use the following facts of Brownian motion and normal distribution: independent increments, $B_s-B_r\sim N(0,s-r), E((B_s-B_r)^4)=3(r-s)^2$.Let $\triangle t_i := t_{i+1}-t_i$.

$$\begin{aligned}
E\left[\left(\sum_{i=0}^{m(\pi)-1}(B_{t_{i+1}}-B_{t_i})^2-t\right)^2\right] &= \sum_i E[(B_{t_{i+1}}-B_{t_i})^4]+\sum_{i\neq j}[(B_{t_{i+1}}-B_{t_i})^2(B_{t_{j+1}}-B_{t_j})^2]\\
&-2t\sum_i E[(B_{t_{i+1}}-B_{t_i})^2]+t^2\\
&=3\sum_i(\triangle t_i)^2+\sum_{i\neq j}(\triangle t_i)(\triangle t_j)-2t\sum_i\triangle_i+t^2\\
&=3\sum_i(\triangle t_i)^2+\sum_{i\neq j}(\triangle t_i)(\triangle t_j)-2t^2+t^2\\
&=2\sum_i(\triangle t_i)^2+\sum_i(\triangle t_i)^2+\sum_{i\neq j}(\triangle t_i)(\triangle t_j)-t^2\\
&=2\sum_i(\triangle t_i)^2+\sum_{i,j}(\triangle t_i)(\triangle t_j)-t^2\\
&=2\sum_i(\triangle t_i)^2\le 2mesh(\pi)t
\end{aligned}$$

This proves the first two assertions.

Using Chebychev's inequality,

$$P\left(\left|\sum_{i=0}^{m(\pi)-1}(B_{t_{i+1}^n}-B_{t_i}^n)^2-t\right|\geq\epsilon\right)\leq e^{-2}E\left[\left(\sum_{i=0}^{m(\pi)-1}(B_{t_{i+1}^n}-B_{t_i}^n)^2-t\right)\right]$$
$$\leq 2t\epsilon^{-2}mesh(\pi^n)$$

If $\sum_n mesh(\pi^n)<\infty$, we conclude by Borel-Cantelli the almost surely convergence. □

1.6 Poisson processes

Definition 1.58. *Let $0<\alpha<\infty$.A nonnegative integer valued random variable X has Poisson distribution with parameter α (Poisson(α)-distribution) if*

$$P(X=k)=e^{-\alpha}\frac{\alpha^k}{k!}\quad for\ k\in\mathbb{Z}_+$$

The case $\alpha=0$ is a random variable X with $P(X=0)=1$, and the case $\alpha=\infty$ is a random variable X with $P(X=\infty)=1$.

The following are some facts about Poisson random variables:

Lemma 1.59.

i) If $X\sim Poisson(\alpha)$,then $E(X)=Var(X)=\alpha$.

ii) If $X_n\sim Poisson(\alpha_n)$ for $n\in\mathbb{N}$ are independent random variables, then $\sum_n X_n\sim Poisson(\sum_n\alpha_n)$

Definition 1.60. *Let $(S,\mathcal{A},\mu)$ be a $\sigma-$finite measure space. A process $\{N(A):A\in\mathcal{A}\}$ indexed by the measurable sets is a Poisson point process with mean measure μ if*

i) Almost surely, $N(\cdot)$ is a $\mathbb{Z}\cup\{\infty\}$-valued measure on $(S,\mathcal{A})$.

ii) N(A) is Poisson distributed with parameter $\mu(A)$.

iii) For any pairwise disjoint $A_1,A_2,\cdots A_n\in\mathcal{A}$,the random variables $N(A_1),\cdots N(A_n)$ are independent.

Remark 1.61. *Parts i) and ii) in the definition above give a complete description of the finite-dimensional distributions of $\{N(A)\}$. For arbitrary $B_1, B_2 \cdots B_m \in \mathcal{A}$,there exists disjoints sets $A_1, \cdots A_n \in \mathcal{A}$ such that each B_j is union of some $A_j's$,so $N(B_j)$ is a certain sum of $N(A_i)'s$(because N is a measure),and then the joint distribution of $N(B_1), \cdots N(B_m)$ is determined by the joint distribution of $N(A_1), \cdots N(A_n)$.*

Proposition 1.62. *Let $(S, \mathcal{A}, \mu)$ be a σ-finite measure space. Then exists a Poisson point process $\{N(A) : A \in \mathcal{A}\}$ with mean measure μ.*

Proof. Let $S_1, S_2, \cdots$ be disjoint measurable sets such that $S = \cup_i S_i$ and $\mu(S_i) < \infty$.First we define a Poisson point process N_i supported on S_i.If $\mu(S_i) = 0$, define $N_i \equiv 0$. If $0 < \mu(S_i) < \infty$, let $\{X_j^i : k \in \mathbb{N}\}$ be i.i.d S_i-valued random variables with common probability distribution

$$P(X_j \in B) = \frac{\mu(B \cap S_i)}{\mu(S_i)} \text{ for measurable sets } B \in \mathcal{A}.$$

Independently of the $\{X_j^i : k \in \mathbb{N}\}$, let K_i be a $Poisson(\mu(S_i))$ random variable. Define

$$N_i(A) := \sum_{j=1}^{K_i} 1_A(X_j^i) \text{ for measurable sets } A \in \mathcal{A}.$$

We show that N_i is a Poisson point process whose mean measure is μ_i given by $\mu_i(A) := \mu(A \cap S_i)$. To show this, we prove first that if $A \in \mathcal{A}$,then $N_i(A) \sim Poisson(\mu(A \cap S_i))$.We have

$$\begin{aligned}
P(N_i(A) = k) &= P(K_i \geq k, N_i(A) = k) = \sum_{l=k}^{\infty} P(K_i = l, \sum_{j=1}^{l} 1_A(X_j^i) = k) \\
&\overset{indep}{=} \sum_{l=k}^{\infty} e^{-\mu(S_i)} \frac{\mu(S_i)^l}{l!} P(\sum_{j=1}^{l} 1_A(X_j^i) = k) = e^{-\mu(S_i)} \sum_{l=k}^{\infty} \frac{\mu(S_i)^l}{l!} \binom{l}{k} \frac{\mu(A \cap S_i)^k}{k!(\mu(S_i)^l)} \frac{\mu(A^c \cap S_i)^{l-k}}{(l-k)!} \\
&= \mu(A \cap S_i)^k \frac{1}{k!} e^{-\mu(S_i)} \sum_{l=k}^{\infty} \frac{1}{(l-k)!} \frac{\mu(A^c \cap S_i)^{l-k}}{(l-k)!} = \mu(A \cap S_i)^k \frac{1}{k!} e^{-(\mu(A \cap S_i))}
\end{aligned}$$

Therefore $N_i(A) \sim Poisson(\mu(A \cap S_i))$.To prove the independence, we consider $A_1, \cdots A_n$ pairwise disjoint, and assume that $\cup_i A_i = S_i$(otherwise, add the complement).Let $k_1, \cdots, k_n \in$

$\mathbb{N}$ and $k := \sum k_1 + \cdots k_n$.We have

$$
\begin{aligned}
P(N_i(A_1) = k_1, \cdots N_i(A_n = k_n) &= P(K_i = k)P(N_i(A_1) = k_1, \cdots N_i(A_n = k_n | K_i = k) \\
&= \mu(S_i)^k \frac{e^{-(\mu(S_i))}}{k!} P(\sum_{j=1}^{k} 1_{A_1}(X_j^i) = k_1, \cdots \sum_{j=1}^{k} 1_{A_n}(X_j^i) = k_n) \\
&= \mu(S_i)^k \frac{e^{-(\mu(S_i))}}{k!} \frac{k!}{k_1! \cdots k_n!} \prod_{j=1}^{n} \frac{\mu(A_j \cap S_i)^{k_j}}{\mu(S_i)^{k_j}} \\
&= \prod_{j=1}^{n} \mu(A_j \cap S_i) \frac{e^{-\mu(S_i) \cap A_j}}{k_j!} \\
&= \prod_{j=1}^{n} P(N_i(A_j) = k_j)
\end{aligned}
$$

That concludes the claim.

We can repeat this construction to each S_i, and take the resulting random processes N_i mutually independent by a suitable product space construction.Finally, define

$$N(A) = \sum_i N_i(A)$$

Then, N(A) satisfies the required. □

Definition 1.63. *Let $(\Omega, \mathcal{F}, P)$ be a probability space ,$\{\mathcal{F}_t\}$ a filtration on it, and $\alpha > 0$. A (homogeneous) Poisson process with rate α is an adapted stochastic process $N = \{N_t : t \in \mathbb{R}_+\}$ with the following properties:*

i) $N_0 = 0$ almost surely.

ii) For almost every ω, the path $t \to N_t(\omega)$ is cadlag.

iii) For $0 \leq s < t, N_t - N_s$ is independent of $\mathcal{F}_s$, and has Poisson distribution with parameter $\alpha(t-s)$

Proposition 1.64. *Homogeneous Poisson processes on $[0, \infty)$ exist.*

Proof. Let $\{N(A) : A \in \mathcal{B}((0, \infty))\}$ be a Poisson point process on $(0, \infty)$ with mean measure αm,and m is Lebesgue measure.Define $N_0 = 0, N_t = N(0, t]$ for $t > 0$,and $\mathcal{F}_t^N = \sigma\{N_s : 0 \leq s \leq t\}$.Let $0 = s_0 < s_1 < \cdots < s_n \leq s < t$.Then

$$N(s_0, s_1], N(s_1, s_2], \cdots, N(s_{n-1}, s_n], N(s, t]$$

are independent random variables, from which follows that the vector $(N_{s_1}, \cdots, N_{s_n})$ is independent of $N(s,t]$.If we consider all such n-tuples while fixing $s < t$ (and any n) we conclude that $\mathcal{F}_s^N$ is independent of $N(s,t] = N_t - N_s$.By definition,it's clear that $N_t - N_s$ is Poisson of parameter $\alpha(t-s)$.To prove the cadlag property, note that almost every ω has the property that $N(0,T] < \infty$ for all $T < \infty$.For such ω and t, there exists $t_0 < t < t_1$ with $N(t_0,t) = N(t,t_1) = 0$(Recall that $N(\cdot)$ is an integer measure on $(0,\infty)$).Consequently, N_s is constant for $t_0 < s < t$(because $N_s - N_{s'} = N(s,s'] \leq N(t_0,t) = 0$ if $t_0 < s < s' < t$), and so the left limit N_{t-} exists.Also, $N_s = N_t$ for $t \leq s < t_1$, which gives the right continuity at t. □

We have an analogue to Brownian motion(so the proof is omitted):

Proposition 1.65. *Suppose $N = \{N_t\}$ is a homogeneous Poisson process with respect to the filtration $\{\mathcal{F}_t\}$ on $(\Omega, \mathcal{F}, P)$.*

i) N is a Poisson process also with respect to the augmented right-continuous filtration $\{\overline{\mathcal{F}_{t+}}\}$

ii) Define $Y_t = N_{s+t} - N_s$ and $\mathcal{G}_t := \mathcal{F}_{(s+t)+}$.Then $Y = \{Y_t : t \in \mathbb{R}_+\}$ is a homogeneous Poisson process with respect to the filtration $\{\mathcal{G}_t\}$, and independent of $\overline{\mathcal{F}_{s+}}$

Note that the Poisson process is monotone nondecreasing , so it cannot be a martingale. However, if we define the process

$$M_t := N_t - \alpha t$$

then M is a martingale.

Proposition 1.66. *M is a martingale.*

Proof. Using the independence of increments,

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

□

Remark 1.67. *The Markov Property for the Poisson process follows as with Brownian motion.In this case, the state space is $\mathbb{Z}$(or $\mathbb{Z}_+$).A Poisson process with initial state $x \in \mathbb{Z}$*

would be defined as $x + N_t$*, and* P^x *would be the distribution of* $\{x + N_t\}_{t \in \mathbb{R}_+}$ *on the space* $D_{\mathbb{Z}}[0, \infty)$ *of* $\mathbb{Z}$*-valued cadlag paths.Because the space is discrete, it's automatically a Feller process, and then the strong Markov property is satisfied.The semi-group for a rate* α *Poisson process is*

$$E^x(g(X_t)) = E^0[g(x + X_t)] = \sum_{k=0}^{\infty} g(x+k) e^{-\alpha} \frac{\alpha^k}{k!} \quad for \ x \in \mathbb{Z}$$

References

[1] Timo Seppäläinen. Basics of stochastic analysis. *University of Wisconsin–Madison. Available at the University of Wisconsin-Madison: http://www. math. wisc. edu/˜ seppalai/courses/735/notes. pdf*, 2012.